Developing Sensor Proxies for “Chemical Cocktails” of Trace Metals in Urban Streams

Carol J. Morel 1,*, Sujay S. Kaushal 1, Maggie L. Tan 1 and Kenneth T. Belt 2

1 Department of Geology & Earth System Science Interdisciplinary Center—College Park, University of Maryland, College Park, MD 20742, USA; skaushal@umd.edu (S.S.K.); maggieltan@gmail.com (M.L.T.)
2 Geography and Environmental Systems, University of Maryland—Baltimore County, Baltimore, MD 21250, USA; belt@umbc.edu

* Correspondence: morel@umd.edu

Received: 8 August 2020; Accepted: 29 September 2020; Published: 14 October 2020

Abstract: Understanding transport mechanisms and temporal patterns in the context of metal concentrations in urban streams is important for developing best management practices and restoration strategies to improve water quality. In some cases, in-situ sensors can be used to estimate unknown concentrations of trace metals or to interpolate between sampling events. Continuous sensor data from the United States Geological Survey were analyzed to determine statistically significant relationships between lead, copper, zinc, cadmium, and mercury with turbidity, specific conductance, dissolved oxygen, and discharge for the Hickey Run, Watts Branch, and Rock Creek watersheds in the Washington, D.C. region. We observed a significant negative linear relationship between concentrations of Cu and dissolved oxygen at Rock Creek ($p < 0.05$). Sometimes, turbidity had significant positive linear relationships with Pb and Hg concentrations. There were negative or positive linear relationships between Pb, Cd, Zn, and Hg and specific conductance. There also appeared to be relationships between watershed areal fluxes of Pb, Cu, Zn, and Cd in streams with turbidity. Watershed monitoring approaches using continuous sensor data have the potential to characterize the frequency, magnitude, and composition of pulses in concentrations and loads of trace metals, which could improve the management and restoration of urban streams.

Keywords: nonpoint source pollution; total maximum daily loads; proxy; surrogate; restoration

1. Introduction

Urbanization is expected to increase over the next century with significant impacts on water quality in streams and rivers across space and time [1–4]. Metals, such as lead (Pb), cadmium (Cd), copper (Cu), zinc (Zn), and mercury (Hg), are commonly found together at elevated concentrations in many urban streams [4–7]. Urbanization has significant impacts on the transport and transformation of novel mixtures of these and other metals, which can form ‘chemical cocktails’ in many watersheds across regional and global scales [2,3,8]. Characterizing changes in concentrations of metals cocktails in urban streams is important for predicting impacts on aquatic life, ecosystem functions, and human health [5,7,9]. Rapid fluctuations in concentrations of metals are sometimes difficult to track in urban streams, which can occur over short time scales in response to hydrologic events [10–12]. In this work, we explore changes in metal concentrations across three urban watersheds varying in size and impervious surface cover, and analyze relationships between trace metals and in-stream continuous sensor data (turbidity, dissolved oxygen, specific conductance, temperature, and discharge). Analyzing concentrations of metals in urban watersheds and relationships between sensor data and multiple metals can provide insights on tracking the transport of complex chemical mixtures in urban streams.
Many metals undergo a range of rapid transport and transformation processes due to aqueous chemical reactions, particularly before, during, and after storm events [13]. Previous studies indicate positive correlations between metals contamination and urban land use; frequent floods and storm events enhance metal mobility and lead to environmental and public health risks [5,7,12,14–18]. Urban areas have a much higher percentage of impervious surfaces, which can allow for metals to be transported to local streams and rivers more efficiently via runoff. For example, urban drainage networks can act as channeled conduits for direct inputs of metals into nearby streams, and the erosional power of stormwater flow can allow for greater overall transport of metals in particulate form [3,19].

Metals can be rapidly transformed by processes including organometallic complexation and sorption to sediments; these transformations could be tracked by sensors [13,20]. In seasonally cold regions, salinization from road salts can affect metal transport by increasing their solubility due to anion complexation and divalent cation competition, which can increase the concentration of total dissolved metals in solution via the ‘salting in’ effect [8,21]. ‘Salting in’ relates to the formation of more soluble and labile compounds via anion complexation. Divalent cation competition displaces metals adsorbed onto sediments, which makes them more mobile [22]. Similarly, some metals may be released into solution from iron and manganese oxyhydroxides in response to changes in redox conditions, and would be sensitive to fluctuations in dissolved oxygen [13,23]. Using sensor proxies to estimate metal concentrations is possible if there are significant statistical relationships between metal concentrations and sensor parameters, such as turbidity, dissolved oxygen, specific conductance, and discharge. Sensors measure continuously and can characterize rapid changes in environmental parameters that could be used to evaluate whether stream maximum contaminant limits have been exceeded.

Our work focused on analyzing relationships between metals (lead, copper, zinc, cadmium, and mercury) concentrations and in-stream sensor data such as turbidity, dissolved oxygen (DO), specific conductance (SC), pH, temperature (T), and discharge (Q). Concentrations of metals were analyzed from discrete sampling events across urban watersheds. We used a database maintained by the United States Geological Survey (USGS) from three urban streams in the Washington DC metropolitan region. Concentrations of metals in streams are regulated for safe drinking water and protection of aquatic life in the U.S. [24]. Analyzing and predicting the high-frequency resolution dynamics of metals in urban streams over time is difficult due to labor intensive sampling and costly geochemical analyses. There is currently a dearth of cost-effective and reliable alternatives.

There is a need to develop high-frequency monitoring approaches to understand hydrochemical processes, even when sensors don’t exist to make direct measurements of chemical parameters [25]. Some technological advancements may allow direct quantification of a few metals such as Cu and Zn in the future, but in most cases indirect sensor proxies will need to be developed [26]. There is evidence that turbidity can be used as a potential proxy for metals in streams [27,28]. There is growing work on development of sensor networks with turbidity, conductivity, dissolved oxygen, nutrients, and other parameters, but more work needs to expand to metals and other parameters in the future [29–31]. This study is unique in that it evaluates relationships between trace metals and multiple in-stream sensor parameters. Predicting frequency and magnitude of water quality exceedances of metals in urban watersheds via the use of continuous sensor data as proxies could improve the monitoring and management of metals in urban watersheds and also identify potential limitations in relying on traditional sampling approaches.

**Site Descriptions**

Continuous sensor data and metals concentrations from three separate watersheds in the Washington, DC, area were analyzed from the USGS National Water Information System. These three watersheds were: Rock Creek at Joyce Road (USGS 01648010), Watts Branch (USGS 01651800), and Hickey Run (USGS 01651770) (Figure 1).
While all located in the DC metropolitan area, the aforementioned watersheds vary in drainage area, percentage of impervious surfaces, proximity to potential point sources (sites with National Pollutant Discharge Elimination System (NPDES) or Municipal Separate Storm Sewer System Permits (MS4)) and land use distribution (Table 1). The Rock Creek watershed and its tributaries, specifically, have set total maximum daily loads (TMDLs) for arsenic, copper, lead, mercury, and zinc [32].

Figure 1. Land use for the Hickey Run (a), Rock Creek (b), and Watts Branch (c) watersheds.
Table 1. Watershed characteristics table for the Rock Creek, Hickey Run and Watts Branch watersheds.

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Drainage Area (Square Miles)</th>
<th>Percent Impervious Surface</th>
<th>Dominant Land Use</th>
<th># of Potential Point Sources within the Watershed</th>
<th>Dominant Metal of Concern</th>
<th>Other Contaminants of Concern</th>
<th>Direction of Effluent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock Creek</td>
<td>63.7</td>
<td>32.4%</td>
<td>Residential (45%)</td>
<td>3</td>
<td>Lead</td>
<td>Fecal coliform bacteria, mercury, zinc, copper, arsenic</td>
<td>Potomac</td>
</tr>
<tr>
<td>Watts Branch</td>
<td>3.36</td>
<td>41.4%</td>
<td>Residential (73.2)</td>
<td>0</td>
<td>Arsenic</td>
<td>Organic pollutants, arsenic, copper, chlordane</td>
<td>Anacostia</td>
</tr>
<tr>
<td>Hickey Run</td>
<td>0.99</td>
<td>52.7%</td>
<td>Parkland (34%)</td>
<td>0</td>
<td>Zinc</td>
<td>Arsenic, organic pollutants, copper</td>
<td>Anacostia</td>
</tr>
</tbody>
</table>

Rock Creek is one of the major streams in the Washington D.C. area. The 33-mile creek, with a drainage area of 63.7 square miles, extends from Laytonsville, Maryland, to the Potomac River in the District of Columbia. In the 1890’s, a large portion of the Rock Creek river valley was designated to become one of the nation’s first national parks [33]. As populations rapidly grew during the 1940s and 1950s, forests and farmlands were converted to suburban developments, which adversely affected water quality as storm sewer systems diverted runoff directly into the creek [33]. The 64 square mile watershed now comprises of 45% residential, 31% grassland, 23% agriculture and recreational, and 2% wetlands [34]. After complaints filed by third party interests in 2009, revisions have been made to the total maximum daily loads to better reflect its current state in water quality. Studies have shown that fecal coliform, mercury, copper, zinc, and lead are major contaminants of concern in this tributary [32].

Watts Branch is a small watershed (3.36 mile$^2$) tributary to the Anacostia River. It begins in Prince George’s County, MD and joins the Anacostia R. in Kenilworth Park. Land use consists of 73.2% residential areas, 11% forested areas, 8.2% open space, 7.3% government lands, and 0.3% wetlands [35]. In the 19th century, a vast majority of the watershed was converted from forests and meadows to croplands [35]. After World War II, deforestation rates increased in order to develop residential and industrial areas [35]. During the 1980s, much of the land became urbanized. As a result of the increased impervious surfaces, total suspended sediment loads significantly increased, calling for a national restoration effort [36]. Along with sediment loads, pollutants like lead, organic pollutants, copper, and arsenic also became major contaminants of concern in the watershed. Sediment loads have been reduced by 33% since the 1990s due to the efforts stemming from the Watts Branch Project [36]. Despite significant reductions in bank erosion, storm water and nutrient runoff, Watts Branch is still listed impaired as of 2018 due to high organic pollutant, metal, and bacteria levels [37]. DDOE is addressing these concerns by conducting ongoing monitoring to ensure water quality improvement [36].

Hickey Run is also a small tributary of the Anacostia River, located within the District of Columbia. The 1 km$^2$ watershed contains 34% parkland, 30% industrial, and 29% residential [38]. It was also deforested as population increased to provide for suburban and residential developments [35]. In the early 1900s, the stream has been a waste site for grease and oil discharge [39]. Water
quality became a concern in the 1930's when the stream caught on fire after decades of pollution. Currently, it is one of the most altered watersheds in the Anacostia River, with 41% of the watershed covered by impervious surfaces [38]. Watts branch water quality is restoration challenge for the District Department of the Environment as they try to reduce stormwater runoff and hydrocarbon pollution [38]. Major pollutants of concern include organic pollutants, copper, zinc, and arsenic [39]. Restoration efforts are focused on stormwater management, revising total maximum daily load standards, wetland creation, and providing incentives for homeowners to get involved with storm water management [40]. As a result, there has been an increase in landscape enhancements, such as rain barrels, rain gardens, and shade trees designed to reduce stormwater runoff [40].

2. Materials and Methods

2.1. Discrete Sampling for Metals by the USGS

In 2013, the USGS started a project in collaboration with the EPA and DC’s Department of Energy and the Environment to enhance Non-Tidal Network monitoring by adding analytes such as trace metals and mercury at poorly represented small, urban catchments in the DC area. Both storms and fixed frequency samples are collected on a monthly basis [11]. Water samples are collected and analyzed for dissolved metals. At all three sites, monthly sampling for metals is conducted for discharges ranging from baseflow to stormflow, including the intervals during and after storms; please see Supplementary Materials.

Water samples for all metal analyses (except for mercury) are filtered through a 0.45 micrometer pore-size disposable capsule filters in the field [41]. All water samples are acidified in the field by the USGS [41]. For major and minor cations and trace elements, nitric acid is added until the pH of the sample was below 2. Two milliliters of 6N ultrapure HCl are added to samples designated for mercury analysis [41]. Water samples are then analyzed on an Inductively Coupled Plasma Mass Spectrometer with reporting limits of 1 µg/L [42]. Data collected from March 2014 through May 2018 were used. Data points that were known to be less than the reported value were assumed to be half the value of the detection limit (i.e., midway between the detection limit and 0). Due to assumptions with samples near the detection limits, all statistical relationships derived from these data used only data points above the detection limits.

2.2. Stream Water Quality Sensors

The three selected USGS gage stations are equipped with sensors to continuously measure water quality parameters such as temperature, specific conductance, DO, pH and turbidity. Data from these sensors is made publicly available on the USGS National Water Information site. Temperature is measured with a thermistor that can read within 0.1 degree Celsius of the actual temperature of the stream [43]. Specific conductance is measured as micro-siemens per centimeter at 25 degrees Celsius using contact sensors with electrodes that measure in the range of 0–2000 µS/cm [43]. Dissolved oxygen is measured using the amperometric method. The amperometric method uses a temperature-compensated polarographic membrane-type sensor. Continuous water quality pH sensors employ the electrometric pH-measurement method, which uses a hydrogen-ion electrode [43]. This electrode can accurately measure pH within 0.2 pH units. Submersible turbidity sensors used by the USGS emit a light beam from a light-emitting diode into the water and measure the light that scatters or is absorbed [43]. Turbidity sensors typically have a range of 0–1000 nephelometric turbidity units (NTU) with an accuracy of 2 NTU or within 5% [43].

2.3. Statistical Analyses of Relationships between Water Quality Parameters and Metals

A variety of statistical tests were used to compare metal concentrations among the sites and to develop statistical relationships between metal concentrations and water quality parameters. These methods included: (a) box and whisker plots of normalized data, (b) multiple linear regression, to compare concentrations among the sites, and (c) regression relationships between metal concentrations and environmental parameters.
Box and whisker plots were generated to compare concentrations of trace metals across sites and analysis of variance tests (ANOVAs) were performed to check for statistical differences. Non-metric Multidimensional Scaling analysis was also conducted to spatially assess similarities and differences between sites. Then, multiple linear regression (in the statistical program R) was used to assess the contribution of each predictor variable (continuous sensor data) to the response variable (metals concentrations) for each site. All parameters available for each site were used to define the model. Data were standardized and checked for normality using the Shapiro–Wilk normality test.

Multicollinearity of the multiple regression models was assessed using the variance inflation factor (VIF) function; this function was repeated until all values were below 10 (multicollinearity is highly suggested above 10). After normality and collinearity were checked, the model was run through the ‘glmulti’ function that computes the Akaike information criterion (AIC) for several models and finds the best model to fit the data. The ‘glmulti’ function also corrects for small sample sizes. The AIC does not choose a model based on absolute quality, but will choose the best model available. The delta score for the AIC model is the difference in AIC between the best model and each other model in the set. According to Burnham and Anderson (2004), a delta score of less than or equal to 2 gives substantial evidence that the model is the best model for that dataset [44]. In addition, single linear regression models were developed for predicting concentrations of individual metals to assess the ability of single sensors to predict metals concentrations.

3. Results

3.1. Comparing Concentrations and Mixtures of Trace Metals among Sites

Metal concentrations for copper were plotted against dissolved oxygen and a linear correlation plotted to ascertain the relationship between the in-stream sensor parameter and metal (Figure 2).

![Figure 2](https://example.com/figure2.png)

**Figure 2.** Relationship between copper and dissolved oxygen at the Rock Creek watershed ($R^2 = 0.14$).

Characteristics of metals concentrations and mixtures were analyzed across sites. Metals concentrations varied by several orders of magnitude at each site (Figure 3).
Figure 3. Concentrations of each metal (a-e) for the Rock Creek, Hickey Run, and Watts Branch watersheds.

For example, copper concentrations ranged from 0.25 to 2934 ppb in Hickey Run over the sampling period. Concentrations of lead peaked at 27, 290, and 1134 ppb at Rock Creek, Watts Branch, and Hickey Run, respectively.

Concentrations of all metals, with the exception of cadmium, appeared to increase with increasing impervious surface cover (Figure 3). No metal exceeds its maximum or secondary maximum contaminant limit as regulated by the U.S. Environmental Protection Agency.

ANOVAs were used to test for statistical differences among mean concentrations of each metal at each site (Table 2).
Table 2. ANOVA results for the Rock Creek, Hickey Run, and Watts Branch watersheds where DF is degrees of freedom and F is the F-statistic.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>F</th>
<th>p Value</th>
<th>F Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cu</td>
<td>98</td>
<td>0.156</td>
<td>0.856</td>
<td>3.091</td>
</tr>
<tr>
<td>Cd</td>
<td>185</td>
<td>7.678</td>
<td>0.001</td>
<td>3.045</td>
</tr>
<tr>
<td>Hg</td>
<td>52</td>
<td>0.516</td>
<td>0.600</td>
<td>3.183</td>
</tr>
<tr>
<td>Pb</td>
<td>136</td>
<td>5.462</td>
<td>0.005</td>
<td>3.064</td>
</tr>
<tr>
<td>Zn</td>
<td>149</td>
<td>17.289</td>
<td>0.000</td>
<td>3.058</td>
</tr>
</tbody>
</table>

The results indicate a significant difference in mean concentrations for Cu and Hg across the three sites (F was less than F critical). No significant differences were found for concentrations of Pb, Zn and Cd across the three sites (F was greater than F critical). For instance, the calculated F statistic for Pb was 5.462 and F critical was 3.064. The p-value was greater than 0.5 for Cu and Hg.

The differences in metals concentrations were visualized using non-metric multidimensional scaling (NMDS) (Figure 4).

Figure 4. Non metric multidimensional scaling plot for Rock Creek, Hickey Run and Watts Branch watersheds.

NMDS is a rank-based approach where objects more similar to each other are ordinated closer together [37]. The two small watersheds with larger impervious surface cover, Watts Branch and Hickey Run, were more similar to each other than to Rock Creek, which had more forested parkland (Figure 4). Watts Branch and Hickey Run were also more similar in mean metals concentrations for Cu, Cd, Pb and Zn (Figure 3 and Table 2).

3.2. Relationships between Metals Concentrations and Continuous Sensor Data

Each metal, Cd, Cu, Zn, Pb, and Hg, for Rock Creek, Watts Branch and Hickey Run was analyzed against available continuous sensor parameters such as turbidity, temperature (T), specific conductance (SC), pH, dissolved oxygen (DO), and instantaneous discharge (Q) to assess the contribution of each predictor variable to the response variable. Relationships between metal fluxes and turbidity, specific conductance, and instantaneous discharge were plotted in Figure 5.
Figure 5. Relationships between metal fluxes and turbidity (a), specific conductance (b), and instantaneous discharge (c) for the Rock Creek watershed.

Examples of relationships between metals concentrations and fluxes and turbidity, specific conductance and instantaneous discharge are plotted in Figure 5. Instantaneous discharge has a strong positive relationship with most metal fluxes (Figure 5c). Turbidity also appears to have a positive relationship with fluxes of the aforementioned metals (Figure 5a). Specific conductance appears to show a weak negative correlation with concentrations of all metals at all sites (Figure 5b),
although this relationship is bifurcated as, at higher ranges of specific conductance, it appears that metal fluxes are increasing. However, this relationship was not statistically significant via a t-test for Cd and Zn for Rock Creek.

3.3. Evaluating Proxies for Metals Concentrations Using Multiple Sensors

We only consider and report statistical relationships where $R^2$ values are greater than 0.70 as robust statistical models for predicting metals concentrations using multiple linear regression analyses and AIC (Table 3).

Table 3. Multiple linear regression models chosen with Akaike’s Information Criterion predicting metals concentrations for Rock Creek, Watts Branch and Hickey Run watersheds. Rank is the model rank relative to all possible models, Wi is the weighting evidence or the probability of that model being the best model, and the adjusted $R^2$ is the correlation coefficient adjusted for the addition of parameters. Parameters in bold were found to be statistically significant ($p < 0.05$).

<table>
<thead>
<tr>
<th>Site</th>
<th>Metal</th>
<th>Model</th>
<th>Rank</th>
<th>Wi</th>
<th>ΔAIC C</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock Creek</td>
<td>Hg</td>
<td>Hg - 0 + 0.2654T + -0.7443Q + -0.3343SC + 1.303TURBIDITY</td>
<td>1</td>
<td>0.3</td>
<td>0</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hg - 0 + -0.9159Q + -0.4187SC + 1.271TURBIDITY</td>
<td>2</td>
<td>0.1</td>
<td>5</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>Pb</td>
<td>Pb - 0 + 0.1999Q + -0.4503SC + 0.4118TURBIDITY</td>
<td>1</td>
<td>0.2</td>
<td>8</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pb - 0 + -0.4662SC + 0.5661TURBIDITY</td>
<td>2</td>
<td>0.2</td>
<td>5</td>
<td>0.8</td>
</tr>
<tr>
<td>Watts Branch</td>
<td>Hg</td>
<td>Hg - 0 + -4.739 × 10^4Q + 1.350TURBIDITY</td>
<td>1</td>
<td>0.7</td>
<td>8</td>
<td>0.9</td>
</tr>
<tr>
<td>Hickey Run</td>
<td>Cd</td>
<td>Cd - 0 + 8.372 × 10^4SC + 1.878 × 10^4TURBIDITY</td>
<td>1</td>
<td>0.3</td>
<td>9</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cd - 0 + 1.261 × 10^4T + 8.860 × 10^4SC + 2.401 × 10^4TURBIDITY</td>
<td>2</td>
<td>0.2</td>
<td>4</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>Hg</td>
<td>Hg - 0 + -4.093 × 10^4SC + 5.708 × 10^4TURBIDITY</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Temperature, instantaneous discharge, specific conductance, and turbidity were all significant predictor variables. The highest contribution was from turbidity, which showed a positive relationship with mercury. The best lead model for Rock Creek according to the AIC explained 89% of the variability of lead and had an evidence weight of 28%. Specific conductance and turbidity were significant continuous variables with the highest contribution from specific conductance. The model indicated a positive relationship between specific conductance and predicted lead.

For the Watts Branch site, the best mercury model using multiple linear regression and AIC had a correlation coefficient ($R^2$) of 0.93 and an evidence weight of 78% (Table 3). Instantaneous discharge and turbidity were the two significant continuous variables with the highest contribution (highest predictor variable coefficient) from turbidity. The best mercury model indicated a positive relationship between turbidity and unfiltered mercury concentrations.

The best multiple linear regression model for cadmium for Hickey Run explained 74% of the variability in cadmium concentrations and had an AIC evidence weight of 39% (Table 3). Turbidity and specific conductance were the only significant continuous variables with the highest contribution from specific conductance. The multiple linear regression model indicated a positive relationship between specific conductance and cadmium and between turbidity and cadmium. For mercury, the best multiple linear regression model had a correlation coefficient of 0.74 and an evidence weight of 0.51. Turbidity and specific conductance were also the only significant continuous variables with the highest contribution from turbidity. The best multiple linear regression model for mercury indicated a positive relationship between turbidity and specific conductance.
3.4. Evaluating Proxies for Metals using Individual Sensors

For comparative purposes, the most robust multiple linear regression model as indicated by the $R^2$ for each site was tested against a data set of metals not used to make the model (i.e., the most recent data collected post May 2018). Using the multiple-sensor approach resulted in predicted values that varied significantly from the actual values, despite the high correlation coefficients of the models. This was true for all models. A single sensor approach was also then analyzed to assess the predictability of metals concentrations and was proven to be more effective. The simple linear regression models, along with the correlation coefficient and $p$-values of the models are displayed in Table 4. Only relationships with a correlation coefficient of 0.70 or higher are displayed.

Table 4. Simple Linear Regression models for the Rock Creek, Hickey Run, and Watts Branch watersheds.

<table>
<thead>
<tr>
<th>Site</th>
<th>Metal</th>
<th>Model</th>
<th>$R^2$</th>
<th>$p$-Value of $x$ Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock Creek</td>
<td>Pb</td>
<td>$[\text{Pb}] - 0.0009 \text{SC} + 0.5438$</td>
<td>0.71</td>
<td>$2.88 \times 10^{-8}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$[\text{Pb}] - 0.0016 \text{Turbidity} + 0.1398$</td>
<td>0.77</td>
<td>$1.73 \times 10^{-9}$</td>
</tr>
<tr>
<td>Watts Branch</td>
<td>Hg</td>
<td>$[\text{Hg}] - 0.2084 \text{Turbidity} - 0.0295$</td>
<td>0.89</td>
<td>$5.35 \times 10^{-7}$</td>
</tr>
<tr>
<td>Hickey Run</td>
<td>Cd</td>
<td>$[\text{Cd}] - 3 \times 10^{-5} \text{SC} + 0.0227$</td>
<td>0.72</td>
<td>$1.81 \times 10^{-11}$</td>
</tr>
</tbody>
</table>

Figure 6 graphically shows the relationship between specific conductance and lead at the Rock Creek watershed.

![Relationship between Specific Conductance and $[\text{Pb}]$ at Rock Creek](image)

**Figure 6.** Relationship between specific conductance and lead concentrations at the Rock Creek watershed ($R^2 = 0.71$).

The resulting predicted values’ percent differences from actual values were within 40% for some mid-range metals concentrations for the linear regression model relating specific conductance and Pb concentrations (Table 5).
Table 5. Percent differences below 40% between predicted and actual Pb concentrations for the linear regression model displayed in Figure 6

<table>
<thead>
<tr>
<th>[Pb] (µg/L)</th>
<th>Predicted [Pb] (µg/L)</th>
<th>Percent Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.29</td>
<td>0.34</td>
<td>15.04</td>
</tr>
<tr>
<td>0.13</td>
<td>0.16</td>
<td>17.74</td>
</tr>
<tr>
<td>0.40</td>
<td>0.38</td>
<td>6.06</td>
</tr>
<tr>
<td>0.22</td>
<td>0.27</td>
<td>15.61</td>
</tr>
<tr>
<td>0.13</td>
<td>0.15</td>
<td>19.48</td>
</tr>
<tr>
<td>0.39</td>
<td>0.41</td>
<td>4.05</td>
</tr>
<tr>
<td>0.51</td>
<td>0.43</td>
<td>18.12</td>
</tr>
<tr>
<td>0.30</td>
<td>0.30</td>
<td>1.53</td>
</tr>
<tr>
<td>0.36</td>
<td>0.38</td>
<td>3.15</td>
</tr>
<tr>
<td>0.51</td>
<td>0.38</td>
<td>28.90</td>
</tr>
<tr>
<td>0.11</td>
<td>0.12</td>
<td>12.29</td>
</tr>
<tr>
<td>0.52</td>
<td>0.46</td>
<td>11.46</td>
</tr>
<tr>
<td>0.49</td>
<td>0.43</td>
<td>14.82</td>
</tr>
<tr>
<td>0.25</td>
<td>0.22</td>
<td>10.03</td>
</tr>
</tbody>
</table>

4. Discussion

As expected, concentrations of trace metals increased with impervious surface cover and land development (e.g., concentrations of mixtures in metals were typically greater in smaller watersheds with greater impervious surface cover). Interestingly, we observed strong relationships between concentrations of copper and dissolved oxygen and between the trace metals (Cu, Pb, Zn, Cd, and Hg) and turbidity. There was also a sharp increase in some metals during winter months when road salt use and specific conductance increased (Cu, Pb, Zn, and Cd). These observations prompted us to explore whether concentrations of metals such as lead, mercury, cadmium, copper, and zinc could be influenced by sensor variables such as temperature, discharge, pH, specific conductance, turbidity, and dissolved oxygen in order to evaluate water quality.

The output from multiple linear regression analyses suggests that none of the studied metals could be robustly estimated from multiple continuous variables consistently across all sites in a universal model. The metals, models, and amounts of variation explained by a multiple linear regression model varied considerably among sites. In some cases, data from individual sensors could be used to predict metal concentrations. The usefulness of different environmental variables in proxy models of metal transport are explored below in addition to the implications of tracking chemical cocktails on urban water quality.

Across all sites, there were linear relationships between turbidity and Hg and turbidity and Pb suggesting the potential for universal models (Supplementary Figures S6 and S7), but these linear models explained less of the variation than site specific models. Our study sites differed in concentrations and composition of metals based on our analysis and results. Urban watersheds may be more complex to model due to altered hydrology, nonpoint sources, differences hydrologic connectivity and impervious surfaces, traffic density, road layout, etc. Universal models may be more applicable with a larger number of urban watersheds and greater number of observations.

Previous work has used linear regression approaches for developing water quality proxies such as relationships between turbidity and phosphorus, nitrate, and total suspended solids fluorescence as a proxy for DOC and BOD [29,31,45,46]. In addition, other work has shown that linear relationships can be used to develop proxies of multiple chemical cocktails of metals mostly with turbidity [27,28]. However, there was a gap in the literature for work on proxies of metals using a high-frequency sensor framework suggesting the novelty and importance of this study. We acknowledge that there can also be nonlinear relationships, especially between specific conductance and metals, which requires further investigation. For example, further work could employ machine learning approaches for predictions of metals concentrations [47]. Further, more work needs to capture potentially large changes in concentrations during storms [48]. The advantage of this study is that
sampling was conducted across a wide range of streamflow conditions including targeted storm sampling for proxies (Supplementary Figure S1).

Concentrations of major ions increase with increased impervious surface cover and hydrologic connectivity in these and other nearby watersheds; we do not present data on major ions here so we can focus on trace metals, but this detailed information can be found elsewhere in our previous papers [2,3,8,49]

4.1. Variations in Metals Concentrations and Chemical Cocktails across Sites

A factor contributing to variations in concentrations of metals across sites could be location and proximity to urban pollution sources; Watts Branch and Hickey Run gage stations are located in more urban settings with more anthropogenic sources of heavy metals. The Watts Branch gage station, in particular, is located in close proximity to a roadway and a railway. Even though parkland surrounds much of the Hickey Run gage station, it is located close to a major roadway. The Rock Creek at Joyce Road gage location, in comparison, is surrounded by acres of protected parkland. Concentrations of metals could therefore vary across sites as some are located closer to anthropogenic sources of metals (e.g., major roadways).

Interestingly, we found that there were no significant differences in mean metals concentrations across sites such as Cd, Pb, and Zn, but there were significant differences in mean metals concentrations for As Hg and Cu using ANOVA (Table 2). Thus, the chemical mixtures and metal concentrations varied across sites. Metals related to automobiles may dominate the chemical cocktail in watersheds with greater hydrologic connectivity [19]. Our results from NMDS showed that not only do the concentrations of metals vary across the gradient of hydrologic connectivity, but so too can the mixtures of metals.

4.2. Comparing Metals Concentrations with Environmental Data from Multiple vs Individual Sensors

There were some discrepancies in the expected relationship between the continuous variables and modeled relationships. For instance, specific conductance generally has a negative relationship with metals but in some models a positive relationship is shown. This could be a result of small sample size or changes in the relationship across seasons due to mobilization in winter from road salts. In addition, there is also some variability in the significant variables. In contrast, the individual sensor approach was able to predict some metals within reasonable ranges such as Pb and Zn at Rock Creek. This single sensor approach to developing proxies for metals concentrations is an alternative to the multiple-sensor predictive models. In some cases, a single sensor approach may be simpler to develop and thus more practical and adaptable across watersheds. The influence of different sensor parameters on metal concentrations is discussed below.

4.3. Discharge Influences Metals Concentrations across Urban Streams

The concentrations of metals in urban watersheds remains higher than reference watersheds, with the most common contributing factor being stormwater runoff [5,7,12,14–17]. A major reason why urbanized watersheds have a higher concentration of metals contamination is that urban areas have a much higher percentage of impervious surfaces that allow for contaminants to be transported to local streams and rivers via urban runoff. Metals from a variety of anthropogenic sources accumulate on these impervious surfaces and are transported directly to nearby streams during a rain event. Since impervious surfaces do not allow for the rain to percolate into the subsurface and because storm water systems act as channeled conduits for direct input into nearby streams, the erosional power of the storm water flow allows for greater transport of metals.

In this study, samples for metals analyses were conducted across a range in streamflow conditions, including storms, which contributed to the large variability in concentrations at specific sites. During the “first flush,” most of the metals that accumulate on impervious surfaces during a dry period are washed away in the first wave of runoff during a new rainfall [15,50]. The more mobile these metals and the more concentrated they become in local waters, the higher the chance of
exposure and thus the higher the hazard or risk of the contaminant [50]. Therefore, increasingly more frequent storm events that flush metals contaminants into local waters constitute a primary pulsed stressor that leads to a chronic issue of contamination in urban watersheds [51].

4.4. Turbidity Influences Concentrations of Metals across Urban Streams

We found that turbidity played a significant role in influencing the concentration of some metals according to the simple linear regression modelling. Many of the factors affecting discharge in urban stream environments also impact turbidity. Combined sewer and storm overflows can also contribute to the amount of suspended particulate matter or organic matter that contribute to overall turbidity in urban streams. Previous work has shown a significant and reliable relationship between total suspended solids (TSS) and turbidity in streams and rivers [43,52]. These solids will often adsorb as they contain functional surface groups that are able to complex with metals [53]. Turbidity is strongly influenced by the properties of transported sediment, which varies in shape size and mineral composition [54]. Desorption of weakly sorbed metals and transitions from sediment brought in after a storm to the aqueous dissolved state could explain elevated concentrations of dissolved metals with turbidity, even though one would expect a lower concentration of dissolved metals with an increase in turbidity and thus availability of particulate sorption sites. Indeed, Bourg 1987’s conceptual model of adsorption and turbidity show that with increased turbidity, the percentage of adsorbed metals like Cu and Zn increases as a result of simple Le Chatelier’s principle [55]. Affinity for the solid phase also appears to increase with increasing organic matter content [56]. The suspended particulates associated with turbidity in Rock Creek, Hickey Run, and Watts Branch could be courser material low in organic matter.

4.5. Specific Conductance Influenced Metals Concentrations with Differing Seasonal Patterns

There was also an interesting seasonal increase in all metals except Hg during winter months when specific conductance simultaneously peaked, which suggested mobilization of metals associated with road salts [8]. Recently, it has been recognized that salt pollution in urban watersheds can contribute to Freshwater Salinization Syndrome, which is the suite of interrelated water quality impacts associated with salinization such as ion exchange and metals and nutrient mobilization [8,57,58]. As the salinity of a solution increases, the solubility of a non-polar or weakly polar organic compound decreases because the organic compound is now competing with ionic species that are much easier for the water to solubilize. The salting-in relationship can also be defined with the Setschenow equation, meaning the solubility of the compound is dependent on the concentration of the salt or salts [59]. The salting-in effect is the increase in solubility of compounds in solution. In the case of metals, it is most related to the formation of more soluble complexes through anion-complexation [22]. Chlorocomplexation with cadmium, for example, decreases the activity of Cd2+ adsorbed in the sediment and increases the solubility and thus mobility of the heavy metal [60]. Cd-chlorocomplexes are also highly stable. Divalent cations from salts such as Ca2+ and Mg2+ can also work to compete with metals like Zn and Cd adsorbed to sediment, mobilizing them in solution [22]. These two processes were found to control a large variance in reversible cadmium portioning between suspended sediment and solution, accounting for a large variance in cadmium mobility under controlled laboratory conditions meant to simulate natural conditions [61].

4.6. Dissolved Oxygen Weakly Influenced Metals Concentrations but Showed Relationships with Cu

Dissolved oxygen was significantly related to Cu concentrations when analyzing relationships between concentrations and individual DO sensor data at Rock Creek according to a regression analysis (p < 0.05). However, dissolved oxygen was a weak predictor and was taken out of the Rock Creek model due to multicollinearity as reported by the vif function. While the trends in the data show a weak negative relationship between dissolved oxygen and Cu, dissolved oxygen has a complex relationship with metals in regard to adsorption and release. Li et al. (2013) observed release of Zn, Cu, Cr, and Pb increased under aerobic conditions while the adsorption of metals and release
of Cd occurred under anaerobic conditions [62]. During reducing conditions, co-precipitated metals with Mn and Fe oxides were also found to be released and at stronger reductive conditions sulphide can complex with metals and reduce mobility [62]. Li et al. also found greater adsorption of metals at lower DO concentrations than at higher DO concentrations possibly resulting from adsorption to iron and manganese (hydr)oxide solid phases and sulphide complexation [62]. In contrast, Kang et al. (2019) showed that that Pb, Zn and Fe were released from sediments in an anoxic environment and adsorbed from the overlying water in an aerobic environment [54]. They found that adsorption onto surfaces of iron and manganese oxy-hydroxides occurred more in oxidizing conditions and metals were released under reducing conditions. Thus, higher dissolved oxygen better facilitated the release of metals than lower dissolved oxygen conditions.

5. Conclusions

The impact of urbanization on the transport and transformation of multiple metals in urban streams has led to a need for characterizing the constituents of this chemical cocktail more holistically. Using multiple linear regression approaches, our analysis suggests that predicting metals concentrations from continuous variables only is effective for some metals at some sites. However, there were cases where data from individual sensors showed relatively strong statistical relationships with trace metal data. One potential implication is that more continuous sensor variables should be considered for additional and more diverse measurements to try to explain more of the variation in metals concentrations. Difficulty in applying some sensors to predict metals concentrations in some cases could also be the result of a limited sample size for discrete characterization of stream water chemistry for grab samples. Furthermore, the metals data (with the exception of mercury) are represented by free or dissolved metals. Perhaps, analysis of total metals would produce better correlations with more continuous sensor variables. Future work could focus on augmenting current data sets with collection of dissolved and total metals data and additional sensor data to create more robust predictive models that could eventually serve to streamline the process of assessing water quality with regards to metals. Overall, results from continuous environmental sensors can help inform monitoring and management efforts for metals in urban streams. The use of sensor data to develop predictions would not replace the importance of direct measurements of metals concentrations via grab samples and traditional analyses using ICP-MS and ICP-OES. Rather, it could help to predict datasets at a higher-frequency temporal resolution and assess whether more time-consuming and costly measurements should be employed based on using statistical models. For future predictor models with low correlation coefficients, further research should include utilization of more predictor variables and expanding the calibration range of current sensors. There are several implications of this work in terms of machine learning, which is conceptually similar to this work [63,64]. Artificial intelligence shows great promise over traditional statistical techniques and is well suited to real time sensors [65]. The combination of artificial intelligence and linear regression models can also be used to predict chemical cocktails of base cations and anions [66]. Eventually, the integration of artificial intelligence and sensor approaches could be used in optimizing urban stormwater management best management practices [67]. All of these future possibilities could potentially enhance machine learning approaches to predict and manage metal concentrations across a wider range of watersheds.

Supplementary Materials: The following are available online at www.mdpi.com/2073-4441/12/10/2864/s1,
Figure S1: Sampling of metals over instantaneous discharge for the Rock Creek (a), Hickey Run (b), and Watts Branch (c) watersheds; Figure S2: Daily fluxes of cadmium (a), copper (b), lead (c), zinc (d) and mercury (e) for the Rock Creek, Hickey Run, and Watts Branch watersheds; Figure S3: Simple Linear Regression Relationships for turbidity and Hg (a), turbidity and Pb (b), specific conductance (c) and discharge (d) across the Rock Creek, Hickey Run, and Watts Branch watersheds; Figure S4: Linear regression model for instantaneous discharge and Cu flux at the Rock Creek watershed (R² = 0.95); Figure S5: Linear Regression Model for discharge and Pb flux at the Rock Creek watershed (R² = 0.95); Figure S6: Linear regression model for turbidity and Hg at all sites. (R² = 0.13); Figure S7: Linear regression model for turbidity and Pb at all sites (R² = 0.20); Table S1: Table of max, median, mean, standard error of the mean (SEM), min and quartiles for daily flux concentrations at the Rock
Creek watershed; Table S2: Table of max, median, mean, standard error of the mean (SEM), min and quartiles for flux concentrations at the Watts Branch watershed; Table S3: Table of max, median, mean, standard error of the mean (SEM), min and quartiles for flux concentrations at the Hickey Run watershed; Table S4: Percent differences below 40% between predicted and actual Cu fluxes for the linear regression model displayed in Figure S4; Table S5: Percent differences below 40% between predicted and actual Pb fluxes for the linear regression model displayed in Figure S5.

Author Contributions: Conceptualization, C.J.M. and S.S.K.; methodology, C.J.M. and S.S.K.; writing—original draft preparation, C.J.M. and M.L.T.; writing—review and editing, S.S.K. and K.T.B.; funding acquisition, C.J.M. and S.S.K. All authors have read and agreed to the published version of the manuscript.

Funding: Funding was provided by NSF EAR1521224, Maryland Sea Grant SA75281870W and a USGS Pathways Fellowship to C.J.M.

Acknowledgments: This project would not have been possible without data collection and sharing by the USGS. In particular, Joe Bell and Joel Blomquist provided valuable information on sampling design and insights related to the study sites. Karen Prestegaard and Shuiwang Duan kindly provided helpful comments and ideas that significantly improved this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References


Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).