



Drivers of cyanotoxin and taste-and-odor compound presence within the benthic algae of human-disturbed rivers

Zane Rider^a, Abigail Percich^a, Yasawantha Hiripitiyage^a, Ted D. Harris^b, Belinda S.M. Sturm^a, Alan E. Wilson^c, Erik D. Pollock^d, John R. Beaver^e, Admin Husic^{a,*}

^a Department of Civil, Environmental and Architectural Engineering, University of Kansas, 2150 Learned Hall, Lawrence, KS 66045, United States

^b Kansas Biological Survey, University of Kansas, Lawrence, KS 66045, United States

^c School of Fisheries, Aquaculture, and Aquatic Sciences, Auburn University, Auburn, AL 36849, United States

^d Stable Isotope Laboratory, University of Arkansas, Fayetteville, AR 72701, United States

^e BSA Environmental Services, Beachwood, OH 44122, United States

ARTICLE INFO

Keywords:

Benthic cyanobacteria
Microcystin
Geosmin
Land use
Water quality

ABSTRACT

Freshwater benthic algae form complex mat matrices that can confer ecosystem benefits but also produce harmful cyanotoxins and nuisance taste-and-odor (T&O) compounds. Despite intensive study of the response of pelagic systems to anthropogenic change, the environmental factors controlling toxin presence in benthic mats remain uncertain. Here, we present a unique dataset from a rapidly urbanizing community (Kansas City, USA) that spans environmental, toxicological, taxonomic, and genomic indicators to identify the prevalence of three cyanotoxins (microcystin, anatoxin-a, and saxitoxin) and two T&O compounds (geosmin and 2-methylisoborneol). Thereafter, we construct a random forest model informed by game theory to assess underlying drivers. Microcystin ($11.9 \pm 11.6 \mu\text{g}/\text{m}^2$), a liver toxin linked to animal fatalities, and geosmin ($0.67 \pm 0.67 \mu\text{g}/\text{m}^2$), a costly-to-treat malodorous compound, were the most abundant compounds and were present in 100 % of samples, irrespective of land use or environmental conditions. Anatoxin-a ($8.1 \pm 11.6 \mu\text{g}/\text{m}^2$) and saxitoxin ($0.18 \pm 0.39 \mu\text{g}/\text{m}^2$), while not always detected, showed a systematic tradeoff in their relative importance with season, an observation not previously reported in the literature. Our model indicates that microcystin concentrations were greatest where microcystin-producing genes were present, whereas geosmin concentrations were high in the absence of geosmin-producing genes. Together, these results suggest that benthic mats produce microcystin *in situ* but that geosmin production may occur *ex situ* with its presence in mats attributable to adsorption by organic matter. Our study broadens the awareness of benthic cyanobacteria as a source of harmful and nuisance metabolites and highlights the importance of benthic monitoring for sustaining water quality standards in rivers.

1. Introduction

Harmful algal blooms, and the compounds they produce, are becoming more prevalent across the globe (Wells et al., 2020). While the last half-century of algal research has primarily focused on lake and reservoir systems, lotic (including rivers and streams) and benthic zones are now recognized as active areas of cyanotoxin production and transport (Bouma-Gregson et al., 2018; Graham et al., 2020; Ibelings et al., 2021). Benthic algal mats may include a diverse community structure including green algae, cyanobacteria, and diatoms. The timing and magnitude by which the drivers of benthic productivity, such as flow, light, and nutrients, are delivered to moving systems is highly

influenced by human activity (Busse et al., 2006). In rural basins, application and leaching of fertilizers causes elevated nutrients, while in urban basins nutrient-rich wastewater effluent often constitutes much of baseflow (Zarnaghsh and Husic, 2021). To this end, the relative impact of different land uses on the presence of harmful and nuisance compounds within benthic algae is not well understood.

Benthic cyanobacteria are the portion of the phylum of cyanobacteria that grow attached to substrates in streams, rivers, and lakes. Some cyanobacteria in benthic communities are capable of making secondary metabolites, termed cyanotoxins, such as microcystin (MC), anatoxin-a (ANA), and saxitoxin (SAX) (Kaloudis et al., 2022). High benthic cyanotoxin concentrations are a serious threat and have been associated

* Corresponding author.

E-mail address: ahusic@ku.edu (A. Husic).

<https://doi.org/10.1016/j.watres.2024.121357>

Received 1 December 2023; Received in revised form 18 January 2024; Accepted 21 February 2024

Available online 23 February 2024

0043-1354/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

with pet and livestock deaths (Mez et al., 1997; Wood et al., 2007). These toxins are not regulated in the United States, but the Environmental Protection Agency does have a health advisory of 1.6 µg/L for microcystin in finished drinking water (Wood et al., 2020). In addition to harmful toxins, benthic cyanobacteria can also produce nuisance substances, such as geosmin (GEO) and 2-methylisoborneol (MIB), which impart a noxious taste when consumed, are detectable at very low concentrations (10 ng/L), and require expensive treatment (Dunlap et al., 2015; Watson et al., 2008). The ability of benthic mats to grow and produce the aforementioned substances is a function of climate, light-conditions, flow velocities, nutrient availability, wastewater inputs, among other drivers.

Numerous field, laboratory, and modeling approaches have been developed to enumerate and describe the benthic algal community and its associated metabolites. Rapid field assessments via handheld fluorometric devices have been applied with varying degrees of success for pigment and community composition estimates (Harris and Graham, 2015; Rosero-López et al., 2021). Enzyme linked immunosorbent assays (ELISA) can provide inexpensive and quick laboratory method for estimating cyanotoxin concentrations (Gaget et al., 2017). More involved approaches, such as taxonomic identification and enumeration by microscope can describe the algal and bacterial species in a sample (Echenique-Subiabre et al., 2018). Another breakthrough process is describing the abundance of genes that encode for toxin production through quantitative polymerase chain reaction (qPCR) (Otten et al., 2015). Lastly, statistical regression models can provide predictive capability of cyanotoxin and T&O compound concentrations as well as explain the importance of various drivers (Otten et al., 2016). Random forest models are particularly useful as they can capture non-linear trends in data (De'Ath and Fabricius, 2000), while interpretable machine learning methods derived from game theory, such as Shapley values, can identify the respective contribution of individual variables to overall prediction strength (Park et al., 2022). Application of isolated methods has often been performed in the literature, but a comprehensive assessment of cyanobacterial mats using a coupled field-laboratory-modeling approach is yet lacking.

The objective of this research was to identify the drivers of cyanotoxin and T&O compound production and presence within the benthic algal mats of three human-disturbed rivers that span rural-to-urban land use. We hypothesized that – because urbanization results in higher nutrient concentrations, less riparian canopy, and shallower flow depths – urban rivers would have 1) greater proliferation of benthic algal mats, 2) larger concentrations of cyanotoxins and T&O compounds, and 3) increased density of toxin producing genes within mat matrices. We tested these hypotheses by sampling benthic material from nine river reaches, across a steep land use gradient, and generating a unique dataset spanning physicochemical, isotopic, taxonomic, toxicological, and genomic analyses. Thereafter, we developed predictive random forest models to quantify the relative importance of the environmental factors driving harmful and nuisance compound production within benthic algae mats.

2. Materials and methods

Study site and sample collection: Johnson County, Kansas, USA, has experienced rapid population growth in the previous two decades (up 20 %) compared to the rest of the state (up 4 %). This growth has been accompanied by urbanization of agricultural and forested lands (Fig. 1 and Table S1). Three rivers with similar topography, geology, soils, and climate, but drastically different land use serve as our study sites: Blue River, Mill Creek, and Indian Creek, which we refer to in this manuscript for simplicity as Rural River, Mixed River, and Urban River, respectively. In summer 2020, an initial longitudinal survey was conducted by walking each river in waders to visually identify and locate benthic algal mats. Approximately 10 km per river were surveyed. In total, we identified eighty-three stream reaches with benthic algal mats: 47 in the Urban River, 31 in the Mixed River, and 5 in the Rural River. While the more urban sites have greater mat density, the aim of the survey was not to map every mat in the rivers, but to build an inventory of potential mat locations for intensive data collection in the following summer. The true number and density of mats is likely greater than we report, particularly for the rural system where turbid water complicated

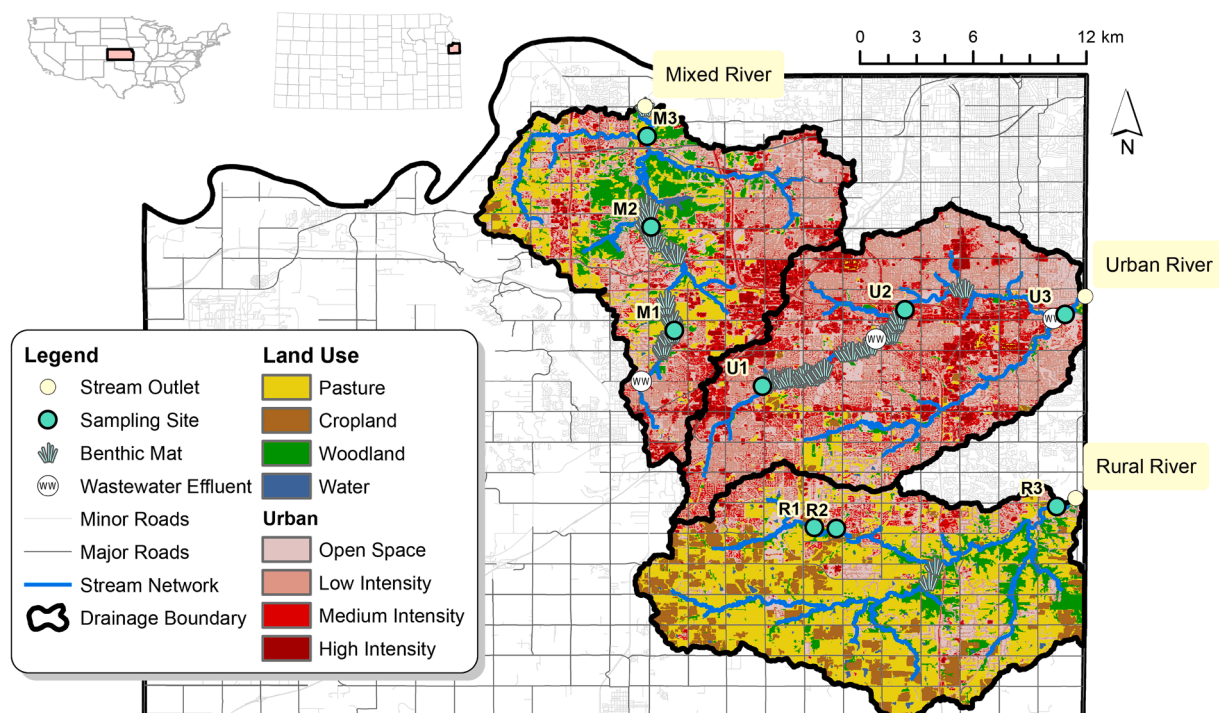


Fig. 1. Map showing the location of Johnson County, Kansas, within the USA. The land use gradient is highlighted as are the surveyed benthic mats and the sampling sites.

rapid visual identification of mats.

A total of nine mats, three per stream, were selected for intensive monitoring in the late summer and early fall of 2021. Where possible, mats were selected such that they evenly spanned the length of a river to capture within-basin heterogeneities (Fig. 1). Over the course of three months, we conducted eight field visits, on a weekly-to-biweekly basis, where we assessed environmental field conditions and collected benthic material for laboratory analysis. At each site, a variety of data were collected, which broadly included *in situ* sensing, water grab samples, and algal mat scrapes. For water analysis, a Horiba U-52 was used to collect temperature, pH, electrical conductivity, turbidity, and dissolved oxygen readings. Rivers were fairly shallow during sampling so near-surface water grab samples were taken for analysis of total suspended solids, dissolved organic carbon, total nitrogen, total phosphorus, nitrate, ammonium, and orthophosphate. Nitrate, ammonium, orthophosphate, and dissolved organic carbon were all filtered in the field (0.45-micron filter). Dissolved organic carbon was additionally preserved with phosphoric acid. For benthic assessment, fifteen cobbles were collected along a 100-m longitudinal transect at each site. A handheld spectrofluorimetric probe (BenthoTorch, bbe Moldaenke) was used on each cobble to get an instantaneous estimate of diatom, cyanobacteria, green, and total algal chlorophyll concentration ($\mu\text{g}/\text{cm}^2$). A bar clamp sampler was used to scrape a known surface area (13.16 cm^2) from each cobble with a clean laboratory spatula. The fifteen cobble scrapings were aggregated into a single opaque bottle and placed immediately on ice in a cooler. A small subsample of algae was placed separately in Lugol's iodine for taxonomic identification and enumeration. In total, 72 benthic algal samples were collected during the field campaign.

Laboratory analyses: Upon return to the lab, the algal samples were freeze-dried for preservation. The dry mass of each bulk algal sample was recorded, the sample was gently homogenized, and aliquots were taken for the subsequent analyses: organic (500 mg), pigment (250 mg), genomic (250 mg), and toxicologic (350 mg). The elemental concentrations and stable isotope ratios of carbon and nitrogen were estimated at the University of Arkansas Stable Isotope Ratio Laboratory using combustion and reduction in an elemental analyzer (EA IsoLink™) coupled to an isotope ratio mass spectrometer (Delta V Plus IRMS). Results were normalized to USGS 41a and USGS 8573 reference material. For photosynthetic pigment analysis at Auburn University, chlorophyll-*a* was extracted in the dark at 4 °C for 24 h using a 90 % ethanol mixture. Chlorophyll-*a* was measured at 664 and 750 nm with a spectrophotometer before and after acidification to correct for phaeophytin readings. Phycocyanin was measured with a fluorometer after extracting each sample in a 50 mM phosphate buffer, also in the dark at 4 °C (Kasinak et al., 2015).

The University of Kansas Center for Microbial Metagenomic Community Analysis (CMMCA) extracted DNA from each sample using a Qiagen DNeasy PowerSoil Pro-DNA Isolation kit, following manufacturer protocol on a QIAcube Connect automated extraction system. Extracted DNA was quantified using the Qubit HS 1X Assay on a Qubit 2 fluorometer. Thereafter, qPCR was performed to quantify genes of Cyanobacteria (16S rRNA) and *Phormidium* (16S rRNA) as well as genes associated with bacterial production of microcystin (*mycE*), anatoxin-*a* (*anaC*), saxitoxin (*sxtA*), and geosmin (*geoA*). Primer names, gene sequences, annealing temperatures, and references to methods are explained in further detail in Rider (2022) and listed in Table S2. At the time of study, consensus gene primers for MIB were not available, thus no genomic analysis was performed for MIB. qPCR reactions were prepared using a Bio-Rad SsoAdvanced Universal SYBR Green Supermix. Thermal cycling was carried out on a Bio-Rad CFX Connect Real-Time PCR Detection System. Data analysis of the qPCR data was performed using Bio-Rad CFX Maestro Version 2.1.

In the University of Kansas Ecohydraulics Laboratory, toxins from the dry algal mass were lysed for assessment of intracellular microcystin, anatoxin-*a*, and saxitoxin concentrations using enzyme-linked

immunosorbent assay (ELISA) kits from Eurofins Abraxis. ELISA detects all congeners of each specific cyanotoxin; thus, toxin concentrations represent all congeners for a given cyanotoxin. To begin, samples were placed in 40 mL of DIDO water and three freeze-thaw plus ultrasonic bath cycles were repeated to lyse intracellular toxins from the algal mass into the overlying water. Anatoxin-*a* and saxitoxin diluent was added after the final lysing step, following manufacturer instructions. The samples were then centrifuged and the resulting supernatant was passed through a 0.45 μm syringe filter. The filtered samples were frozen for preservation and thawed later for batch analysis on a full ELISA tray. Toxin results were fitted to the standard curve determined by the manufacturer's standard solutions. The same lysing procedure was followed for geosmin and MIB assessment, but the method of analysis was solid-phase microextraction and GC-MS at the WaterOne laboratory, the testing center for a local drinking water utility. dry benthic algal mass Geosmin and MIB analysis could not be conducted for sample run six due to machine issues, thus T&O compound concentrations are reported for 64 of the 72 total samples. Extracted intra-cellular metabolite concentrations ($\mu\text{g}/\text{L}$) were normalized to scrape area ($\mu\text{g}/\text{m}^2$) and dry mass of algae ($\mu\text{g}/\text{kg}$) for further comparison.

A subset of 24 samples, equally distributed among sites, was delivered to the BSA Environmental Services Laboratory for taxonomic identification and enumeration. Samples were quantified on a per milliliter basis using the Utermohl method at 800X and 1260X magnification on a Leica DMi1. The count, density, and biovolume for a library of algal genera were recorded. Counts for diatoms, green algae, cyanobacteria, cryptophyta, and euglenophyta were recorded, but only the first three phyla are considered in this analysis. In total, the count, density, and biovolume of 44 algal genera were estimated, spanning green algae, diatoms, and cyanobacteria phyla.

Statistical analyses: For initial exploratory analysis, boxplots of all environmental, toxicological, taxonomic, and genomic data-points were created to quickly assess trends. Given the large quantity of data generated for each collected sample, principal component analysis (PCA) was conducted to reduce the dimensionality of the dataset and reveal underlying patterns among variables. To visually assess separation of variables across land use, confidence ellipses were generated for each grouping using a one standard deviation bound. While visual assessment of PCA can be informative, it is also subjective thus we conduct a permutational multivariate analysis of variance (perMANOVA) test using the Euclidean distance between points as the dissimilarity measure. The perMANOVA test assesses if the centroid and dispersion between groupings is significantly different. Linear regressions were performed on the time-series of anatoxin-*a* and saxitoxin concentrations. We evaluated the significance of each regression ($\alpha = 0.05$) and plotted bounds showing one standard deviation of uncertainty in the regression. All statistical analyses were performed in MATLAB 2021.

Machine learning model: Classification and regression trees are a supervised learning approach for the prediction of discrete or continuous values. As individual trees tend to overfit, we developed a random forest model that utilizes an ensemble of bagged regression trees. We weighed the contribution of individual trees to the overall prediction using each tree's coefficient of determination (R^2). Thereafter, a linear regression was fitted between the observed data and the random forest prediction to assess the model's performance. We constructed models for microcystin and geosmin as their presence was detected in all samples, encompassing a large predictor dataset. To determine the relative importance of each predictor feature to microcystin and geosmin concentration, we calculated the Shapley value that is a concept from cooperative game theory that explains the average expected marginal contribution of a feature to the overall model prediction (Shapley, 1953). The absolute Shapley value is considered to assess the magnitude of a feature's importance, but it is vital to note that features can be positive or negative drivers to model predictions. All machine learning analyses were performed in MATLAB 2021.

3. Results

We located eighty-three benthic mats during the 2020 field surveys: 41 in Urban River, 37 in Mixed River, and 5 in Rural River (Fig. 1). A subset of nine sites, evenly distributed across the rivers, were sampled 8 times each during the summer and fall of 2021. The seventy-two samples collected from the water column and benthos represent a wide range of physicochemical conditions (Table S3). Specific conductance varies considerably from 546 $\mu\text{S}/\text{cm}$ in Rural River to 897 $\mu\text{S}/\text{cm}$ in Urban River, highlighting the influence of wastewater effluent to stream chemistry. The rural system is typically more turbid and has greater suspended solids than the mixed-use and urban systems. Nutrient concentrations are substantially greater, both in terms of magnitude and number of detections, in the mixed and urban rivers compared to the rural system (Table S3). For example, Urban River was more polluted than the Rural River in NO_3^- (4.90 vs 1.26 mg/L), NH_4^+ (0.35 vs 0.07 mg/L), PO_4^{3-} (0.63 vs 0.17 mg/L), TN (5.86 vs 1.53 mg/L), and TP (0.62 vs 0.14 mg/L).

Within the benthos, we see variability in the range of organic, pigment, and fluorometric properties of algal material (Table S4). The thickness of individual mats we observed ranged from sub 1 mm films to 50 cm long filaments (Fig. S2). The overall mean thickness for all sites was 2.8 mm but ranged from less than 1 mm to 4.5 mm depending on the site. Mat morphology included algal filament structures, slimy mat structures, thin films, and macro-algae formations (Fig. S2). The carbon and nitrogen contents of algal scrapes were typically 10 % and 1 %, respectively. The chlorophyll (Chl-*a*) and phycocyanin (PC) measurements indicate no significant patterns across land use. On the other hand, the BenthosTorch fluorometric readings suggest that Mixed and Urban Rivers were more productive than rural rivers; however, the BenthosTorch chlorophyll estimates were only linearly correlated with laboratory results up to a limit of 4 $\mu\text{g}/\text{cm}^2$ ($R^2 = 0.44$; Fig. S3), above which no relationship existed.

Taxonomic results show that across the green algae, diatom, and cyanobacterial taxa, 44 unique genera were observed up to a mean cell density of 3×10^5 cells/mL (Fig. 2). The most common genera of cyanobacteria, which are the primary producers of cyanotoxins, were *Leptolyngbya*, *Phormidium*, and *Eucapsis* with some counts of *Cylindrospermopsis*. By counts, taxonomic enumeration indicated that cyanobacteria comprised 9 to 21 % of the benthic algal community. *Phormidium*, a genus commonly reported to produce cyanotoxins, was detected in 46 % of samples containing at least one cyanobacterial genus and comprised 2 to 6 % of the total community. Cyanobacterial density was slightly greater in Rural Stream. The handheld BenthosTorch provides an estimate of cyanobacterial density, but these estimates were uncorrelated to the laboratory counts ($R^2 = 0.05$; Fig. S3).

Genomic results showed the presence of cyanobacteria and *Phormidium* 16 s rRNA gene sequences in all 72 samples (Table S4). Likewise, *mcyE* and *sxtA* gene sequences were present in all samples. The *anaC* and *geoA* gene sequences were detected less frequently in 38 % and 11 % of samples, respectively. The detection and quantity of cyanobacterial, cyanotoxin, and T&O compound genes – apart from *mcyE* – was greater in the Rural River than the Mixed and Urban Rivers.

Microcystin and geosmin were the most abundant cyanotoxin and T&O compounds, respectively, and were present in 100 % of samples (Table S5), irrespective of land use or environmental conditions. Anatoxin-*a*, saxitoxin, and MIB detections were less-frequent at 67 %, 26 %, and 86 %, respectively. Mean cyanotoxin concentrations of MC, ANA, and SAX were 11.92, 12.11, and 0.67 $\mu\text{g}/\text{m}^2$, respectively. Mean T&O compound concentrations of GEO and MIB were 0.67 and 1.38 $\mu\text{g}/\text{m}^2$, respectively. There were a few significant differences in the concentrations of ANA, GEO, and MIB between pairs of sites, but these differences were not systematically related to land use.

Regarding seasonal effects on cyanotoxins and T&O compound concentrations (Fig. 3), there was a significant increase in MC from late summer into early fall for the Rural ($p = 0.02$) and Mixed ($p = 0.05$)

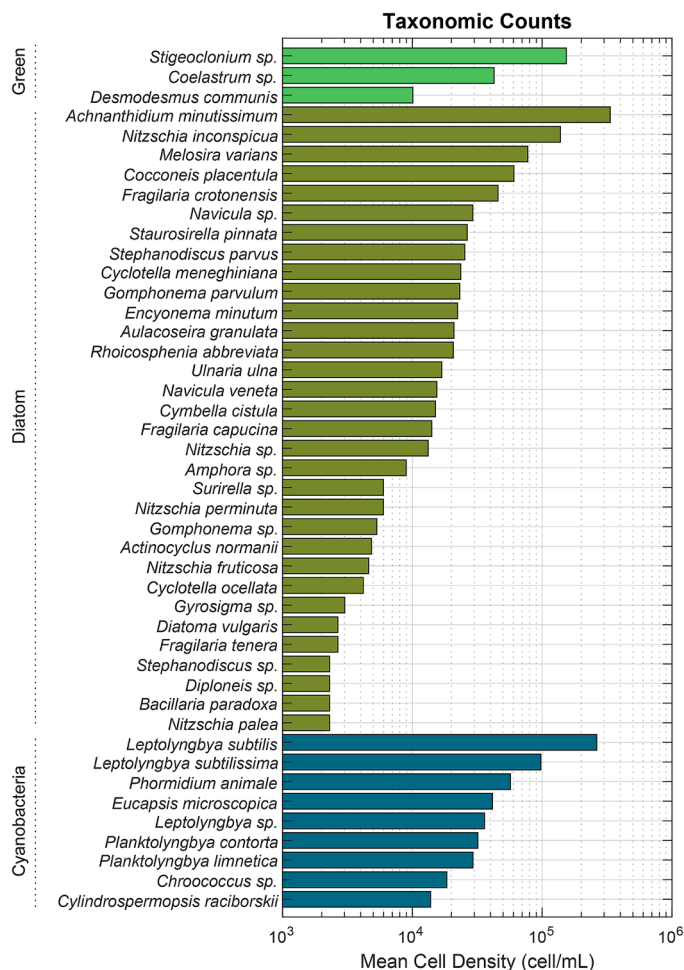


Fig. 2. Taxonomic cell density results of three dominant taxa (green algae, diatoms, and cyanobacteria) for a subset of samples ($n = 24$). Note the logarithmic x-axis.

Rivers, but not for Urban River ($p = 0.66$). Geosmin and MIB concentrations were unaffected by seasonality at any of the sites. The slopes of these trends are broadly similar when the concentrations are normalized by dry mass ($\mu\text{g}/\text{kg}$) rather than scrape area ($\mu\text{g}/\text{m}^2$), but they lose their statistical significance ($p > 0.05$) under dry mass normalization. Both anatoxin-*a* and saxitoxin showed a unique systematic tradeoff in their relative importance with season when normalized by dry weight ($p < 0.05$; Fig. 4), but only the ANA trend remains significant when results are normalized to scrape area.

The random forest model for both microcystin and geosmin adequately predicted observed concentrations (Fig. S6). The fraction of variance explained (R^2) for the training and out-of-bag testing splits, respectively, was 0.75 and 0.57 for microcystin and 0.80 and 0.45 for geosmin. With confidence in the random forest model to accurately predict concentrations, we investigated which of the physicochemical, nutrient, organic, pigment, fluorometric, and genomic predictor variables provided the greatest explanatory strength. Shapley value analysis of microcystin ($\mu\text{g}/\text{cm}^2$) highlights the influence of genomic predictors with *mcyE*, 16S_{Cyano}, and 16S_{Sphor} ranking as the greatest contributors to successful microcystin prediction (Fig. 5). These results suggest that when the algal material lacks the genetic information necessary to encode for microcystin, concentrations of MC will be lower. However, the same is not true for geosmin, whose predictive performance relies primarily on the specific conductance of the overlying water than it does on genetic information within the benthic matrix, although the latter does impart secondary influence.

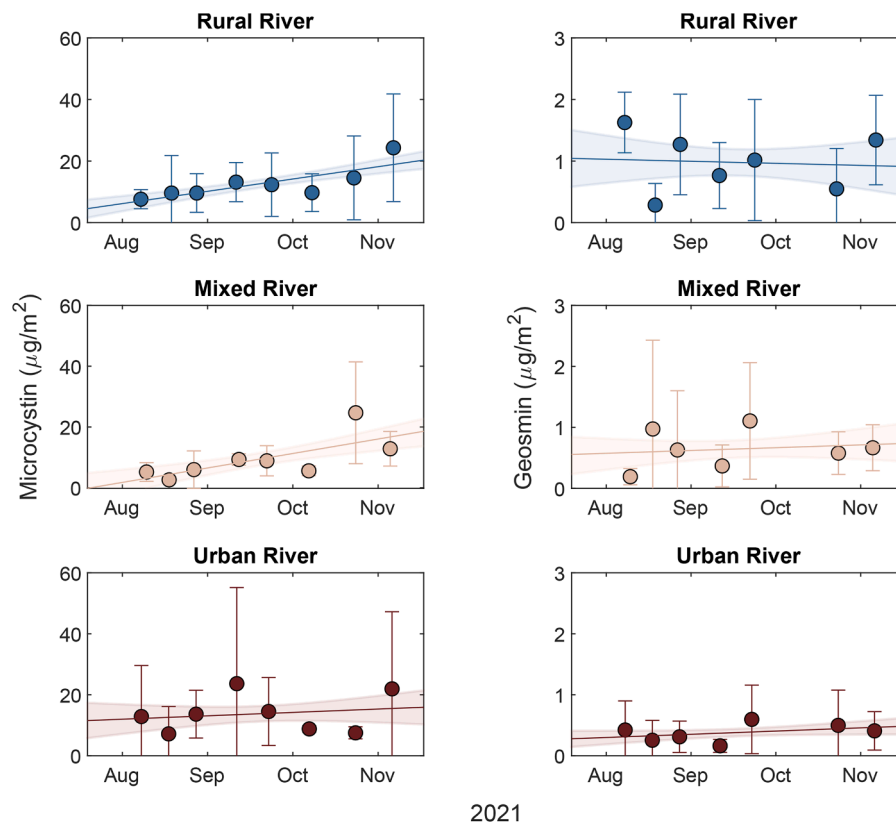


Fig. 3. Microcystin and geosmin concentrations during the study period. Markers represent the mean concentration of all sites in a basin on a sampling date and error bars indicate one standard deviation. Lines-of-best-fit are plotted with shaded bounds representing one standard deviation of uncertainty in each linear regression. Significant regressions include Rural River ($p = 0.02$) and Mixed River ($p = 0.05$) for microcystin.

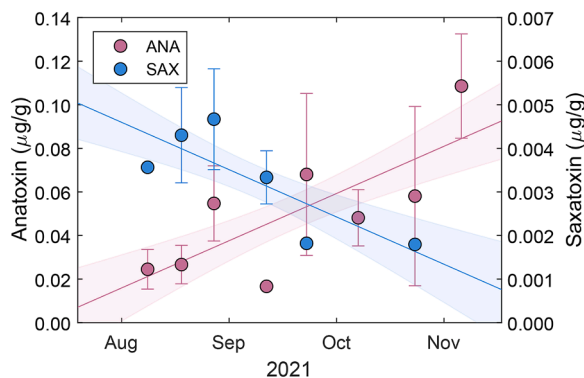


Fig. 4. Anatoxin-a and saxitoxin concentrations during the study period. Markers represent the mean concentration of all sites on a sampling date with error bars indicating one standard deviation. Saxitoxin was not detected in any samples on two dates: early October and November. Lines-of-best-fit are plotted with shaded bounds representing one standard deviation of uncertainty in each linear regression: ANA ($p = 0.03$) and SAX ($p = 0.05$).

4. Discussion

We hypothesized that urban rivers would have (1) greater proliferation of benthic algae mats, (2) larger concentrations of cyanotoxins and T&O compounds, and (3) increased density of toxin producing genes within mat matrices, because urbanization typically causes higher nutrient concentrations, less riparian canopy, and shallower flow depths. Our results lend support to the first point that urban rivers have a more biologically productive benthos, as we noted vastly more mats in the urbanized rivers and higher overall pigment concentrations.

2021

However, our results lead us to reject the latter two points; the benthos of urban rivers do not seem to be more potent in their MC and GEO concentrations, nor do they seem to harbor an environment that is more conducive to toxin-producing genes.

A trend of more frequent harmful and nuisance algal events has been observed within lakes and along coasts (Watson et al., 2015), and these hazards are being noted for inland benthic fluvial systems, like the rivers in Kansas (Paerl et al., 2018). While land use change, and the impacts associated with it, are often cited as drivers of cyanotoxin and T&O compounds production (Beaver et al., 2018; Ibelings et al., 2021), we find presence of these metabolites irrespective of land use type, complementing recent findings (Fetscher et al., 2015; Schulte et al., 2022). Distinct from those studies, we detected GEO and MC in every river and from every sample we analyzed. Though our rural and urban sites are characterized by distinct water chemistry differences ($p < 0.05$ per-MANOVA test; Fig. 6), the toxicological concentrations we aim to understand (MC, ANA, SAX, GEO, and MIB) are not strongly aligned with the principal component that describes land use (PC 2). However, the loadings for the metabolites of interest are, for the most part, oriented perpendicular (along PC 1) to the land-use axis.

Whether rural, mixed, or urban, it could be that benthic algae are saturated in nutrients and thus are not limited (Buley et al., 2021; Scott and Marcarelli, 2012), which could indicate that toxin production is de-coupled from nutrient addition in disturbed systems. Ideally, a forested watershed that was relatively unaffected by rural and urban land use could have offered a control for comparison (Beaver et al., 2018), but humans so heavily impact our region that no such streams exist in close proximity. Thus, the toxicological response of cyanotoxins and T&O compounds is not a simple linear function of land use alone and more complex behavior exists.

Non-linear trends in drivers of microcystin presence in benthic algae were revealed with random forest and Shapley value analysis (Fig. 5).

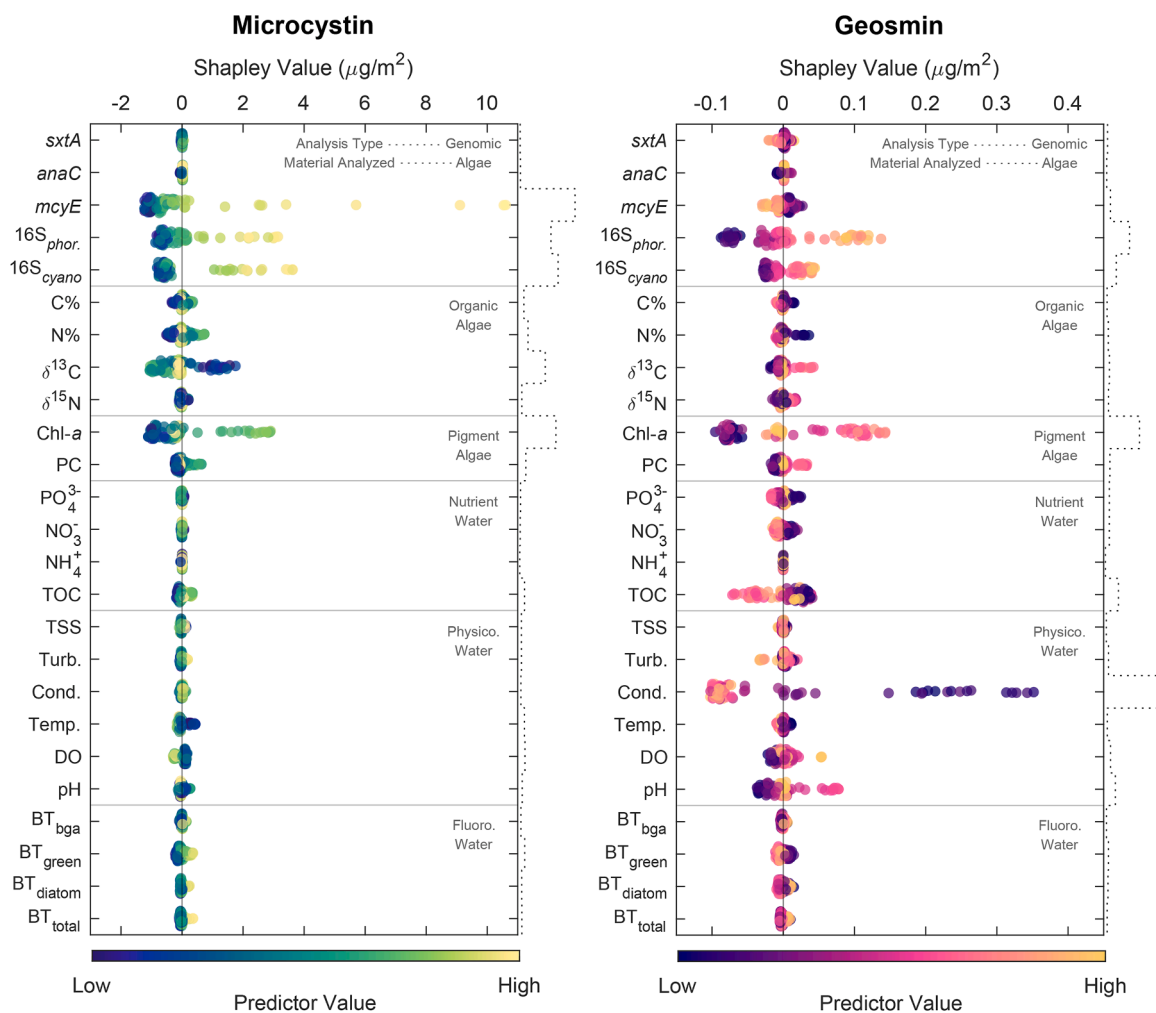


Fig. 5. Shapley values for predictive model features for microcystin (left) and geosmin (right) for six analysis types. Larger (absolute) values indicate a greater importance of a feature to overall prediction. Shapley values for each predictor are color-coded by the predictor's value. The mean absolute Shapley values are shown as a histogram to the right of each subplot to indicate relative importance of predictors.

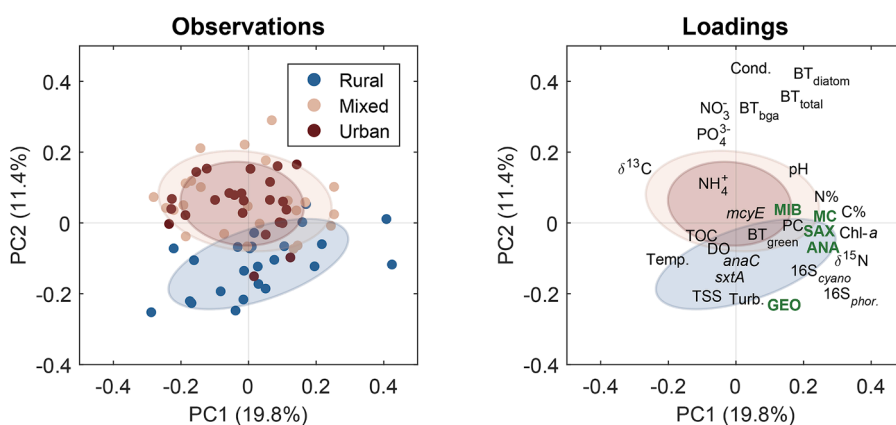


Fig. 6. Principal component analysis showing the first two components. Left panel: data observations ($n = 72$) are color-coded based on sampling location and are encircled by a confidence ellipse. PC2 effectively separates rural and urban land use. Right panel: variable loadings ($n = 30$) show that most toxin and taste-and-odor indicators plot along PC1 with little influence of PC2, indicating little dependence on land use. Toxin and taste-and-odor concentrations (denoted in bold green font) are MC (microcystin), ANA (anatoxin-a), SAX (saxitoxin), GEO (geosmin), and MIB (2-methylisoborneol).

Perhaps unsurprisingly, microcystin concentrations were greatest where microcystin-producing (*mcyE*) and cyanobacterial ($16S_{phor}$ and $16S_{cyano}$) genes were present in relatively large quantities, suggesting *in situ* production of microcystin. The only other variable with substantial

influence on microcystin was Chl-*a*, an indicator of total primary production. Interestingly, the marginal contribution of Chl-*a* to increasing microcystin concentrations was positive up to a value of $25 \mu\text{g}/\text{cm}^2$, after which additional Chl-*a* did not result in greater MC. Studies have

shown that in lotic systems with excess nutrients and plant growth, the relative cyanobacterial domination of benthic assemblages is inhibited in favor of green and diatom algae (Scott and Marcarelli, 2012), which may be the case when Chl-*a* concentrations are at their greatest. Thus, our study suggests that moderately productive benthic mats with genes that encode for microcystin present the greatest potential source of microcystin.

The drivers of geosmin presence in benthic algae were more complex with important variables spanning physicochemical, pigment, and genomic analyses (Fig. 5). Unlike microcystin, geosmin concentrations were large irrespective of the presence of geosmin-producing genes, suggesting *ex situ* production of geosmin. While we detected geosmin in 100 % of benthic scrapes, the geosmin-encoding gene (*geoA*) was found in only 11 % of algal samples. Although the *geoA* primer can identify geosmin synthase for a wide range of cyanobacteria (Kutovaya and Watson, 2014), there are nearly a dozen primers for the cyanobacterial geosmin gene (Devi et al., 2021) and it is possible that another primer could have yielded greater detection. Notwithstanding this, we found that specific conductance was the most influential variable with the largest geosmin concentrations occurring at the smallest conductivity values. This suggests that geosmin presence in the benthos is associated with runoff arrival, which causes reductions in stream conductivity (Zarnaghsh and Husic, 2023). We hypothesize that the runoff-derived geosmin in the water column adsorbs to the organic matter substrate (Guttman and van Rijn, 2009), explaining the presence of geosmin in the benthos even in the absence of genes that encode for it. It is important to note that we measure only intra-cellular MC and GEO and while it would be reasonable to expect that compounds measured from the lysing of intracellular material were produced within the material itself, our results show that potential exists for material produced *ex situ* to adsorb to benthic organic matter. Without a full suite of data analyses, it would be difficult to distinguish allochthonous and autochthonous geosmin production.

We found a novel tradeoff between intracellular anatoxin-*a* and saxitoxin (Fig. 3), which may indicate that anatoxin-*a* occurrence is more likely under relatively cooler water temperatures (November mean: 15.9 °C) compared to saxitoxin (August mean: 27.5 °C) in benthic cyanobacterial mats. In contrast to our results on anatoxin and saxitoxin, we did not find an optimum temperature for microcystin (Fig. S5), which was recently found for pelagic cyanobacterial blooms (Merder et al., 2023). Thus, we show that benthic cyanobacterial mats may have different temperature optimums than lake phytoplankton communities. Although specific reasons why we and others (Merder et al., 2023) have found cyanotoxin temperature optimums are currently unclear, recent research has shown that cyanobacterial toxins may help cyanobacterial cells better acclimate to changing temperature conditions (Martin et al., 2020). Regardless of the cellular mechanisms, our results indicate that utilities may need heightened awareness of this temperature-driven toxin tradeoff to proactively prepare for different toxin types depending on season/water temp. Future studies could examine temperature-toxin relations to better understand how temperature and other physicochemical variables alter toxin production in benthic cyanobacterial taxa.

This work presents several opportunities for future study, particularly with regard to monitoring extracellular cyanotoxins and T&O compounds, comparing data-normalization techniques (scrape area vs. dry weight) and assessment of the handheld fluorometric probe. First, we exclusively look at the intracellular concentrations of cyanotoxins and T&O compounds, which are bound to benthic algae as opposed to extracellular material that advects with water. While metabolism and cell death release intracellular toxins to the water column, it remains unknown if the trends and drivers observed of intracellular toxins match those of extracellular toxins (Romera-García et al., 2021). Thus, this presents a large source of uncertainty in our analysis, and future work should look to close the gap in understanding the interaction between the benthos and water column (Wood et al., 2020). Second, the selection

of normalization method (scrape area vs. dry weight) influences the direct relationships that can be inferred across sites and time (Wood et al., 2020). Area-normalized methods have the advantage of standardizing each measurement to a known scrape area and accounting for intra-site variability. Normalizing by dry weight can skew results if one particular cobble contributes disproportionately to the sampled mass. However, at the same time, this heterogeneity represents potential hot spots of toxin production. Third, regarding the BenthosTorch handheld fluoroprobe, we find that the probe satisfactorily represents low Chl-*a* concentration ($< 4 \mu\text{g}/\text{cm}^2$; $R^2 = 0.44$), but not more productive and dense mats ($> 4 \mu\text{g}/\text{cm}^2$; $R^2 = 0.0$), agreeing with other studies (Harris and Graham, 2015; Kaylor et al., 2018). While some researchers have shown success in estimating cyanobacterial composition with the BenthosTorch (e.g., study (Rosero-López et al., 2021) showed $R^2 = 0.65$), our study shows substantial overestimation of cyanobacteria by the BenthosTorch (38.1 %) compared to laboratory cell counts (11.6 %) and poor prediction in general ($R^2 = 0.05$). More study is needed to resolve the challenges surrounding extracellular concentrations, data-normalization techniques, and fluoroprobe estimates.

Rivers serve as vital municipal, agricultural, and recreational waterways. Our results suggest human activity, whether rural or urban, promotes environments where these harmful and nuisance compound can be present within benthic algae. Ranked in terms of occurrence, we found intracellular MC, GEO, MIB, ANA, and SAX in 100 %, 100 %, 86 %, 67 %, and 26 % of benthic samples, respectively. Literature-reported rates of detection in less-disturbed landscapes tend to be lower (Beaver et al., 2018), highlighting that human activity promotes the presence of these compounds. However, others have noted no – or even negative, although not statistically significant – trends between the urban and agricultural land use near a site and microcystin concentrations in wadeable rivers (Fetscher et al., 2015). We conducted a review of thirteen research articles that report benthic concentrations for one of the five metabolites we list and found results that span orders of magnitude (Table S6). While our results show frequent detection of toxins, the benthic microcystin concentrations we observed were at nuisance levels (max value: 0.67 mg/kg) rather than at lethal doses that have been associated with pet and livestock deaths ($> 25 \text{ mg}/\text{kg}$) (Quiblier et al., 2013; Wood et al., 2010). Nonetheless, geosmin and MIB – although not harmful to humans – are detectable at extremely low levels compared to cyanotoxin guidance values and removing them from water supplies is an expensive process. Likewise, while MC, ANA, and SAX concentrations are not presently dangerous, studies suggest long-term influences of climate change may increase toxin concentrations (Merder et al., 2023), which presents a potential concern. Beyond scientific inquiry, there is also a need for better assessment, risk communication, and mitigation practices for public areas where toxic benthic mats may cause future harm (Wood et al., 2020). Further, while the Environmental Protection Agency and the World Health Organization have developed guidance for metabolite concentrations in drinking water, the safe thresholds for metabolites in recreational waters and in biomass are less certain. The challenge of ensuring adequate water quality in rivers requires cooperation between scientists, regulators, and stakeholders whose expertise spans the domains necessary to adequately identify, treat, and raise awareness of harmful and nuisance algal compounds.

5. Conclusions

Microcystin and geosmin were present in 100 % of samples collected from benthic algae in our study area, irrespective of land use or environmental conditions. On the other hand, anatoxin-*a* and saxitoxin showed a systematic seasonal tradeoff, an observation not previously reported in the literature. Random forest and Shapley value modeling indicated that microcystin concentrations were greatest where microcystin-producing genes were present, whereas geosmin concentrations were high in the absence of geosmin-producing genes. Together, these results suggest that benthic mats produce cyanotoxins *in situ* but

that geosmin production may occur *ex situ* with its presence in mats attributable to adsorption by organic matter. As our system did not have a pristine freshwater stream environment to use as an “undisturbed benchmark”, further analysis into these conclusions is needed. Due to the complex nature in which human-disturbed environments impact freshwater ecosystems, it is suggested that more research not only be done in lab-controlled tests but also in real-world environments. As climate change and urbanization are both projected to increase in the coming decades, the potential safety and environmental impact benthic-sourced cyanotoxins and T&O compounds can have will increase as well. Land managers need to create, test, and adopt benthic cyanobacterial monitoring programs and increase testing for cyanotoxins when blooms occur to reduce animal deaths or to prevent human harm.

CRedit authorship contribution statement

Zane Rider: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Abigail Percich:** Data curation, Investigation. **Yaswantha Hiripitiyage:** Methodology. **Ted D. Harris:** Conceptualization, Funding acquisition, Writing – review & editing. **Belinda S.M. Sturm:** Conceptualization, Funding acquisition, Writing – review & editing. **Alan E. Wilson:** Methodology, Writing – review & editing. **Erik D. Pollock:** Methodology, Writing – review & editing. **John R. Beaver:** Methodology, Writing – review & editing. **Admin Husic:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing.

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.

Data availability

All data used in this study are available in the Supplementary Information. The code used for the random forest model, the Shapley value analysis, and generation of figures can be found at the following Open Science Framework link: <https://doi.org/10.17605/OSF.IO/KBQ9Z>.

Acknowledgment

The authors would like to thank the NSF EPSCoR RII Track-1 grant (#OIA-1656006) for funding a First Award that helped support this project. We also thank the Kansas Water Resources Research Institute 104 Award (#G16AP00054) for supporting this work. We are grateful for the chemistry section of WaterOne for providing geosmin and MIB analysis. We acknowledge the assistance of Aaron Tuschhoff and Devin Gouger for help and assistance with field work. The authors would like to thank two anonymous reviewers whose comments helped to improve the manuscript. The authors declare no conflict of interests.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.watres.2024.121357](https://doi.org/10.1016/j.watres.2024.121357).

References

Beaver, J.R., Tausz, C.E., Scotese, K.C., Pollard, A.I., Mitchell, R.M., 2018. Environmental factors influencing the quantitative distribution of microcystin and common potentially toxigenic cyanobacteria in U.S. lakes and reservoirs. *Harmful Algae* 78, 118–128. <https://doi.org/10.1016/j.hal.2018.08.004>.

- Bouma-Gregson, K., Kudela, R.M., Power, M.E., 2018. Widespread anatoxin-a detection in benthic cyanobacterial mats throughout a river network. *PLoS One* 13, 1–21. <https://doi.org/10.1371/journal.pone.0197669>.
- Buley, R.P., Correia, H.E., Abebe, A., Issa, T.B., Wilson, A.E., 2021. Predicting microcystin occurrence in freshwater lakes and reservoirs: assessing environmental variables. *Int. Waters* 11, 430–444. <https://doi.org/10.1080/20442041.2021.1938491>.
- Busse, L.B., Simpson, J.C., Cooper, S.D., 2006. Relationships among nutrients, algae, and land use in urbanized southern California streams. *Can. J. Fish. Aquat. Sci.* 63, 2621–2638. <https://doi.org/10.1139/F06-146>.
- De'Ath, G., Fabricius, K.E., 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81, 3178–3192. [https://doi.org/10.1890/0012-9658\(2000\)081\[3178:CARTAP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2). *Ecology*.
- Devi, A., Chiu, Y.T., Hsueh, H.T., Lin, T.F., 2021. Quantitative PCR based detection system for cyanobacterial geosmin/2-methylisoborneol (2-MIB) events in drinking water sources: current status and challenges. *Water Res.* 188, 116478. <https://doi.org/10.1016/j.watres.2020.116478>.
- Dunlap, C.R., Sklenar, K.S., Blake, L.J., 2015. A costly endeavor: addressing algae problems in a water supply. *J. Am. Water Works Assoc.* 107, E255–E262. <https://doi.org/10.5942/jawwa.2015.107.0055>.
- Echenique-Subiabre, I., Tenon, M., Humbert, J.F., Quiblier, C., 2018. Spatial and temporal variability in the development and potential toxicity of *Phormidium* biofilms in the Tarn River, France. *Toxins* 10, 1–14. <https://doi.org/10.3390/toxins10100418> (Basel).
- Fetscher, A.E., Howard, M.D.A., Stancheva, R., Kudela, R.M., Stein, E.D., Sutula, M.A., Busse, L.B., Sheath, R.G., 2015. Wadeable streams as widespread sources of benthic cyanotoxins in California, USA. *Harmful Algae* 49, 105–116. <https://doi.org/10.1016/j.hal.2015.09.002>.
- Gaget, V., Humpage, A.R., Huang, Q., Monis, P., Brookes, J.D., 2017. Benthic cyanobacteria: a source of cylindrospermopsin and microcystin in Australian drinking water reservoirs. *Water Res.* 124, 454–464. <https://doi.org/10.1016/j.watres.2017.07.073>.
- Graham, J.L., Dubrovsky, N.M., Foster, G.M., King, L.R., Loftin, K.A., Rosen, B.H., Stelzer, E.A., 2020. Cyanotoxin occurrence in large rivers of the United States. *Int. Waters* 10, 109–117. <https://doi.org/10.1080/20442041.2019.1700749>.
- Guttman, L., van Rijn, J., 2009. 2-Methylisoborneol and geosmin uptake by organic sludge derived from a recirculating aquaculture system. *Water Res.* 43, 474–480. <https://doi.org/10.1016/j.watres.2008.10.018>.
- Harris, T.D., Graham, J.L., 2015. Preliminary evaluation of an in vivo fluorometer to quantify algal periphyton biomass and community composition. *Lake Reserv. Manag.* 31, 127–133. <https://doi.org/10.1080/10402381.2015.1025153>.
- Ibelings, B.W., Kurmayer, R., Azevedo, S.M.F.O., Wood, S.A., Chorus, I., Welker, M., 2021. Understanding the occurrence of cyanobacteria and cyanotoxins. *Toxic Cyanobacteria in Water*. CRC Press, pp. 213–294. <https://doi.org/10.1201/9781003081449-4>.
- Kaloudis, T., Hiskia, A., Triantis, T.M., 2022. Cyanotoxins in bloom: ever-increasing occurrence and global distribution of freshwater cyanotoxins from planktonic and benthic cyanobacteria. *Toxins* 14, 264. <https://doi.org/10.3390/toxins14040264> (Basel).
- Kasinak, J.M.E., Holt, B.M., Chislock, M.F., Wilson, A.E., 2015. Benchtop fluorometry of phycoerythrin as a rapid approach for estimating cyanobacterial biovolume. *J. Plankton Res.* 37, 248–257. <https://doi.org/10.1093/plankt/fbu096>.
- Kaylor, M.J., Argerich, A., White, S.M., VerWey, B.J., Arismendi, I., 2018. A cautionary tale for *in situ* fluorometric measurement of stream chlorophyll a: influences of light and periphyton biomass. *Freshw. Sci.* 37, 287–295. <https://doi.org/10.1086/697239>.
- Kutovaya, O.A., Watson, S.B., 2014. Development and application of a molecular assay to detect and monitor geosmin-producing cyanobacteria and actinomycetes in the Great Lakes. *J. Great Lakes Res.* 40, 404–414. <https://doi.org/10.1016/j.jglr.2014.03.016>.
- Martin, R.M., Moniruzzaman, M., Stark, G.F., Gann, E.R., Derminio, D.S., Wei, B., Hellweg, F.L., Pinto, A., Boyer, G.L., Wilhelm, S.W., 2020. Episodic decrease in temperature increases *mcy* gene transcription and cellular microcystin in continuous cultures of *Microcystis aeruginosa* PCC 7806. *Front. Microbiol.* 11, 1–15. <https://doi.org/10.3389/fmicb.2020.601864>.
- Merder, J., Harris, T., Zhao, G., Stasinopoulos, D.M., Rigby, R.A., Michalak, A.M., 2023. Geographic redistribution of microcystin hotspots in response to climate warming. *Nat. Water* 1, 844–854. <https://doi.org/10.1038/s44221-023-00138-w>.
- Mez, K., Hanselmann, K., Beattie, K., Codd, G., Hauser, B., Naegeli, H., Preisig, H., 1997. Identification of a microcystin in benthic cyanobacteria linked to cattle deaths on alpine pastures in Switzerland. *Eur. J. Phycol.* 32, 111–117. <https://doi.org/10.1080/09670269710001737029>.
- Otten, T.G., Crosswell, J.R., Mackey, S., Dreher, T.W., 2015. Application of molecular tools for microbial source tracking and public health risk assessment of a *Microcystis* bloom traversing 300km of the Klamath River. *Harmful Algae* 46, 71–81. <https://doi.org/10.1016/j.hal.2015.05.007>.
- Otten, T.G., Graham, J.L., Harris, T.D., Dreher, T.W., 2016. Elucidation of taste- and odor-producing bacteria and toxigenic cyanobacteria in a midwestern drinking water supply reservoir by shotgun metagenomic analysis. *Appl. Environ. Microbiol.* 82, 5410–5420. <https://doi.org/10.1128/aem.01334-16>.
- Paerl, H.W., Otten, T.G., Kudela, R., 2018. Mitigating the expansion of harmful algal blooms across the freshwater-to-marine continuum. *Environ. Sci. Technol.* 52, 5519–5529. <https://doi.org/10.1021/acs.est.7b05950>.
- Park, J., Lee, W.H., Kim, K.T., Park, C.Y., Lee, S., Heo, T.Y., 2022. Interpretation of ensemble learning to predict water quality using explainable artificial intelligence. *Sci. Total Environ.* 832, 155070. <https://doi.org/10.1016/j.scitotenv.2022.155070>.

- Quiblier, C., Wood, S., Echenique-Subiabre, I., Heath, M., Villeneuve, A., Humbert, J.F., 2013. A review of current knowledge on toxic benthic freshwater cyanobacteria – Ecology, toxin production and risk management. *Water Res.* 47, 5464–5479. <https://doi.org/10.1016/j.watres.2013.06.042>.
- Romera-García, E., Helmus, R., Ballesteros-Gómez, A., Visser, P.M., 2021. Multi-class determination of intracellular and extracellular cyanotoxins in freshwater samples by ultra-high performance liquid chromatography coupled to high resolution mass spectrometry. *Chemosphere* 274. <https://doi.org/10.1016/j.chemosphere.2021.129770>. *Chemosphere*.
- Rosero-López, D., Walter, M.T., Flecker, A.S., Ontaneda, D.F., Dangles, O., 2021. Standardization of instantaneous fluoroprobe measurements of benthic algal biomass and composition in streams. *Ecol. Indic.* 121, 107185. <https://doi.org/10.1016/j.ecolind.2020.107185>.
- Schulte, N.O., Carlisle, D.M., Spaulding, S.A., 2022. Natural and anthropogenic influences on benthic cyanobacteria in streams of the northeastern United States. *Sci. Total Environ.* 826, 154241. <https://doi.org/10.1016/j.scitotenv.2022.154241>.
- Scott, J.T., Marcarelli, A.M., 2012. Cyanobacteria in freshwater benthic environments. Whitton, B.A. *Ecology of Cyanobacteria II: Their Diversity in Space and Time* 271–289. https://doi.org/10.1007/978-94-007-3855-3_9.
- Shapley, L.S., 1953. A value for n-person games. Kuhn, H.W., Tucker, A.W. *Contributions to the Theory of Games (AM-28)*, Volume II 307–318. <https://doi.org/10.1515/9781400881970-018>.
- Watson, S.B., Ridal, J., Boyer, G.L., 2008. Taste and odour and cyanobacterial toxins: impairment, prediction, and management in the Great Lakes. *Can. J. Fish. Aquat. Sci.* 65, 1779–1796. <https://doi.org/10.1139/F08-084>.
- Watson, S.B., Whitton, B.A., Higgins, S.N., Paerl, H.W., Brooks, B.W., Wehr, J.D., 2015. Harmful Algal Blooms, *Freshwater Algae of North America: Ecology and Classification*. Academic Press, San Diego. <https://doi.org/10.1016/B978-0-12-385876-4.00020-7>.
- Wells, M.L., Karlson, B., Wulff, A., Kudela, R., Trick, C., Asnaghi, V., Berdalet, E., Cochlan, W., Davidson, K., De Rijcke, M., Dutkiewicz, S., Hallegraeff, G., Flynn, K.J., Legrand, C., Paerl, H., Silke, J., Suikkanen, S., Thompson, P., Trainer, V.L., 2020. Future HAB science: directions and challenges in a changing climate. *Harmful Algae* 91. <https://doi.org/10.1016/j.hal.2019.101632>.
- Wood, S.A., Heath, M.W., Holland, P.T., Munday, R., McGregor, G.B., Ryan, K.G., 2010. Identification of a benthic microcystin-producing filamentous cyanobacterium (Oscillatoriales) associated with a dog poisoning in New Zealand. *Toxicon*. 55, 897–903. <https://doi.org/10.1016/j.toxicon.2009.12.019>.
- Wood, S.A., Kelly, L.T., Humbert, K.B.J., Dail, H., Iv, L., Lazorchak, J., Mcallister, T.G., Mcqueen, A., Pokrzywinski, K., Puddick, J., Quiblier, C., Reitz, L.A., Ryan, K.G., Vadeboncoeur, Y., Zastepa, A., Davis, T.W., 2020. Toxic benthic freshwater cyanobacterial proliferations: challenges and solutions for enhancing knowledge and improving monitoring and mitigation. *Freshw. Biol.* 1–19. <https://doi.org/10.1111/fwb.13532>.
- Wood, S.A., Selwood, A.I., Rueckert, A., Holland, P.T., Milne, J.R., Smith, K.F., Smits, B., Watts, L.F., Cary, C.S., 2007. First report of homoanatoxin-a and associated dog neurotoxicosis in New Zealand. *Toxicon*. 50, 292–301. <https://doi.org/10.1016/j.toxicon.2007.03.025>.
- Zarnaghsh, A., Husic, A., 2023. An index for inferring dominant transport pathways of solutes and sediment: assessing land use impacts with high-frequency conductivity and turbidity sensor data. *Sci. Total Environ.* 164931. <https://doi.org/10.1016/j.scitotenv.2023.164931>.
- Zarnaghsh, A., Husic, A., 2021. Degree of anthropogenic land disturbance controls fluvial sediment hysteresis. *Environ. Sci. Technol.* 55, 13737–13748. <https://doi.org/10.1021/acs.est.1c00740>.