United States Environmental Protection Agency

Office of Water
EPA 820-R-15-106
March 2016

Draft Technical Support Document: Recommended Estimates for Missing Water Quality Parameters for Application in EPA's Biotic Ligand Model

Disclaimer Page

This technical support document (herein referred to as the "Missing Parameters TSD") summarizes data analysis approaches EPA used to develop recommendations for default values for water quality parameters used in the copper BLM when data are lacking. When published in final form, this document will provide information to states, tribes, and the regulated community interested in using the Biotic Ligand Model to protect aquatic life from toxic effects of copper. Under the CWA, states and tribes are to establish water quality criteria to protect designated uses. State and tribal decision makers retain the discretion to adopt approaches on a case-by-case basis when appropriate. This document does not substitute for the CWA or EPA's regulations; nor is it a regulation itself. Thus, it cannot impose legally binding requirements on EPA, states, tribes, or the regulated community, and might not apply to a particular situation based upon the specific circumstances. EPA may change this document in the future. This document has been approved for publication by the Office of Science and Technology, Office of Water, U.S. Environmental Protection Agency.

Mention of trade names or commercial products does not constitute endorsement or recommendation for use. This document can be downloaded from:
http://water.epa.gov/scitech/swguidance/standards/criteria/aqlife/copper/2007_index.cfm

Cover photo: Credit USEPA

Acknowledgements

Luis A. Cruz
(Document Coordinator and Contributor)
United States Environmental Protection Agency
Health and Ecological Criteria Division
Washington, D.C. 20460

Reviewers
Elizabeth Southerland, Elizabeth Behl, Kathryn Gallagher
United States Environmental Protection Agency
Office of Science and Technology
Washington, D.C. 20460

Acronyms or Abbreviations

Acronym/Abbreviation	Definition
ACR	Acute to Chronic Ratio
ASE	Average Standard Error
BLM	Biotic Ligand Model
BOD	Biochemical Oxygen Demand
CCC	Criterion Continuous Concentration
CMC	Criterion Maximum Concentration
CWA	Clean Water Act
df	Degrees of Freedom
DOC	Dissolved Organic Carbon
Ecoregion	Areas in which ecosystems and the type, quality, and quantity of environmental resources are generally similar. In this document is represented by EPA Level III Ecoregions
EMAP	Environmental Monitoring and Assessment Program
EPA	United States Environmental Protection Agency
FSC	Fixed Site Criteria
GI	Geochemical Ions; ion parameters for the Biotic Ligand Model (Ca, Mg, Cl, Na, K, SO_4, alkalinity)
GIS	Geographic Information System
GLEC	Great Lakes Environmental Center
HUC	Hydrologic Unit Code (a two to eight digit code that identifies each hydrologic unit)
IQR	Interquartile range
IWQC	Instantaneous Water Quality Criteria
LCL	Lower Confidence Limit
LDC	Legacy Data Center (EPA historical STORET database, recently renamed)
LR	Linear Regression
mg/L	Milligrams per liter
NHD	National Hydrography Dataset
NHD-Plus	National Hydrography Dataset Plus
NOCD	National Organic Carbon Database (combined organic carbon data from two databases: USGS WATSTORE and EPA STORET)
NRC	National Research Council
NRSA	National River and Stream Assessment
NWIS	National Waters Information System
PCS	Permit Compliance System
RPDs	Relative Percent Differences
PCA	Principal Components Analysis
pH	Negative logarithm of the hydrogen ion concentration (in moles per liter); scale range from 0 to 14
POC	Particulate Organic Carbon
POTWs	Publicly Owned Treatment Works
RMSE	Root Mean Square Error

Acronym/Abbreviation	Definition
RSS	Residual Sum of Squares
SPLOM	Scatter Plot Matrices
SO	Strahler Stream Order
STORET	EPA STOrage and RETrieval Data Warehouse (recently renamed Legacy Data Center, LDC)
TDS	Total Dissolved Solids
TSD	Technical support document
TSS	Total Suspended Solids
TOC	Total Organic Carbon
UCL	Upper Confidence Limit
U.S.	United States
USGS	United States Geological Survey
WATSTORE	USGS National WATer Data STOrage and REtrieval System (predecessor of NWIS)
WSA	Wadeable Stream Assessment
WQC	Water Quality Criteria
µS/cm	Microsiemens per centimeter
µg/L	Micrograms per liter

Executive Summary

The United States (U.S.) Environmental Protection Agency (EPA) developed revised freshwater aquatic life criteria for copper using the Biotic Ligand Model (BLM) in 2007. The 2007 Freshwater Copper BLM predicts acute copper toxicity based on site-specific water quality parameters, and calculates aquatic life criteria based on the predicted copper toxicity. The current freshwater copper BLM requires 10 input parameters to calculate copper criteria: temperature, pH, dissolved organic carbon (DOC), alkalinity, calcium, magnesium, sodium, potassium, sulfate, and chloride, the last seven of which are also referred to as geochemical ions (GI). Previously available hardness-based copper criteria incorporated consideration only of the effects of hardness on bioavailability, while the BLM incorporates consideration of all of the water chemistry parameters that have a major influence on metal bioavailability. This allows the BLM-based criteria to be customized to the particular water body under consideration. However, given the broad geographical range over which the BLM is likely to be applied, and the limited availability of data for input parameters in many areas, a practical method to estimate missing water quality parameters was needed to successfully run the BLM. This technical support document (herein referred to as the "Missing Parameters TSD") summarizes data analysis approaches EPA used to develop recommendations for default values for water quality parameters used in the Freshwater Copper BLM when data are lacking. These default values could also be used to fill in missing water quality input parameters in the application of other metal BLM models as well, when data are lacking. EPA used three approaches to develop these default value recommendations:

- Conducted geostatistics and conductivity analyses to predict GI parameters
- Applied stream order to refine prediction of GI parameters
- Mined the National Organic Carbon Database (NOCD) to estimate DOC

In brief, EPA found that an approach that used correlation (with conductivity and discharge as explanatory variables), combined with geostatistical techniques (kriging), and a consideration of stream order produced the best estimates for BLM GI parameters. Tables 8, 9, and 10 present estimated inputs for each GI and water hardness in each ecoregion categorized by stream order for low, medium, and high order streams, respectively. Recommended GI values are based on the 10^{th} percentile of ecoregional Level III values for the appropriate stream order (size) and are expected to yield appropriately protective criteria values when applied in the BLM model. In Table 20 of Section 4, EPA provides estimates for DOC by ecoregion based on an analysis of a compilation of national organic carbon databases. The 10^{th} percentile of ecoregional Level III values are recommended for DOC. There was insufficient data to refine the DOC estimates by stream order. EPA recommends measurement of pH and temperature directly to use as an input in the BLM. Temperature is a commonly measured parameter, and should be easily obtainable for use in the BLM. The following paragraphs summarize the contents of each section in this report.

Section 1 provides an introduction to this study, including background on the BLM. In developing the approaches outlined in this study, EPA relied upon several previous studies that attempted to estimate values for BLM input water quality parameters; these studies are outlined briefly in Section 1 and are described in detail in Appendices A through D. This earlier work demonstrates that protective water

quality criteria (WQC) for copper generally corresponded to a low percentile of the distribution of instantaneous water quality criteria (IWQC) predicted by the BLM.

Section 2 provides a discussion of the approach taken by EPA to estimate BLM GI parameter values using geostatistics, which are a suite of statistical methodologies that use spatial coordinates to formulate models used in estimation and prediction. Section 2 also describes how EPA supplemented the geostatistical approach with conductivity as an explanatory variable, because conductivity data are abundant and correlate well to the BLM GI parameters.

Section 3 provides an analysis and discussion of the EPA approach to estimation of BLM GI parameters incorporating stream order as a variable, with a goal of providing BLM users with tables of recommended GI parameter estimates based upon both ecoregion and stream order. For each Level III ecoregion, we recalculated the 10^{th} percentiles of the distributions of all daily water quality parameters measured at all NWIS stations taking into account stream orders or ranges (groups) of stream orders within each ecoregion. Values of the BLM GI parameters generally increased with stream order. Based upon this trend, we grouped the estimates for each parameter by stream order: 1 through 3 (headwater streams), 4 through 6 (mid-reaches), and 7 through 9 (rivers).

Section 4 discusses the estimation of DOC based on the NOCD and two other databases. The NOCD was compiled from a number of sources, including EPA's Storage and Retrieval Data Warehouse (STORET) and the United States Geological Survey's National Water Data Storage and Retrieval System (WATSTORE) (the predecessor of the National Waters Information System (NWIS)). The two other databases, the Wadeable Stream Assessment (WSA) and the National River and Stream Assessment (NRSA), were used to supplement and update the DOC analysis. Section 4 summarizes the data sources, analysis, and uncertainty associated with ecoregional statistics for the NOCD and outlines how tests for bias in the data influence selection of 10^{th} percentile DOC concentrations from either the NOCD or the WSA or NRSA databases. The importance of field sampling for DOC is highlighted in Section 4 because of limitations of the NOCD and the importance of DOC in criteria calculation.

Section 5 provides a summary of the three approaches used to develop EPA's recommendations. Taken together, the approaches presented in this TSD describe EPA's recommendations for default input parameters in the BLM to derive protective freshwater aquatic life criteria when data are lacking. However, it should be noted that site-specific data are always preferable for developing criteria based on the BLM and should be used when possible. Users of the BLM are encouraged to sample their water body of interest, and to analyze the samples for the constituent (parameter) concentrations as a basis for determining BLM inputs where possible.

Table of Contents

List of Figures

List of Tables

1 INTRODUCTION

1.1 Background and Objective

The United States (U.S.) Environmental Protection Agency (EPA) has a congressional mandate to develop and publish criteria for water quality that reflects the effects of pollutants on aquatic life and human health under 304(a)(1) of the Clean Water Act (CWA). The CWA was intended to protect the chemical, physical, and biological integrity of the Nation's waters. Section 304(a)(1) of the CWA, 33 U.S.C. § 1314(a)(1), directs the Administrator of EPA to publish water quality criteria that accurately reflect the latest scientific knowledge on the kind and extent of all identifiable effects on health and welfare that might be expected from the presence of pollutants in any body of water, including ground water. Under this authority, EPA developed revised aquatic life criteria for copper that are based on the Biotic Ligand Model (BLM) in 2007. The BLM predicts metal toxicity based on site-specific water quality parameters, and derives acute and chronic criteria from the predicted toxicity. Derivation of water quality criteria using the BLM requires 10 input parameters (temperature, pH, dissolved organic carbon (DOC), alkalinity, calcium, magnesium, sodium, potassium, sulfate, and chloride). Data regarding these parameters may not be available for many receiving waters. Given that the BLM is likely to be applied over a broad geographical range, and that limited data are available for many areas, a practical method to estimate missing water quality parameters was needed to facilitate full use of the BLM in water quality standards across the U.S. This technical support document (herein referred to as the "Missing Parameters TSD") summarizes data analysis approaches EPA used to develop recommendations for default values for water quality parameters that may be used in the BLM when data are lacking. The section of the CWA related to the development of the information presented in this technical support document is CWA Section 304(a)(2). CWA Section 304(a)(2) generally requires EPA to develop and publish information on the factors necessary to restore and maintain the chemical, physical, and biological integrity of navigable waters. Section 304(a)(2) also allows EPA to provide information on the conditions necessary for the protection and propagation of shellfish, fish, and wildlife in receiving waters and for allowing recreational activities in and on the water.

The objective of this report is to summarize recommendations that BLM users can apply to estimate values for missing input water quality parameters.

1.2 Input Data and the BLM

The BLM calculates metal toxicity to aquatic organisms as a function of the concentrations of certain chemical constituents of water, including, for example, ions that can complex with metals and limit biological availability, and ions that compete with metals for binding sites at the ion exchange tissues of aquatic organisms (e.g., at the fish gill). The BLM predicts the metal criteria concentrations, such as copper in freshwater, which will vary according to changes in the associated water quality parameters.

An appropriately protective acute and chronic copper (or other metals) criteria must reflect the variability of water quality parameters at the site. In previous analyses, EPA found that protective water quality criteria for copper generally correspond to approximately the 2.5th percentile of the

distribution of instantaneous water quality criteria (IWQC) predicted by the BLM[1] (USEPA, 2002). Thus, predictions made for a site using the corresponding low percentile of the water quality parameter distributions are appropriately protective. Copper BLM predictions are most sensitive to the following five important parameters: DOC, pH, and calcium, magnesium, and sodium concentrations (taken together). Estimates are most sensitive to DOC, and vary in direct proportion to a change in value (i.e., they are 100% sensitive to DOC). Estimates are 50% sensitive to a change in pH, and 20% sensitive to the combined concentrations of calcium, magnesium, and sodium.

1.3 Previous Studies

EPA has conducted previous studies to develop tools to estimate BLM water quality parameters for sites where there may be few (or no) water quality data available. Brief summaries of these previous studies are provided below, and more detailed descriptions are provided in Appendices A through D.

1.3.1 An Examination of Spatial Trends in Surface Water Chemistry in the Continental United States: Implications for the Use of Default Values as Inputs to the BLM for Prediction of Acute Metal Toxicity to Aquatic Organisms (Carlton, 2006)

A large database of surface water chemistry monitoring data was examined to look for spatial trends in five chemical constituents that are key inputs to a model for predicting metal toxicity to aquatic organisms: pH, dissolved organic carbon, alkalinity, calcium, and sodium. Continuous prediction maps of concentrations were generated using various kriging techniques to interpolate between site-median values measured at several thousand separate locations throughout the continental U.S. Continuous concentration surfaces were then averaged over 8-digit Hydrologic Unit (HUC) polygons to produce block-averaged mean estimates of site-median concentrations. Pairwise comparisons indicated distinct trends between various HUC-averaged predicted constituents. The same analyses performed on data from 772 locations where all five constituents had been measured revealed similar relationships between monitored constituents. Principal components analyses (PCA) performed on these data sets showed that 80 to 90 percent of the variance in both cases could be explained by a single component with loadings on three of the five constituents. The use of kriging to produce appropriate quantile maps for block-averaging is suggested as a possible approach for developing regional values to use as default model inputs, when site-specific monitoring data are lacking. Refer to Appendix A for more information.

1.3.2 Approaches for Estimating Missing BLM Input Parameters: Correlation Approaches to Estimate BLM Input Parameters Using Conductivity and Discharge as Explanatory Variables (USEPA, 2007)

In this 2007 report, EPA developed regression models to project BLM water quality parameters from conductivity data. EPA assessed supplementing the geostatistical approach with classical estimation methods, such as regression and correlation by assessing the degree of correlation between conductivity and each of the BLM water quality parameters using National Water Information System

[1] This was the median for 17 sites; the range was 1 to 36%.

(NWIS) data from three states (Colorado, Utah, and Wyoming). These states were selected because of the large spatial and temporal variability observed.

EPA concluded that conductivity is significantly correlated with BLM water quality parameters between sites, especially for the low-end distribution statistics of interest for parameter estimation. Since conductivity data are abundant and correlate well with BLM water quality parameters, EPA determined it is reasonable to incorporate conductivity in spatial projections of BLM parameters. Correlation coefficients were lower for pH and DOC than for the geochemical ions (GIs) and alkalinity, but were also significant. Refer to Appendix B for more information.

1.3.3 Copper Biotic Ligand Model (BLM) Software and Supporting Documents Preparation: Development of Tools to Estimate BLM Parameters (USEPA, 2008)

In order to predict parameters based on geographic location, this 2008 report investigated how to project BLM water quality parameters for a given site based on other site data using geostatistical methods. There are a number of ways in which the conductivity regressions can be used to project BLM water quality inputs. The regressions allow estimates of the BLM water quality inputs from either: (1) a limited number of conductivity measurements, or (2) a low-end conductivity value estimated by geostatistical methods.

The first approach, projecting BLM water quality inputs from conductivity measurements, was demonstrated for a limited number of test sites. Regression models were developed to project 10^{th} percentiles of BLM water quality input parameters from the 10^{th} percentile of measured conductivity distributions at sites in Colorado, Utah, and Wyoming. The 10^{th} percentile is the value below which 10 percent of the observations may be found. The regression models were tested using data and copper BLM predictions for four sites, and produced highly consistent results. The regression models for pH and DOC, the most sensitive of the BLM water quality parameters, were not sufficiently accurate to make reliable BLM parameter predictions. However, regression models for the GI parameters (calcium, magnesium, sodium, potassium, chloride, sulfate, and alkalinity) were reasonably accurate, as judged by comparing model predictions made using projected values of the these BLM input parameters to model predictions made using measured input data. No estimate for site-specific pH was superior to the observed weak conductivity regression. To improve upon this estimate, it was necessary to use actual site-specific pH data. For DOC, the Level III ecoregion (referred to herein as simply “ecoregion”) and water body type-specific DOC concentration percentiles tabulated by EPA for the National Bioaccumulation Factors Technical Support Document (USEPA, 2003) appear to be far better estimates of lower-percentile DOC concentrations than the estimates made using the conductivity regression.

EPA also provided a proof of concept for the second approach, which was to see whether combining the kriged conductivities with the conductivity-hardness regression would project the 10^{th} percentiles of hardness better than direct kriging of the hardness data. EPA found that both approaches produce estimates of hardness that correlate significantly with the measured data (correlation coefficient $r=0.80$ for direct kriging of hardness; $r=0.95$ for conductivity kriging + regression). However, the kriging + regression approach fits the hardness data substantially better than direct kriging. Refer to Appendix C for more information.

1.3.4 Approaches for Estimating Missing BLM Input Parameters: Projections of Total Organic Carbon as a Function of Biochemical Oxygen Demand (USEPA, 2006a)

DOC concentrations downstream of an effluent discharge are necessary inputs for the BLM to predict toxicity associated with a wastewater discharge. Effluent DOC is monitored by very few publicly-owned treatment works (POTWs) according to data retrieved from EPA's Permit Compliance System (PCS). Biochemical oxygen demand (BOD) is monitored by most POTWs. EPA developed regressions to project total organic carbon (TOC) concentrations from BOD values using effluent samples at all POTWs reporting data for both parameters in EPA's PCS. EPA concluded that this regression gives reasonable estimates of TOC in POTWs effluents and are likely the best available estimates of effluent TOC to determine DOC concentrations for the BLM. Refer to Appendix D for more information.

1.4 Approaches to Estimate Water Quality Parameters for the BLM

Building upon the studies described above, this report uses three approaches to develop default estimates for parameters needed for the BLM when empirical data are lacking. The three approaches are listed below and are detailed in the following sections:

- Section 2: Using Geostatistics and Conductivity to Predict GI Parameters
- Section 3: Using Stream Order to Refine Prediction of GI Parameters
- Section 4: Using the National Organic Carbon Database to estimate DOC

EPA recommends that temperature and pH be measured directly in the field.

2 USING GEOSTATISTICS AND CONDUCTIVITY TO PREDICT GI PARAMETERS

The following section describes studies that demonstrate how geostatistical techniques, coupled with conductivity correlations, can be used to predict BLM input parameters for GIs when site-specific data are unavailable. In a previous study (USEPA, 2008) EPA demonstrated that combining kriging with regressions to estimate inputs based on conductivity improves the accuracy of GI estimates. In this section EPA has expanded on this approach and developed national estimates at the Level III ecoregion in the continental U.S.

The current freshwater copper BLM requires 10 input parameters that reflect water chemistry in order to calculate copper criteria: temperature, pH, DOC, alkalinity, calcium, magnesium, sodium, potassium, sulfate, and chloride, the last seven of which are GIs. The concentrations of GIs vary in surface waters due to dissolution, weathering, ground water-surface water interactions, and other geologic processes in the watershed, in addition to dilution by snowmelt and precipitation. Consequently, the concentrations of GI parameters tend to vary according to the regional geology. For example, alkalinity has noticeable geographic trends. Areas dominated by carbonate rocks, such as limestone as in the prairie states, tend toward high alkalinity. Areas dominated by igneous rocks, such as granite, such as parts of the northeast, tend toward low hardness and alkalinity.

In this section we expand on the EPA 2008 proof of concept (in Appendix C) using geostatistics to develop default missing GI parameter values based on geography. Geostatistics are statistical methodologies that use spatial coordinates to help formulate models used in estimation and prediction (ESRI, 2003). Geostatistical techniques are attractive because they explain parameter variation arising from spatial correlations, which are not used in conventional statistics. We have supplemented the geostatistical approach by adding conductivity as an additional explanatory variable. Conductivity is one of the most widely monitored water quality indicators in the U.S. Because conductivity data are abundant and correlate well to the BLM GI parameters, we incorporated conductivity in spatial projections of BLM parameters. Based on the proof of concept described above (and in Appendix C), we expected that this approach, which can be implemented by co-kriging (i.e., an interpolation technique that allows for better estimates by the incorporation of well-sampled, correlated secondary data) in geostatistical software, would allow more robust spatial projections of BLM water quality parameters.

2.1 Data Source and Processing

Water quality data for conductivity and BLM GI water quality parameters were retrieved from the United States Geological Survey (USGS) National Water Information System (NWIS). NWIS contains data from millions of sampling events at tens of thousands of individual sampling locations (stations) in the continental U.S. (Figure 1). Not all water quality parameters of relevance to the BLM were monitored at each location. The numbers of sampling events at individual locations also range widely, with a mean of 15, and a mode of one (i.e., most sites were only sampled once). Examination of the spatial distribution of numbers of sampling events per site reveals that the Midwestern and Western states tended to be sampled most intensively (Carleton, 2006). Because environmental sampling data tend to be lognormally distributed, disparities in numbers of samples may tend to produce higher mean and median values at locations that have been sampled more frequently. As spatial distributions of representative (e.g., median) concentrations are examined, it should be kept in mind that apparent

geographic trends in concentration may be in part simply the result of uneven sampling intensity (Carleton, 2006).

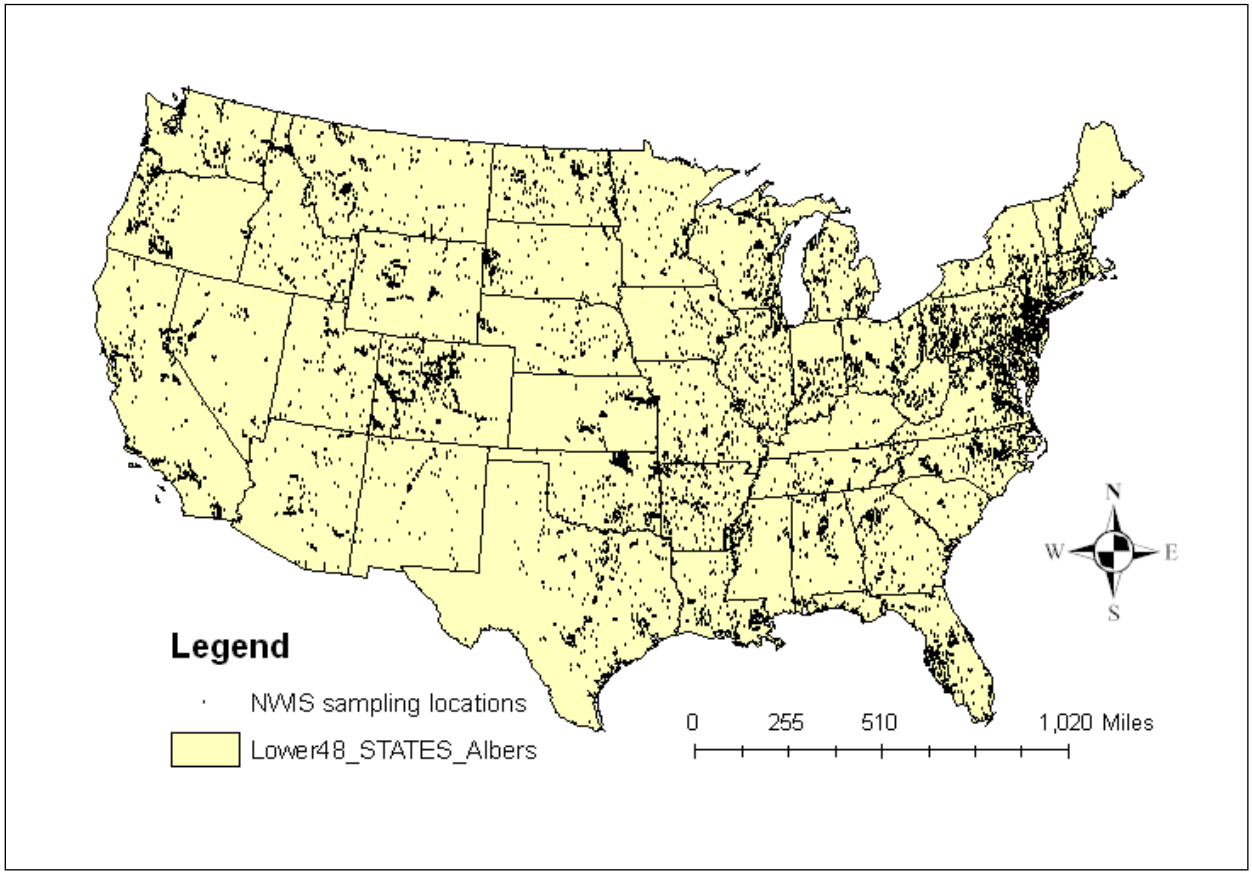

Figure 1. NWIS sample collection locations in the continental U.S. (Carleton, 2006)

We focused our efforts on data collected from rivers and streams between 1984 and 2009. Data collected prior to 1984 were excluded because a number of the analytical methods used by USGS prior to that date have been replaced by methods with improved precision and lower detection limits. Furthermore, only sites with 40 or more samples were included in the analysis. With support from USGS staff, we obtained a complete download of national water quality data from NWIS, which totaled 4,714,165 measurements from 959,946 samples, collected at 5,901 sites. These data included measurements for BLM water quality input parameters required to calculate copper criteria using the BLM: pH, DOC, alkalinity, calcium, magnesium, sodium, potassium, sulfate, and chloride. Data were also collected on filtered (dissolved) copper, and the spatial coordinates (latitude and longitude) of each sampling station were also retrieved. No data were collected on temperature. Only the GI data were included in the geostatistical analysis. A summary of the water quality data retrieved from NWIS is provided in Table 1.

Table 1. Summary of water quality data retrieved from NWIS			
BLM Water Quality Parameter	**NWIS Parameter Code**	**Parameter Description**	**Number of Observations**
Conductivity[1]	00094	Specific conductance, water, unfiltered, field, microsiemens per centimeter at 25 degrees Celsius	553,700
	00095	Specific conductance, water, unfiltered, laboratory, microsiemens per centimeter at 25 degrees Celsius	799
pH	00400	pH, water, unfiltered, field, standard units	352,336
	00403	pH, water, unfiltered, laboratory, standard units	151,161
Dissolver Organic Carbon	00681	Organic carbon, water, filtered, laboratory, milligrams per liter	30,008
Alkalinity	00410	Acid neutralizing capacity, water, unfiltered, fixed endpoint (pH 4.5) titration, field, milligrams per liter as calcium carbonate	35,232
	00417	Acid neutralizing capacity, water, unfiltered, fixed endpoint (pH 4.5) titration, laboratory, milligrams per liter as calcium carbonate	15,264
	00419	Acid neutralizing capacity, water, unfiltered, incremental titration, field, milligrams per liter as calcium carbonate	10,198
	00418	Alkalinity, water, filtered, fixed endpoint (pH 4.5) titration, field, milligrams per liter as calcium carbonate	2,686
Calcium	00915	Calcium, water, filtered, laboratory, milligrams per liter	146,608
Magnesium	00925	Magnesium, water, filtered, laboratory, milligrams per liter	145,938
Sodium	00930	Sodium, water, filtered, laboratory, milligrams per liter	136,310
Potassium	00935	Potassium, water, filtered, laboratory, milligrams per liter	132,659
Sulfate	00945	Sulfate, water, filtered, laboratory, milligrams per liter	147,824
Chloride	00940	Chloride, water, filtered, laboratory, milligrams per liter	146,601

[1] Conductivity is not a BLM parameter, but was used as an explanatory variable for the other water quality parameters.

The data were screened using established quality assurance procedures. All data were checked to confirm that they contained numerical values without null (missing) records and remark codes were

identified and reviewed. Minimum and maximum values for each parameter were confirmed to be within expected ranges and frequency distributions were plotted and examined for each of the parameters to identify outliers. We also confirmed that the spatial coordinate data placed each sampling location within the continental U.S. Additional data processing included the following steps:

- For the data at each station, the observations for each variable were averaged on a daily basis. This was done to reduce the influence of high frequency sampling at a few stations.
- Data were edited by censoring parameter(s) with fewer than 10 to 20 daily values at a station. The 10^{th} percentile for that parameter at that station was censored to improve the reliability of the lower-tail (i.e., 10^{th}) percentile statistics.
- Tenth (rank order/nonparametric) percentiles of the distributions of all water quality parameters measured at each station were calculated.

It should be emphasized that all of the statistical and geostatistical analyses and predictions presented in this report are based on the 10^{th} percentiles of the concentration distribution measured for each parameter at every station. The estimates of water quality parameter values for "missing" data are therefore also 10^{th} percentile concentrations. We selected the 10^{th} percentile of the site parameter distributions as a statistic that is a practical compromise between a lower-bound concentration and a percentile that can be reliably determined from small sample sizes. Initial testing with the BLM suggested that protective water quality criteria (WQC) for copper generally corresponded to approximately the 2.5^{th} percentile of the distribution of instantaneous water quality criteria (IWQC) predicted by the BLM. Thus, BLM predictions made for a site using the corresponding low percentiles of the water quality parameter distributions should (logically) also be a conservative approximation of a protective criterion. As a more reliably determined statistic, the 10^{th} percentile of water quality parameters will also derive reasonably protective criteria, especially for small sample sizes where there may be greater uncertainty at lower percentile estimates. The 10^{th} percentile estimates presented in this document were initially developed to implement the copper BLM published by EPA in 2007 and will apply to other metals as well.

2.2 Geostatistical Analysis of National Data for Geochemical Ions

The ESRI ArcGIS Geostatistical Analyst tool was used to create statistically valid two-dimensional surface models for conductivity and for each of the BLM GI parameters. Using the 10^{th} percentile daily average concentrations at each sampling location from the NWIS data, Geostatistical Analyst was used to create predictions for unmeasured locations throughout the continental U.S. For each parameter, the surface models were fit by minimizing the statistical error of the predicted surface. Surface fitting involved three steps: exploratory spatial data analysis, structural analysis (modeling the semivariogram to analyze surface properties of data from nearby locations), and surface prediction and assessment of the results. The semivariogram represents autocorrelation of measured data points spatially.

Modeling of the semivariogram was based on cross-validation, which calculates error statistics that serve as diagnostics to indicate whether the model is reasonable for map production. Cross-validation was used to select the models that provided the most accurate predictions. The following criteria were used to evaluate goodness of fit for the semivariogram model:

- Mean Standardized Error: close to 0;
- Root Mean Square Error (RMSE): as small as possible;

- Root-Mean-Square Standardized Error: close to 1; and,
- RMSE close to Average Standard Error ASE).

The difference between the prediction and the measured data value is the prediction error. For a model that provides accurate predictions, the mean prediction error should be close to 0 if the predictions are unbiased. The root-mean-square standardized prediction error should be close to 1 if the standard errors are accurate, and RMSE should be small if the predictions are close to the measured values (ESRI, 2003).

A tabulation of the geostatistical model selected for each water quality parameter, the number of data points interpolated, and the resulting error statistics are presented in Table 2. We used the optimal parameters for a spherical semivariogram as calculated by the Geostatistical Analyst. No transformations were applied to the data. Anisotropy (directional influence on the semivariogram) was not incorporated in the semivariogram models.

Table 2. Model selection and cross validation statistics for geostatistical fitting of 10th percentiles of BLM GI parameters

Parameter	Geostatistical model	Number of samples	Mean standardized error	Root mean square error	RMS standardized error	Average standard error
Conductivity	Universal kriging	4833	-0.01038	1361	1.081	1259
Alkalinity	Universal cokriging with conductivity	1372	-0.001115	36.62	1.09	33.23
Calcium	Universal cokriging with conductivity	2590	0.0001694	26.81	1.186	22.02
Magnesium	Universal cokriging with conductivity	2578	-0.002258	15.92	1.16	13.58
Sodium	Universal cokriging with conductivity	2439	-0.002929	156.3	1.583	95.78
Potassium	Universal cokriging with conductivity	2379	-0.001184	3.488	1.429	2.381
Sulfate	Universal cokriging with conductivity	2650	-0.0000225	114.5	1.29	87.04
Chloride	Universal cokriging with conductivity	2792	0.001653	375.2	1.51	247

2.2.1 Kriging of Conductivity Data

Universal kriging with a constant trend was used to map the surface of 10th percentile conductivity values. Kriging weights the surrounding measured values to derive a prediction for each location. The weights are based on the distances between the measured points and the prediction location, as well as the overall spatial arrangement among the measured points. The kriged prediction surface of 10th percentiles of conductivity is mapped in Figure 2. As the kriging results show, conductivities are highest in the south-central and southwestern regions, as well as along the Gulf and southern Atlantic coasts. Regions of lower conductivity are found in a number of parts of the country.

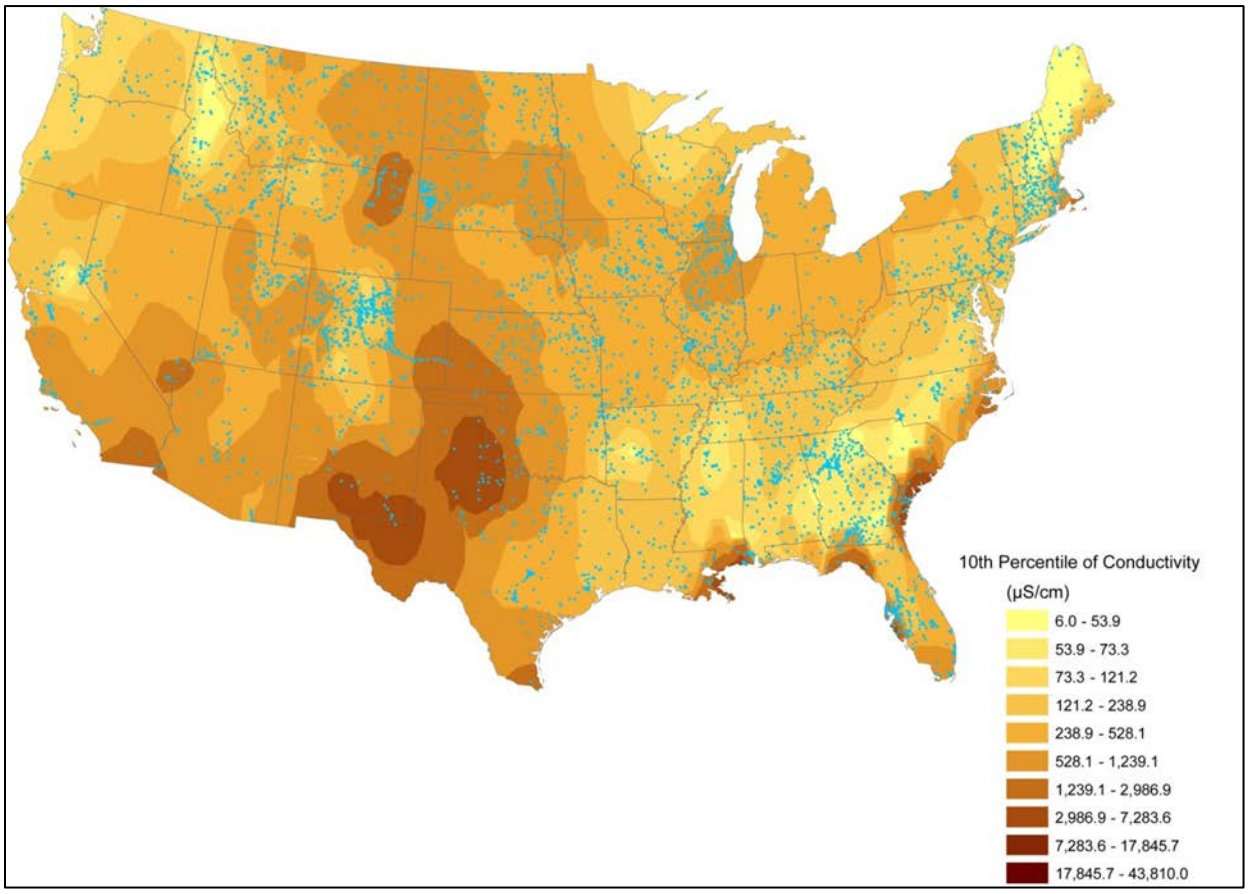

Figure 2. Kriged prediction surface for 10th percentile of conductivity in the continental U.S. (sample locations in blue)

2.2.2 Co-kriging of GI Data

Co-kriging was used to improve surface predictions of the BLM GI parameters by taking into account secondary variables, in this case conductivity. As demonstrated above, conductivity is significantly correlated to all of the BLM GIs. Universal co-kriging with conductivity, assuming a constant trend, was used to map the surface of 10th percentile BLM GIs concentrations. For each of these parameters, co-kriging produced cross-validation errors that were superior in terms of the goodness-of-fit criteria to errors produced by universal kriging. Prediction surfaces for calcium and alkalinity are mapped in Figures 3 and 4. The spatial distribution of calcium (Figure 3) shares a number of similarities with the mapping of conductivity (Figure 2). The co-kriged alkalinity surface (Figure 4) is rather different, with high alkalinity values reflecting geographic features (such as the carbonate geology of the prairie states) and low alkalinity values that reflect the granitic geology of the northeast. Prediction surfaces for the other BLM GI's are generally similar to those for conductivity and calcium.

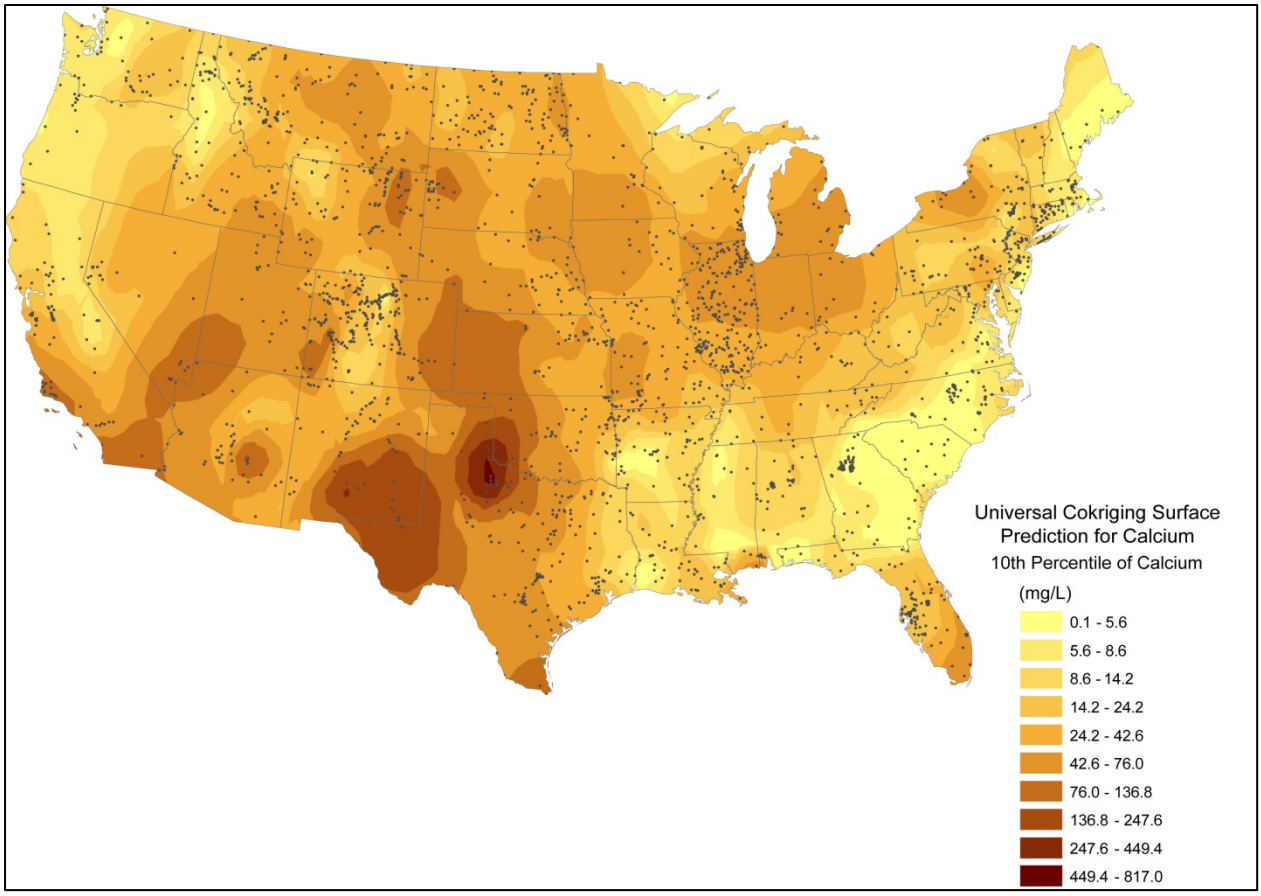

Figure 3. Co-kriged prediction surface for 10th percentile of calcium in the continental U.S.

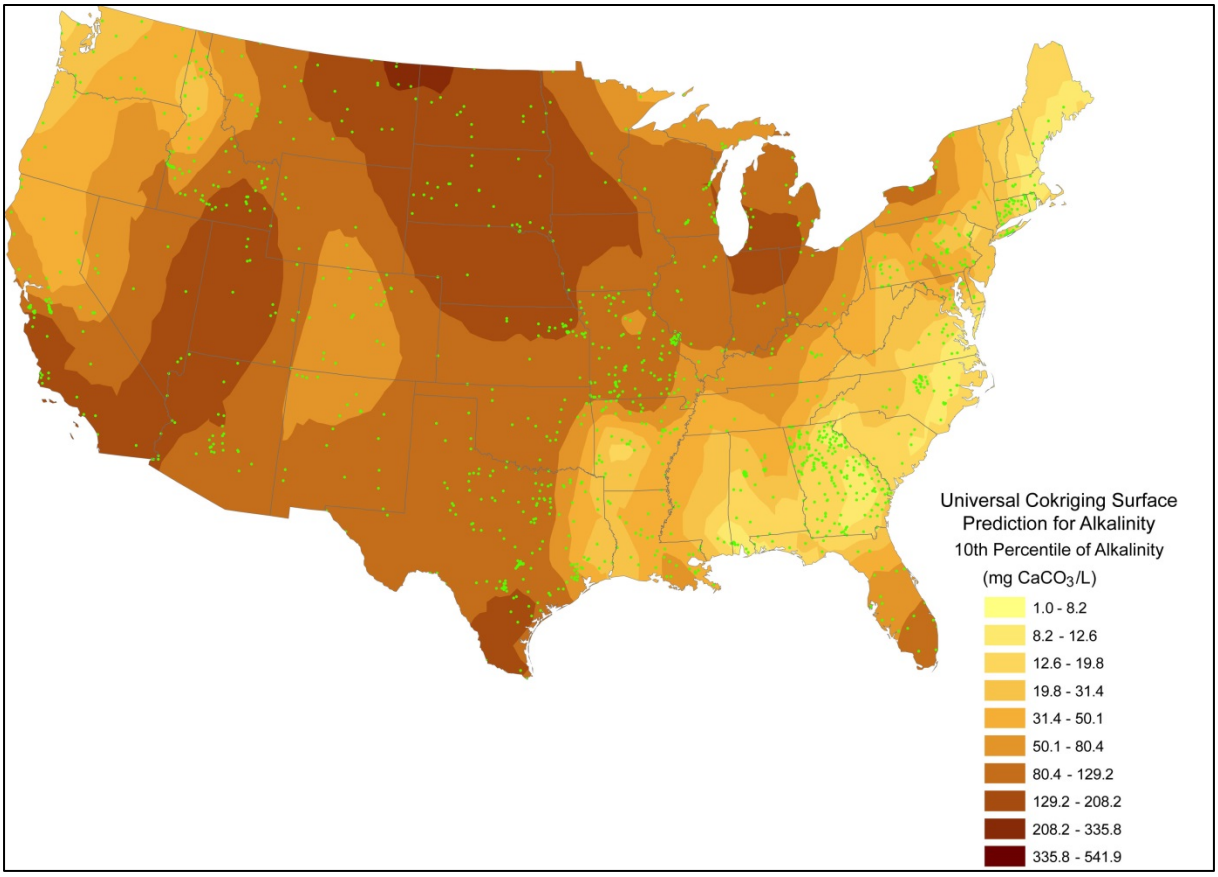

Figure 4. Co-kriged prediction surface for 10th percentile of alkalinity in the continental U.S.

2.2.3 Projection of Geostatistical Predictions onto Level III Ecoregions

Although maps of the geostatistical predictions are informative, a tabulation of the results is preferable for the purpose of providing guidance to BLM users. We chose to spatially average the geostatistical predictions of BLM water quality parameters according to the Level III ecoregions of the continental U.S. (Table 3), as these ecoregions provide a sound basis for spatial averaging of the water quality predictions. Ecoregions are designed to serve as a spatial framework for environmental resource management and denote areas within which ecosystems (and the type, quality, and quantity of environmental resources) are generally similar. They typically provide a logical and useful spatial (geographical) framework for organizing the results of environmental measurements (Omernik and Griffith, 2014). Ecoregions can be distinguished by landscape-level characteristics that cause ecosystem components to reflect different patterns in different regions (Omernik, 1987). "Level III Ecoregions of the Continental U.S." map layer shows ecoregion delineation based on common patterns of geology, physiography, vegetation, climate, soils, land use, wildlife, water quality, and hydrology. The map layer in Figure 5 was compiled by EPA (USEPA, 2013a) (http://www.epa.gov/wed/pages/ecoregions/level_iii_iv.htm).

Table 3. Level III ecoregions of the U.S. organized according to broader Level I ecoregions	
Level I Ecological Regions	
Level III Ecoregion	**Name of Ecoregion**
Marine West Coast Forest	
1	Coast Range
2	Puget Lowland
3	Willamette Valley
111	Ahklun and Kilbuck Mountains
113	Alaska Peninsula Mountains
115	Cook Inlet
119	Pacific Coastal Mountains
120	Coastal Western Hemlock-Sitka Spruce Forests
Northwestern Forested Mountains	
4	Cascades
5	Sierra Nevada
9	Eastern Cascades Slopes and Foothills
11	Blue Mountains
15	Northern Rockies
16	Idaho Batholith
17	Middle Rockies
19	Wasatch and Uinta Mountains
21	Southern Rockies
41	Canadian Rockies
77	North Cascades
78	Klamath Mountains
105	Interior Highlands
116	Alaska Range
117	Copper Plateau
118	Wrangell Mountains
Mediterranean California	
6	Southern and Central California Chaparral and Oak Woodlands
7	Central California Valley
8	Southern California Mountains
North American Deserts	
10	Columbia Plateau
12	Snake River Plain
13	Central Basin and Range
14	Mojave Basin and Range
18	Wyoming Basin
20	Colorado Plateaus
22	Arizona/New Mexico Plateau

Table 3. Level III ecoregions of the U.S. organized according to broader Level I ecoregions	
Level I Ecological Regions	
Level III Ecoregion	**Name of Ecoregion**
24	Chihuahuan Deserts
80	Northern Basin and Range
81	Sonoran Basin and Range
Temperate Sierras	
23	Arizona/New Mexico Mountains
Great Plains	
25	Western High Plains
26	Southwestern Tablelands
27	Central Great Plains
28	Flint Hills
29	Central Oklahoma/Texas Plains
30	Edwards Plateau
31	Southern Texas Plains
34	Western Gulf Coastal Plain
40	Central Irregular Plains
42	Northwestern Glaciated Plains
43	Northwestern Great Plains
44	Nebraska Sand Hills
45	Piedmont
46	Northern Glaciated Plains
47	Western Corn Belt Plains
48	Lake Agassiz Plain
Eastern Temperate Forest	
32	Texas Blackland Prairies
33	East Central Texas Plains
35	South Central Plains
36	Ouachita Mountains
37	Arkansas Valley
38	Boston Mountains
39	Ozark Highlands
51	North Central Hardwood Forests
52	Driftless Area
53	Southeastern Wisconsin Till Plains
54	Central Corn Belt Plains
55	Eastern Corn Belt Plains
56	Southern Michigan/Northern Indiana Drift Plains
57	Huron/Erie Lake Plains
59	Northeastern Coastal Zone
60	Northern Appalachian Plateau and Uplands

Table 3. Level III ecoregions of the U.S. organized according to broader Level I ecoregions	
Level I Ecological Regions	
Level III Ecoregion	**Name of Ecoregion**
61	Erie Drift Plain
63	Middle Atlantic Coastal Plain
64	Northern Piedmont
65	Southeastern Plains
66	Blue Ridge
67	Ridge and Valley
68	Southwestern Appalachians
69	Central Appalachians
70	Western Allegheny Plateau
71	Interior Plateau
72	Interior River Valleys and Hills
73	Mississippi Alluvial Plain
74	Mississippi Valley Loess Plains
75	Southern Coastal Plain
82	Laurentian Plains and Hills
83	Eastern Great Lakes and Hudson Lowlands
84	Atlantic Coastal Pine Barrens
Northern Forests	
49	Northern Minnesota Wetlands
50	Northern Lakes and Forests
58	Northeastern Highlands
62	North Central Appalachians
Tropical Wet Forests	
76	Southern Florida Coastal Plain
Southern Semi-Arid Highlands	
79	Madrean Archipelago
Taiga	
101	Arctic Coastal Plain
102	Arctic Foothills
103	Brooks Range
104	Interior Forested Lowlands and Uplands
106	Interior Bottomlands
107	Yukon Flats
108	Ogilvie Mountains
Tundra	
109	Subarctic Coastal Plains
110	Seward Peninsula
112	Bristol Bay-Nushagak Lowlands
114	Aleutian Islands

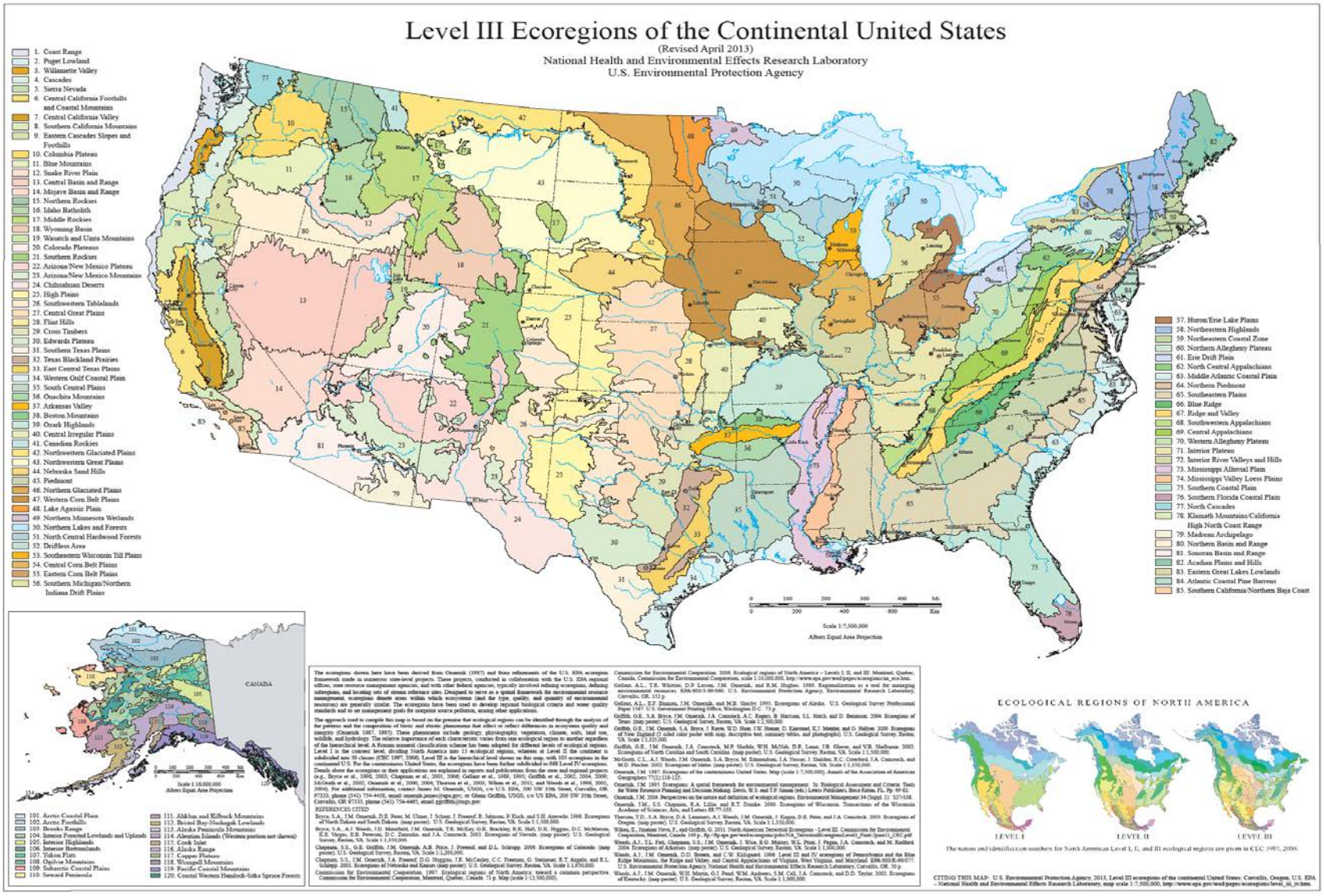

Figure 5. Map of Level III ecoregions in the U.S.

(Image taken from ftp://ftp.epa.gov/wed/ecoregions/us/Eco_Level_III_US.pdf)

Using the differences in land and water interactions, regional variations in attainable water quality, distinct biogeographical patterns (MacArthur, 1972), and similarities and differences in ecosystems to delineate ecoregions makes the application of ecoregions in environmental analyses a powerful tool with which to organize environmental information. The approach can take into account regional factors related to attainable water quality, and thus can be used to designate lakes for protection and to establish lake-restoration goals that are appropriate for each ecoregion (National Research Council (NRC), 1992). The NRC of the National Academy of Sciences has similarly endorsed the use of the concept in restoring and managing streams, rivers, and wetlands (NRC, 1992).

The theory of ecoregion delineation states that natural water quality characteristics of lakes and streams within a single ecoregion will be more similar than the characteristics between ecoregions (Perry and Vanderklein, 1996). Water quality characteristics exist in a landscape framework; neither normal nor impacted conditions of water resources can be separated from controlling influences of the surrounding landscape. The ecoregion concept has been applied and tested rather extensively in streams, rivers, and lakes. Testing and validation has been conducted in many diverse areas of the U.S., including several streams in Arkansas, Colorado, Kansas, Minnesota, Ohio, and Oregon, and in lakes of Michigan, Minnesota, Ohio, and Wisconsin (NRC, 1992).

Carleton (2006) used HUCs instead of ecoregions as the basis for averaging geostatistical results. HUCs are spatial delineations used for river basin management. Although a river basin may offer a logical framework for water supply management, for water quality management river basins are less applicable. The assumption that basins share similar properties is not always borne out, because river basins are often linked only by the water that flows through them. As Carleton (2006) noted, the use of HUCs for spatial averaging of surface water concentrations presents other conceptual difficulties. Only about 45% of HUCs are actual watersheds (Omernik, 2003); the rest receive drainage from additional upgradient areas. Concentrations measured in flowing waters reflect the soil, vegetation, and land use properties of the aggregate upstream drainage areas rather than of the sampling locations themselves (Smith et al., 1997).

2.2.3.1 Averaging Methods

To average the geostatistical predictions, a uniform grid was laid over the predicted surfaces and the predictions were sampled at the grid points falling within the polygons representing each ecoregion. The grid spacing was sized so that at least 30 points were sampled within each ecoregion. Unbiased log means were then calculated from the sampled concentration predictions in each ecoregion. The logarithmic transformation was applied because this normalized the concentration distributions in almost all of the ecoregions.

2.2.3.2 Tabulations of Ecoregional Estimated BLM Water Quality Parameters

The average predicted 10^{th} percentile concentrations for conductivity and the GI parameters for each of the Level III ecoregions in the continental U.S. are presented in Table 4. For each of these parameters, there is considerable variation between the ecoregional averages nationally. Chloride concentrations exhibit the greatest variation, with ecoregional 10^{th} percentile averages that range from 0.7 to 573 milligrams per liter (mg/L). Alkalinity was the least variable, but the ecoregional 10^{th} percentile average concentrations still ranged from 12 to 163 mg/L.

Table 4. Predicted 10th percentile concentrations for conductivity (µS/cm), BLM GI water quality parameters (mg/L) and hardness in each Level III ecoregion in the continental U.S.

Level III Ecoregion	Ecoregion Name	Unbiased log mean of 10th percentile concentrations Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness[1]
1	Coast Range	102	8.4	3.2	4.1	0.64	33	3.2	4.8	34.12
2	Puget Lowland	80	7.1	1.9	2.8	0.64	22	2.3	5.6	25.54
3	Willamette Valley	91	8.2	2.9	4.4	0.90	30	4.7	3.8	32.39
4	Cascades	107	6.6	2.9	3.5	0.74	35	2.2	3.2	28.39
5	Sierra Nevada	195	8.3	4.7	8.8	1.3	58	5.8	11	40.02
6	Southern and Central California Chaparral and Oak Woodlands	600	42	24	48	2.5	124	56	136	203.4
7	Central California Valley	378	21	16	25	1.7	91	21	58	118.1
8	Southern California Mountains	772	63	25	63	3.8	150	54	171	260
9	Eastern Cascades Slopes and Foothills	212	8.2	3.8	6.0	1.0	44	3.2	5.0	36.08
10	Columbia Plateau	166	15	5.2	9.3	1.8	40	3.3	10	58.82
11	Blue Mountains	142	11	3.9	7.7	1.4	49	3.3	7.1	43.49
12	Snake River Plain	273	33	10	13	2.3	109	10	22	123.5
13	Central Basin and Range	426	43	16	45	3.6	120	45	83	173.1
14	Mojave Basin and Range	976	69	27	81	6.3	138	85	258	283.2
15	Northern Rockies	90	11	3.1	2.3	0.67	44	0.72	4.4	40.21
16	Idaho Batholith	91	13	3.8	3.6	0.88	62	1.9	5.9	48.08
17	Middle Rockies	300	30	10	14	1.9	105	7.6	55	116
18	Wyoming Basin	446	35	13	33	1.7	96	7.2	104	140.8
19	Wasatch and Uinta Mountains	426	61	27	61	3.3	155	55	155	263.2
20	Colorado Plateaus	639	65	26	57	2.6	117	28	197	269.1
21	Southern Rockies	259	26	8.0	12	1.4	55	3.8	56	97.8
22	Arizona/New Mexico Plateau	697	50	15	65	3.0	96	65	143	186.5
23	Arizona/New Mexico Mountains	879	66	18	85	3.4	102	100	189	238.8
24	Chihuahuan Deserts	2712	176	50	379	8.6	106	573	608	645
25	High Plains	1770	104	35	191	6.0	112	281	353	403.5
26	Southwestern	2147	114	34	316	4.9	94	512	374	424.4

Table 4. Predicted 10th percentile concentrations for conductivity (µS/cm), BLM GI water quality parameters (mg/L) and hardness in each Level III ecoregion in the continental U.S.

Level III Ecoregion	Ecoregion Name	Unbiased log mean of 10th percentile concentrations Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness[1]
	Tablelands									
27	Central Great Plains	1228	84	24	176	6.9	121	245	204	308.4
28	Flint Hills	406	42	8.5	30	4.3	121	42	45	139.85
29	Central Oklahoma/Texas Plains	925	60	16	107	4.1	95	164	108	215.6
30	Edwards Plateau	596	48	14	38	2.7	98	62	52	177.4
31	Southern Texas Plains	798	56	14	58	3.8	129	73	91	197.4
32	Texas Blackland Prairies	364	39	5.8	21	3.2	92	26	33	121.28
33	East Central Texas Plains	367	36	6.3	23	3.8	98	29	29	115.83
34	Western Gulf Coastal Plain	565	43	10	62	3.9	87	86	78	148.5
35	South Central Plains	160	12	2.8	11	2.3	34	15	15	41.48
36	Ouachita Mountains	116	7.9	2.8	8.4	1.3	34	10	11	31.23
37	Arkansas Valley	192	16	4.7	15	1.8	51	20	16	59.27
38	Boston Mountains	152	18	3.3	4.3	1.3	53	6.7	8.2	58.53
39	Ozark Highlands	258	31	10	4.5	1.6	96	6.0	20	118.5
40	Central Irregular Plains	310	39	8.5	11	3.0	100	13	50	132.35
41	Canadian Rockies	164	22	8.7	15	0.57	80	1.7	38	90.67
42	Northwestern Glaciated Plains	545	37	20	61	5.9	163	8.1	147	174.5
43	Northwestern Great Plains	828	49	24	84	5.3	151	10	247	220.9
44	Nebraska Sand Hills	486	47	13	35	6.9	151	10	96	170.8
45	Piedmont	75	5.8	1.9	4.0	1.5	19	3.9	4.1	22.29
46	Northern Glaciated Plains	524	40	20	38	9.1	163	13	106	182
47	Western Corn Belt Plains	464	48	16	16	3.4	136	16	45	185.6
48	Lake Agassiz Plain	441	42	18	16	5.1	140	8.6	62	178.8
49	Northern Minnesota	229	24	10	3.2	1.4	95	2.6	8.4	101

Table 4. Predicted 10th percentile concentrations for conductivity (µS/cm), BLM GI water quality parameters (mg/L) and hardness in each Level III ecoregion in the continental U.S.

Level III Ecoregion	Ecoregion Name	Unbiased log mean of 10th percentile concentrations Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness[1]
	Wetlands									
50	Northern Lakes and Forests	166	19	6.5	2.5	0.78	83	3.4	6.1	74.15
51	North Central Hardwood Forests	295	31	12	5.6	1.7	115	8.9	15	126.7
52	Driftless Area	348	34	15	5.1	1.6	107	10	17	146.5
53	Southeastern Wisconsin Till Plains	510	39	20	16	2.1	112	31	25	179.5
54	Central Corn Belt Plains	546	53	24	14	1.7	124	30	46	230.9
55	Eastern Corn Belt Plains	463	49	15	11	2.0	116	23	32	184
56	Southern Michigan/Northern Indiana Drift Plains	463	52	15	14	1.9	134	28	29	191.5
57	Huron/Erie Lake Plains	467	52	15	13	2.2	125	27	32	191.5
58	Northeastern Highlands	97	11	1.9	5.7	0.69	24	10	7.4	35.29
59	Northeastern Coastal Zone	176	8.3	2.0	14	1.3	15	22	8.4	28.95
60	Northern Appalachian Plateau and Uplands	271	33	7.3	37	1.3	53	64	22	112.43
61	Erie Drift Plain	364	31	8.1	19	2.3	64	29	38	110.71
62	North Central Appalachians	184	13	3.9	7.1	1.0	41	11	15	48.49
63	Middle Atlantic Coastal Plain	793	6.7	2.3	6.8	1.8	15	83	22	26.18
64	Northern Piedmont	208	21	5.8	10	1.9	39	17	15	76.28
65	Southeastern Plains	121	7.4	2.2	5.5	1.5	19	15	7.2	27.52
66	Blue Ridge	121	11	3.2	3.0	1.3	23	3.4	6.0	40.62
67	Ridge and Valley	163	17	4.5	4.6	1.4	33	6.3	15	60.95
68	Southwestern Appalachians	151	13	3.2	2.5	1.3	42	3.2	11	45.62
69	Central Appalachians	193	16	5.6	4.6	1.3	34	4.5	33	62.96

Table 4. Predicted 10th percentile concentrations for conductivity (µS/cm), BLM GI water quality parameters (mg/L) and hardness in each Level III ecoregion in the continental U.S.										
		Unbiased log mean of 10th percentile concentrations								
Level III Ecoregion	Ecoregion Name	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness[1]
70	Western Allegheny Plateau	276	23	7.3	10	1.7	46	14	47	87.43
71	Interior Plateau	237	27	5.9	4.4	1.5	65	7.0	20	91.69
72	Interior River Valleys and Hills	326	34	12	11	2.4	87	17	46	134.2
73	Mississippi Alluvial Plain	255	14	5.5	17	2.6	44	22	13	57.55
74	Mississippi Valley Loess Plains	102	10	2.9	3.9	1.6	38	4.5	8.4	36.89
75	Southern Coastal Plain	726	18	5.0	20	2.0	41	390	38	65.5
76	Southern Florida Coastal Plain	682	49	6.6	25	2.4	103	43	15	149.56
77	North Cascades	93	6.5	2.3	2.3	0.64	27	1.2	4.9	25.68
78	Klamath Mountains	156	8.7	4.6	4.0	0.66	44	2.1	3.5	40.61
79	Madrean Archipelago	625	42	11	45	2.8	92	39	78	150.1
80	Northern Basin and Range	298	26	8.2	20	2.7	89	15	24	98.62
81	Sonoran Basin and Range	991	64	24	115	4.4	121	131	192	258.4
82	Laurentian Plains and Hills	104	4.8	0.78	2.5	0.48	12	2.7	4.4	15.198
83	Eastern Great Lakes and Hudson Lowlands	294	34	6.8	21	1.3	61	40	26	112.88
84	Atlantic Coastal Pine Barrens	261	7.4	2.7	10	1.7	23	16	11	29.57

[1] Water Hardness calculated as equivalents $CaCO3 = 2.5 (Ca^{2+}) + 4.1 (Mg^{2+})$

2.2.3.3 Confirmation of Results

To confirm the results of the geostatistical predictions, a number of comparisons were made between the ecoregional average predictions and averages based directly on the data. For each GI parameter, we compared the ecoregional average predictions against the corresponding averages calculated from the data for each ecoregion.

Scatter plot matrices provide a visual presentation of the correlations between different parameters. Scatter plot matrices were developed for the ecoregional averages of conductivity and GI parameters. Figure 6 is the scatter plot matrix for ecoregional averages based on the data, and Figure 7 is the scatter plot matrix for ecoregional averages based on the geostatistical predictions. Comparison of

these figures shows that the predicted averages capture the same trends in terms of distributions and parameter correlations as those that are found for the ecoregional data. The similarities in the distributions and correlation structures between the ecoregional averages in Figures 6 and 7 demonstrate that the geostatistical ecoregion predictions are reasonable.

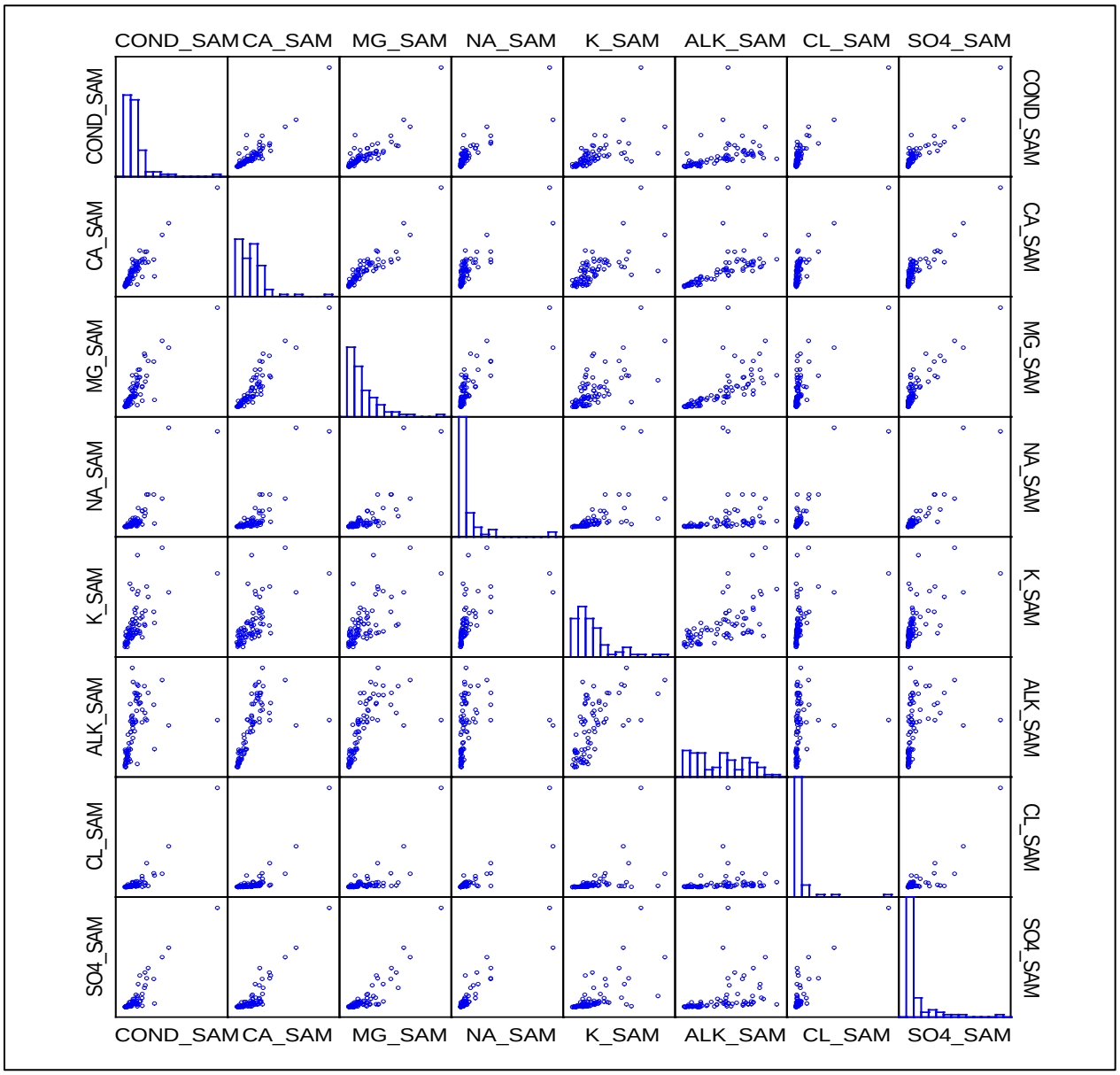

Figure 6. Scatter plot matrix of ecoregional average 10th percentiles of data for conductivity (COND_SAM) and GI parameters (calcium=CA_SAM, magnesium=MG_SAM, sodium=NA_SAM, potassium=K_SAM, alkalinity=ALK_SAM, chloride=CL_SAM, sulfate=SO4_SAM)

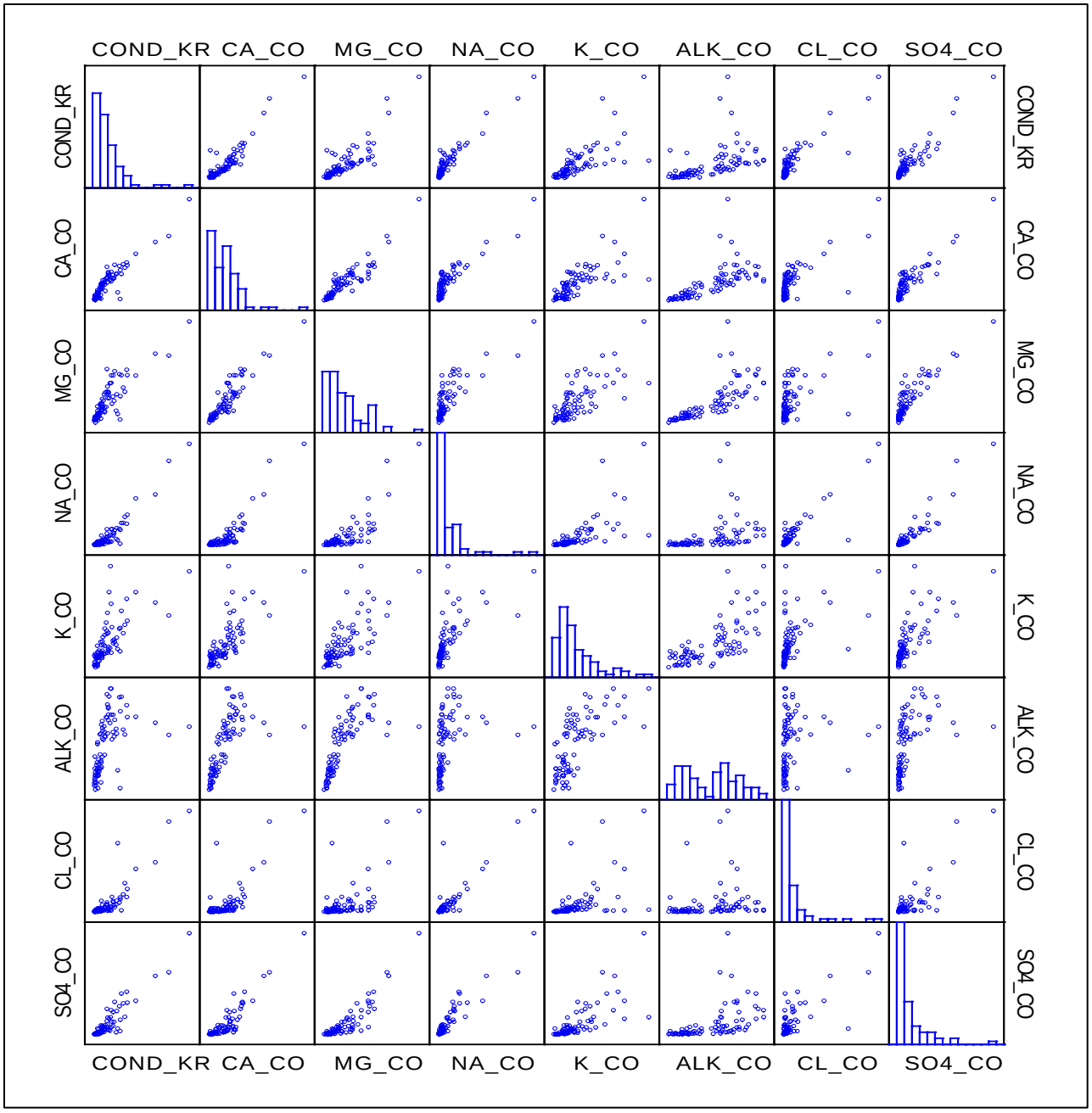

Figure 7. Scatter plot matrix of ecoregional average 10th percentiles of geostatistical predictions of conductivity (COND_KR) and BLM GI parameters (calcium=CA_CO, magnesium=MG_SCO, sodium=NA_CO, potassium=K_CO, alkalinity=ALK_CO, chloride=CL_CO, sulfate=SO4_CO)

In addition to scatter plots, correlation coefficient matrices between the parameters in each of the two data sets were generated. The Spearman (rank order) correlation coefficients for data-based ecoregional averages are presented in Table 5; correlation coefficients for ecoregional average geostatistical predictions are presented in Table 6. Although not identical, the correlation coefficients are similar between the two datasets, again demonstrating that the geostatistical predictions are reasonable.

Table 5. Spearman rank correlation matrix for unbiased log means of 10th percentile concentrations measured in Level III ecoregions

	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate
Conductivity	1							
Calcium	0.895	1						
Magnesium	0.877	0.927	1					
Sodium	0.865	0.819	0.813	1				
Potassium	0.823	0.758	0.73	0.859	1			
Alkalinity	0.769	0.881	0.898	0.698	0.679	1		
Chloride	0.815	0.774	0.702	0.855	0.788	0.595	1	
Sulfate	0.894	0.864	0.85	0.883	0.786	0.725	0.744	1

Table 6. Spearman rank correlation matrix for unbiased log means of 10th percentile predicted (kriged/cokriged) concentrations in Level III ecoregions

	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate
Conductivity	1							
Calcium	0.872	1						
Magnesium	0.843	0.924	1					
Sodium	0.87	0.82	0.778	1				
Potassium	0.842	0.803	0.753	0.836	1			
Alkalinity	0.747	0.868	0.901	0.672	0.745	1		
Chloride	0.829	0.715	0.592	0.826	0.725	0.461	1	
Sulfate	0.889	0.873	0.875	0.893	0.831	0.751	0.715	1

As a final test of the accuracy of the geostatistical predictions, we regressed the ecoregional averages based on the geostatistical predictions against the ecoregional averages based on the data. Scatter plots and fitted regression lines are shown for each of the parameters: conductivity (Figure 8), alkalinity (Figure 9), calcium (Figure 10), magnesium (Figure 11), sodium (Figure 12), potassium (Figure 13), sulfate (Figure 14), and chloride (Figure 15). Statistics for the linear regressions are provided in Table 7. For each of the parameters, the predicted and data-based ecoregional averages are significantly correlated. In each case, the linear regression coefficient is nearly 1.0, with a highly significant P value. As with the previous comparisons, the linear regression results demonstrate that the accuracy of the geostatistical predictions is high.

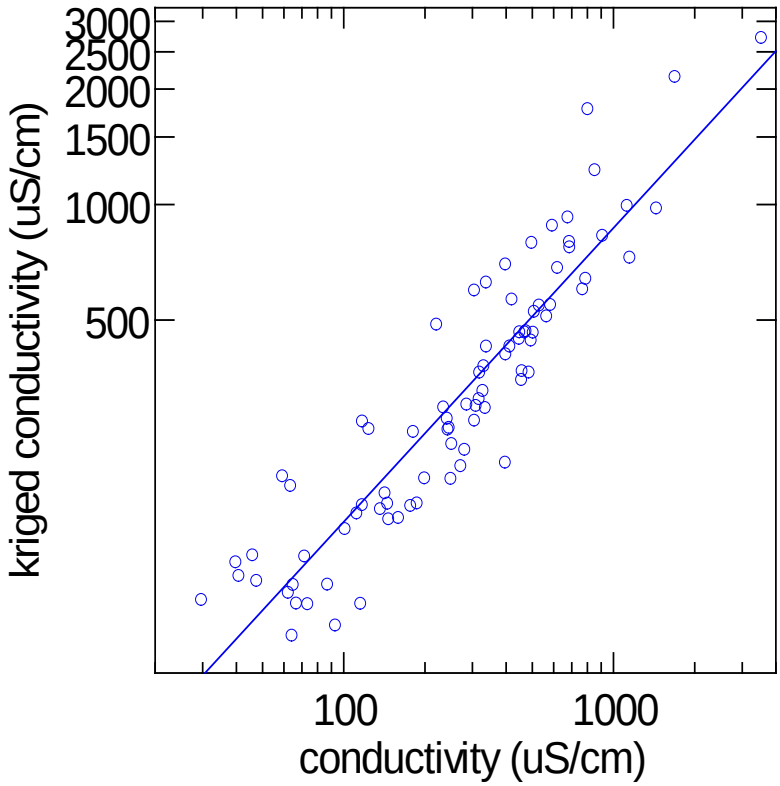

Figure 8. Ecoregional averages of kriged 10th percentiles of conductivity versus data

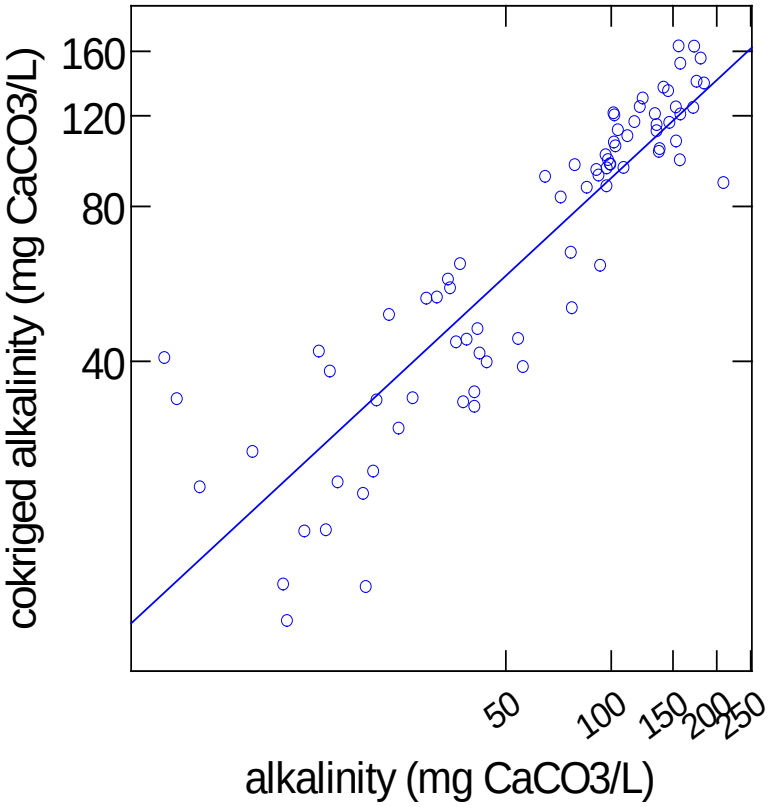

Figure 9. Ecoregional averages of cokriged 10th percentiles of alkalinity versus data

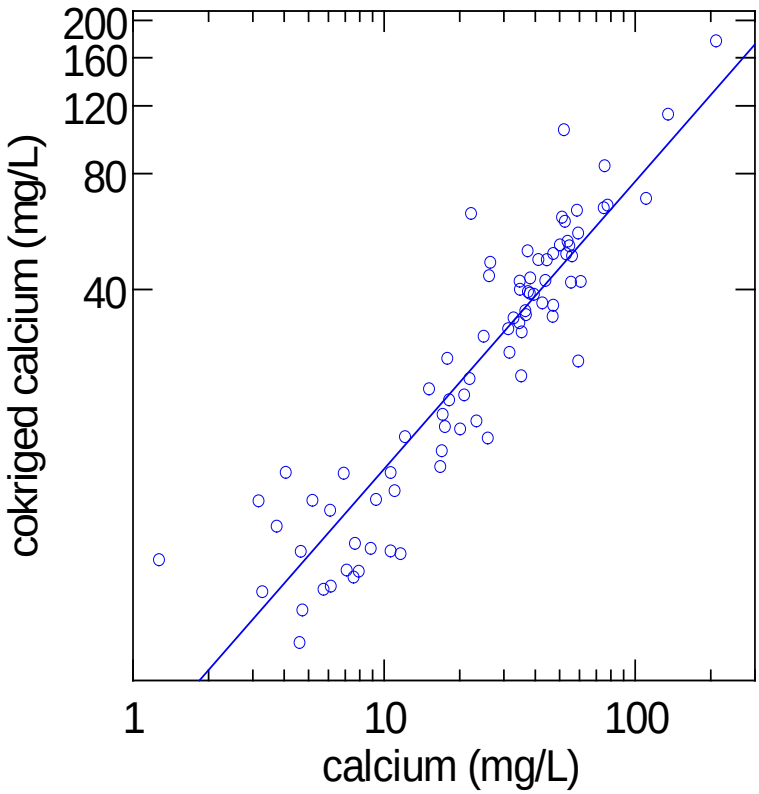

Figure 10. Ecoregional averages of cokriged 10th percentiles of calcium versus data

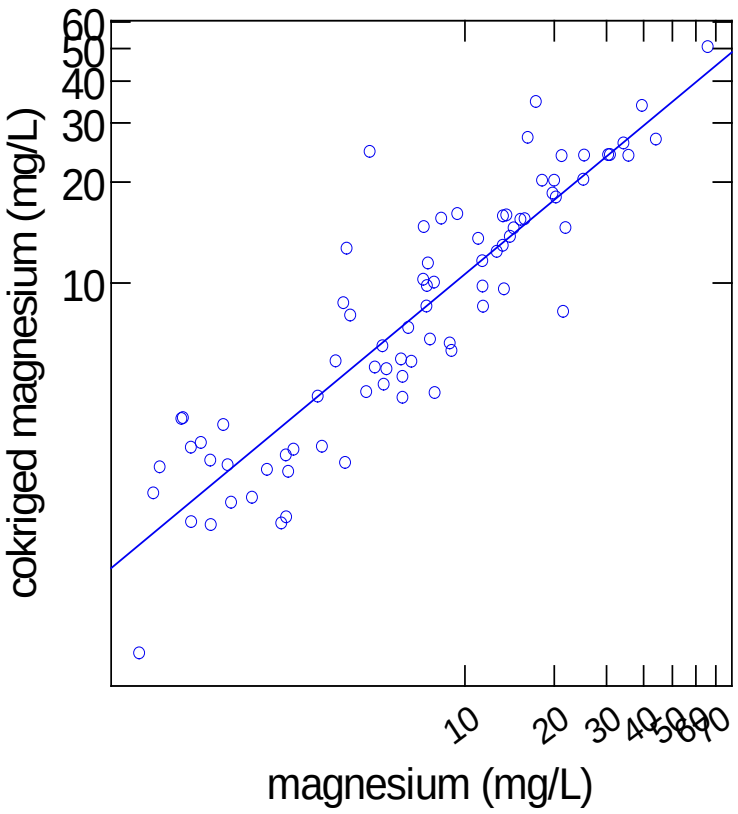

Figure 11. Ecoregional averages of cokriged 10th percentiles of magnesium versus data

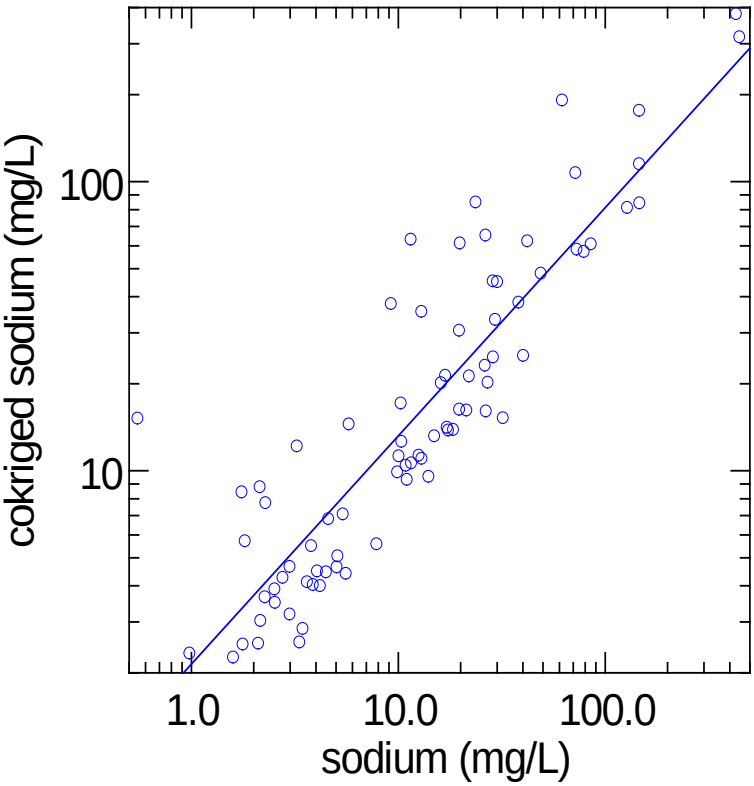

Figure 12. Ecoregional averages of cokriged 10th percentiles of sodium versus data

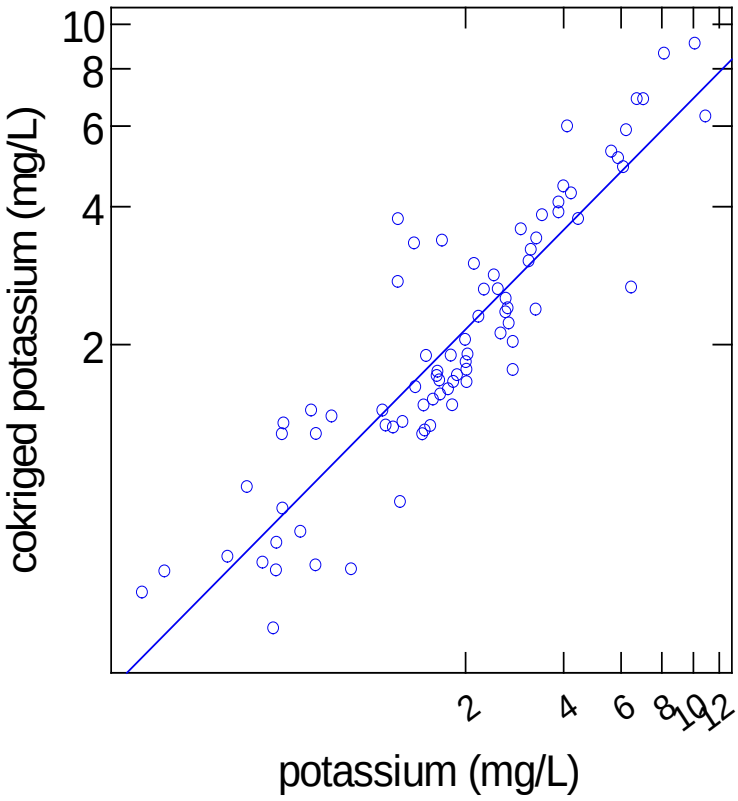

Figure 13. Ecoregional averages of cokriged 10th percentiles of potassium versus data

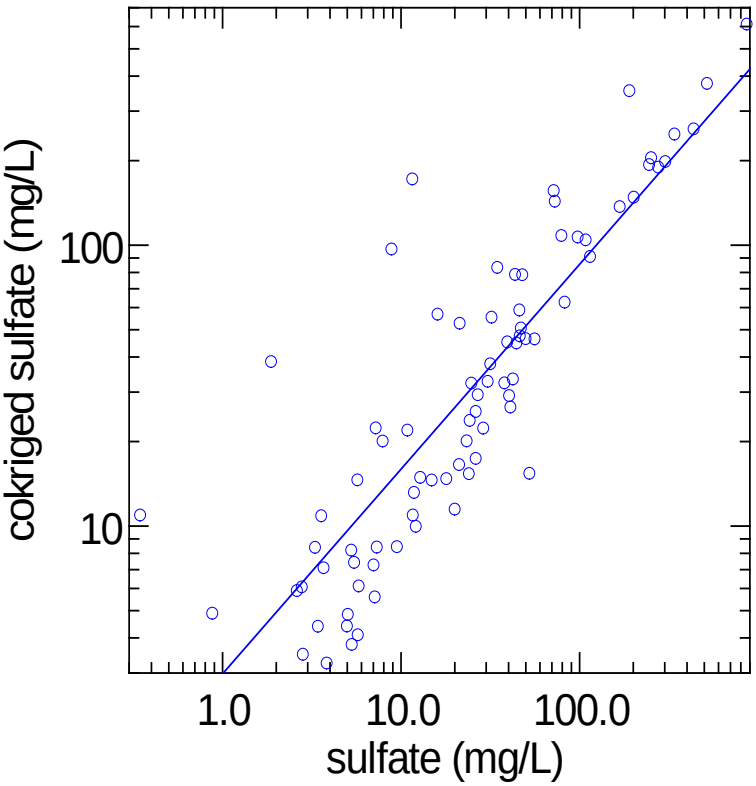

Figure 14. Ecoregional averages of cokriged 10th percentiles of sulfate versus data

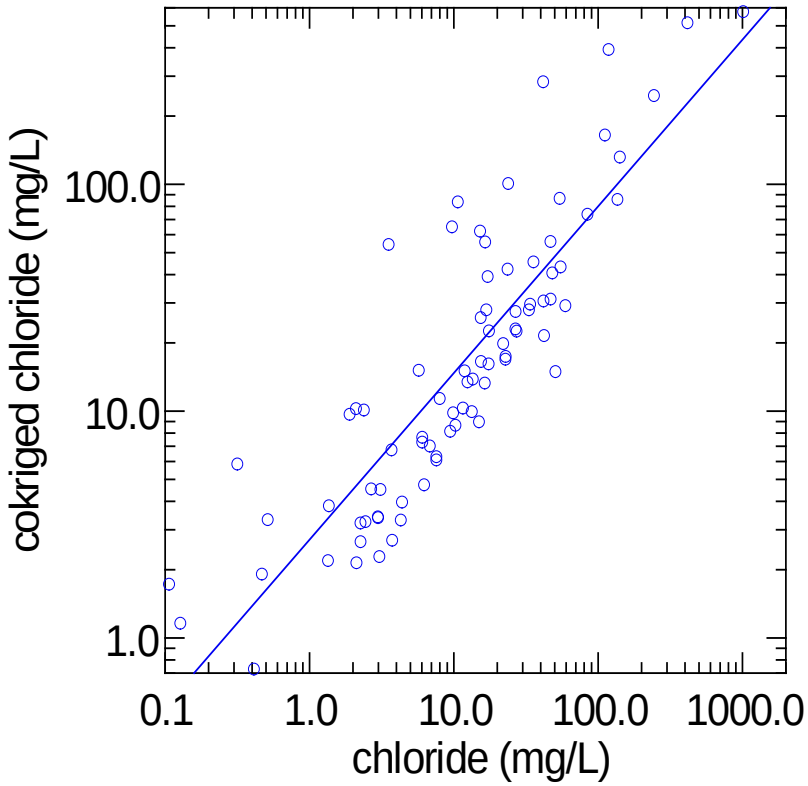

Figure 15. Ecoregional averages of cokriged 10th percentiles of chloride versus data

Table 7. Correlation coefficients and linear regression (LR) statistics between ecoregional average 10th percentiles of data and geostatistical predictions

Parameter	Correlation coefficient, r	LR coefficient	P (2 Tail)	LR constant	P (2 Tail)
Conductivity	0.908	0.98	<0.001	-20.992	0.495
Alkalinity	0.915	1.209	<0.001	-12.161	0.028
Calcium	0.922	1.061	<0.001	-1.994	0.365
Magnesium	0.885	1.097	<0.001	-1.169	0.22
Sodium	0.93	1.101	<0.001	-4.78	0.169
Potassium	0.906	1.073	<0.001	-0.184	0.297
Sulfate	0.865	1.264	<0.001	-13.599	0.041
Chloride	0.839	1.039	<0.001	-7.271	0.386

2.2.3.4 Conclusions for Selection of Water Quality Parameters

In this section we used geostatistics to estimate an intermediate step in generating missing GI parameter values based on geography. We supplemented the geostatistical approach by adding conductivity as an additional explanatory variable to generate a more robust spatial estimate of the GI water quality inputs for the BLM because conductivity is one of the most widely monitored water quality indicators in the U.S. and correlates well with GIs. In section 3, these estimates are further refined by stream order. We present here the average predicted 10^{th} percentile concentrations for the BLM GI water quality parameters, as presented in Table 4 by ecoregions. Because they are based on the 10^{th} percentiles of the daily average data from each USGS monitoring station, they are expected to yield copper criteria that are reasonably protective of aquatic life when applied as missing data for parameters in the BLM model. These data could also be used to fill in missing water chemistry parameters in the application of other metal BLM models. The most appropriate parameter selection however would include consideration of stream order in GIs estimates. Section 3 presents further refinements of estimates of the GI parameters by stream order and EPA's recommendations for default GI parameters for the BLM when data are lacking.

As with any estimate or prediction, it is appropriate to seek alternative estimates for the purpose of comparison or confirmation. If conductivity data are available for the site, either site-specific measurement data or data of opportunity from a database such as the NWIS, the regressions in EPA (2008; Appendix C) can be used to make independent estimates of the missing BLM water quality parameters. If the regression projections differ from the geostatistical average predictions, the lower (more conservative) estimate is recommended for application to ensure protection of aquatic life. As always, users of the BLM should be also encouraged to sample the water body of interest and to analyze for the constituent (parameter) concentrations as a basis for determining reliable BLM inputs.

2.2.3.5 Guidance Regarding Selection of Water Quality Parameters: pH and DOC

Although the geostatistical and regression-based approaches can be used to reliably estimate GI parameters used as BLM inputs, the same approaches do not produce accurate site-specific estimates for the two most important BLM inputs: pH and DOC. The BLM is less sensitive to the GI parameters than to pH and DOC predicting site-specific criteria for copper. Since our analysis indicates that there is little or no trend in relationships between conductivity and pH, and direct kriging produced similarly ambiguous predictions, site-specific data for pH must be used for BLM application at a site.

For DOC, analysis of NWIS data indicated a weak relationship with conductivity, so the regression approach is not appropriate for this parameter. In 2008, EPA recommended use of the ecoregional DOC concentration percentiles tabulated by EPA for the *Development of National Bioaccumulation Factors Technical Support Document* (USEPA, 2003) because they appeared to offer reasonable estimates of lower percentile DOC concentrations, and were based on substantially more DOC data than were available in the NWIS. In Section 4 of this report we further tested these ecoregional DOC concentrations for use in the BLM where site-specific data are not available.

3 USING STREAM ORDER TO REFINE PREDICTION OF GI PARAMETERS

The following section discusses how stream order was used to address anthropogenic impacts. The goal is to provide BLM users with tables of appropriately protective estimates of GI parameters, building on the ecoregional work described in Section 2.

Estimations of values for the GI parameters (alkalinity, calcium, magnesium, sodium, potassium, sulfate, and chloride) tend to vary regionally. As demonstrated in Section 2, the spatial variation of these factors is generally known or at least predictable, and therefore spatial or geographic analysis of data can be used to estimate GI input parameter values. However, these values also vary due to anthropogenic impact. In the case of conductivity and GI parameters, a positive correlation between ion concentrations and measures of human activity, such as population density, urban and agricultural land use, road density, point and nonpoint pollutant sources, among other activities, is expected and may confound the pattern of geographic variability both within and between ecoregions.

One way to account for surface water quality variability within ecoregions is to distinguish water bodies according to the Strahler stream order (SO). The SO is used to define stream size based on a hierarchy of tributaries (Strahler, 1952; 1957) and may range from 1st order (a stream with no tributaries) to 12th order (the Amazon, at its mouth). First through 3rd order streams are called headwater streams (source waters of a stream). Over 80% of Earth's waterways are headwater streams (Strahler, 1957). A stream that is 7th order or larger constitutes a river. For example, the Ohio River is 8th order and the Mississippi River is 10th order. According to the River Continuum Concept, changes in water quality are commonly observed between the upper, middle, and lower reaches of a stream (FISRWG, 1998; Ward, 1992; USEPA, 2015).

In this section we consider variability in GIs by determining the SO of each surface water sampling location in the USGS NWIS[2] database, and explore methods of incorporating SO variation in the parameter estimates. Tables are provided in this section showing tabulations of parameter estimates based upon both ecoregion and SO to maximize the accuracy of estimated input parameters.

3.1 Determining SO of NWIS Surface Water Sampling Locations

GIS was used to determine the SO of each NWIS surface water sampling location. Flowlines and catchments with SO were obtained from the NHD-Plus V2 geospatial hydrologic framework (McKay et al., 2012).[3] The point locations corresponding to the latitude-longitude coordinates of the NWIS sampling stations were snapped to the NHD-Plus flowlines using ArcGIS. A spatial join was then performed between these shapefiles and the NHD-Plus flowlines to link stream order to the sampling locations. Some of the NHD-Plus flowlines did not have SO data associated with the record. When a sampling location occurred on a flowline that didn't have a SO, the SO from the catchment was used. When the catchment also did not have a SO, the SO of the nearest stream was applied. SO was added as an attribute to the information for each station in the database.

[2] http://waterdata.usgs.gov/nwis
[3] http://www.horizon-systems.com/NHDPlus/NHDPlusV2_home.php

3.2 Estimating BLM Parameters for Ecoregions and SO

Estimated (10th percentile) BLM water quality parameters were presented in Section 2 for 84 Level III ecoregions of the continental U.S. In the work presented here, the parameter estimates were recalculated for individual SOs or ranges (groups) of SOs within each ecoregion.

3.3 Results

The distribution of NWIS sampling locations by SO is presented in Figure 16. The largest proportion of sampled locations (78%) was found to be in SO 1 through 4.

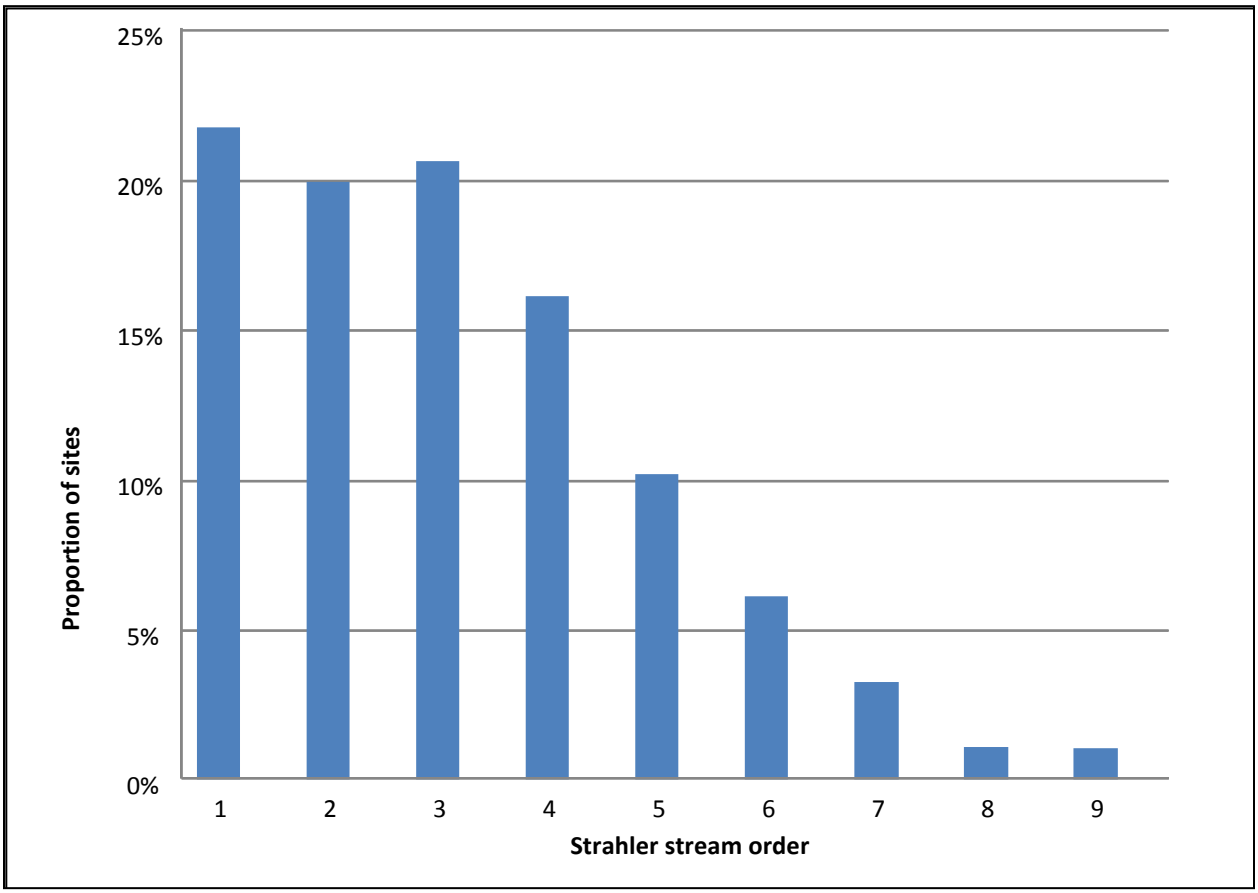

Figure 16. Distribution of NWIS surface water sampling locations by SO

3.3.1 Dependence of Ecoregional Parameter Estimates on SO

Box plots were constructed to examine how the GIs estimates varied with SO. Box plots of conductivity (Figure 17), alkalinity (Figure 18), calcium (Figure 19), magnesium (Figure 20), sodium (Figure 21), potassium (Figure 22), sulfate (Figure 23) and chloride (Figure 24) all show a general increase in the magnitude of the estimate with SO. This trend was most apparent and consistent when comparing medium stream orders (SO 4-6) to higher stream orders (SO ≥7). In addition, the upper quartile parameter estimates were generally higher in SOs 4 through 6 than in lower order streams (SO ≤3). Based upon these trends, we grouped the estimates for each parameter by SO: 1 through 3 (headwater

streams), 4 through 6 (mid-reaches) and 7 through 9 (rivers). There were no data for rivers with SO>9. Grouping simplified the presentation of results and improved the robustness of the parameter estimates, without losing significance of the SO trends. Parameter estimates for these three SO groups are included in the box plots in Figures 18 through 24, labeled as "13," "46," and "79." The classes depicted as 13, 46, and 79 reflect groupings according to SO (i.e., 1 through 3, 4 through 6, and 7 through 9).

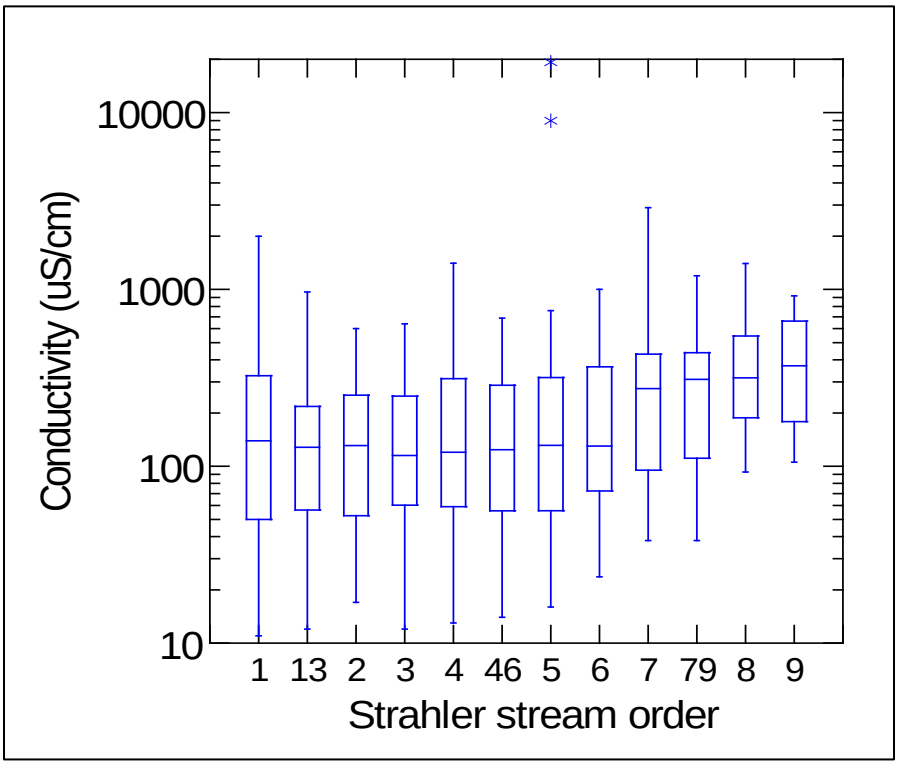

Figure 17. Box plot of estimated ecoregional conductivities as a function of SO

Note: Classifications depicted as 13, 46, and 79 reflect groupings according to stream order (i.e., 1 through 3, 4 through 6, and 7 through 9) as described in the text. For box plots, the bottom and top of each "box" displays the 25th and 75th percentile concentrations defined as the interquartile range (IQR) (i.e., the box contains 50% of the data values), respectively. The median is displayed as the horizontal line within the box. The "whiskers" show the relative distribution of data points outside of the IQR and represent 1.5 times the IQR. Data not included between the whiskers are plotted as outliers with a star/asterisk.

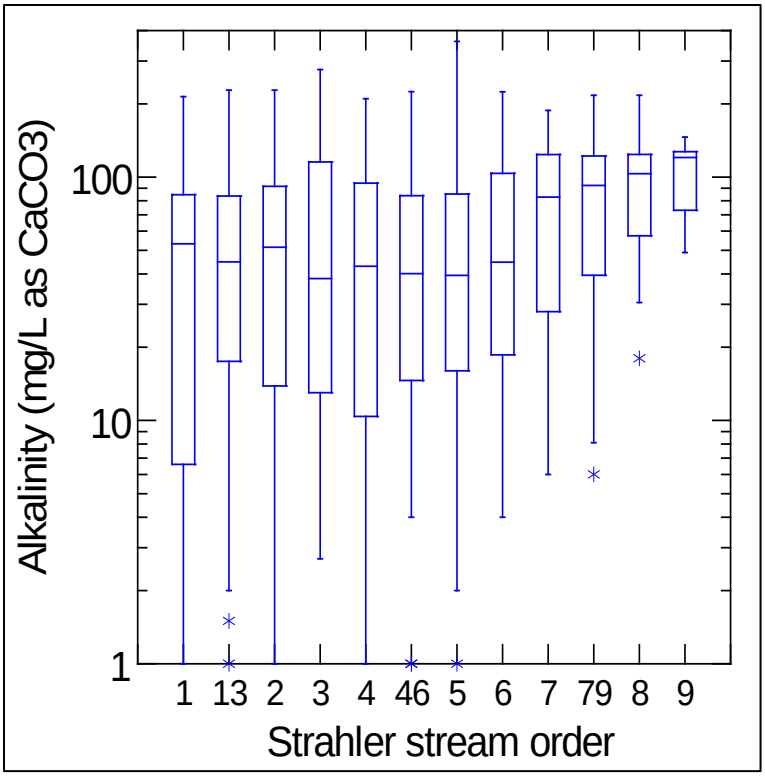

Figure 18. Box plot of estimated ecoregional alkalinity concentrations as a function of SO

(Refer to note in Figure 17.)

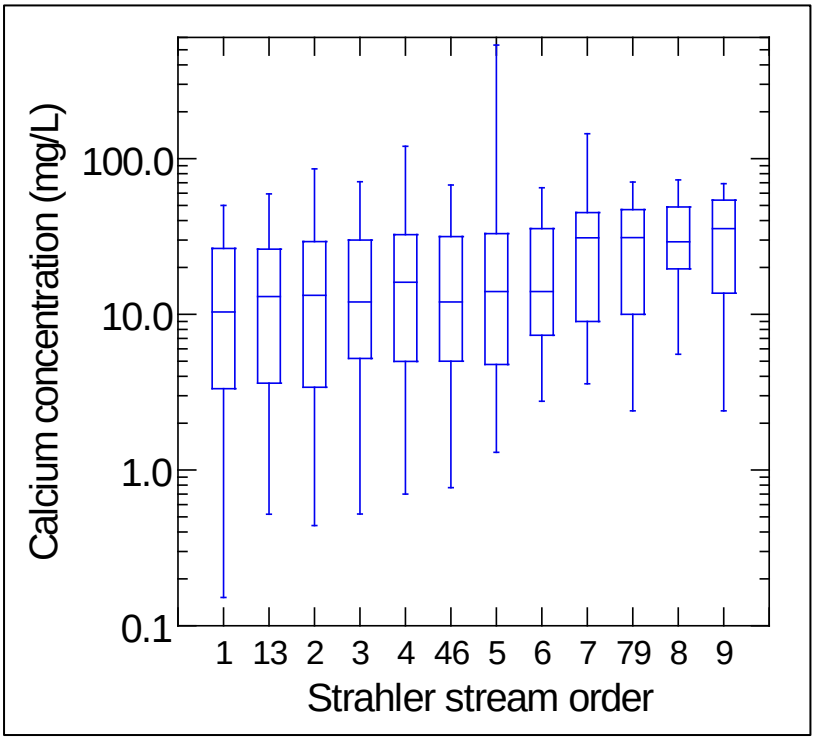

Figure 19. Box plot of estimated ecoregional calcium concentrations as a function of SO

(Refer to note in Figure 17.)

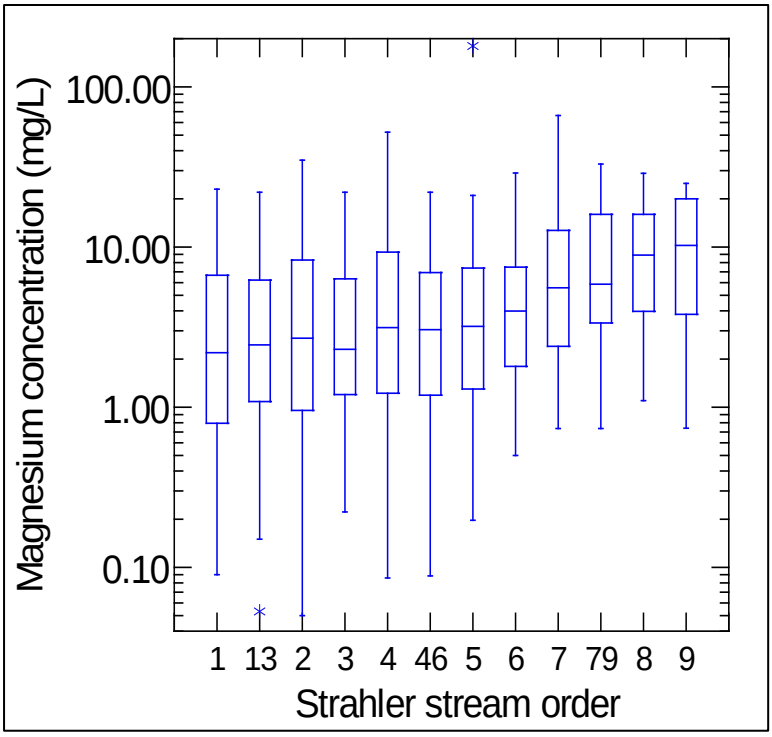

Figure 20. Box plot of estimated ecoregional magnesium concentrations as a function of SO

(Refer to note in Figure 17.)

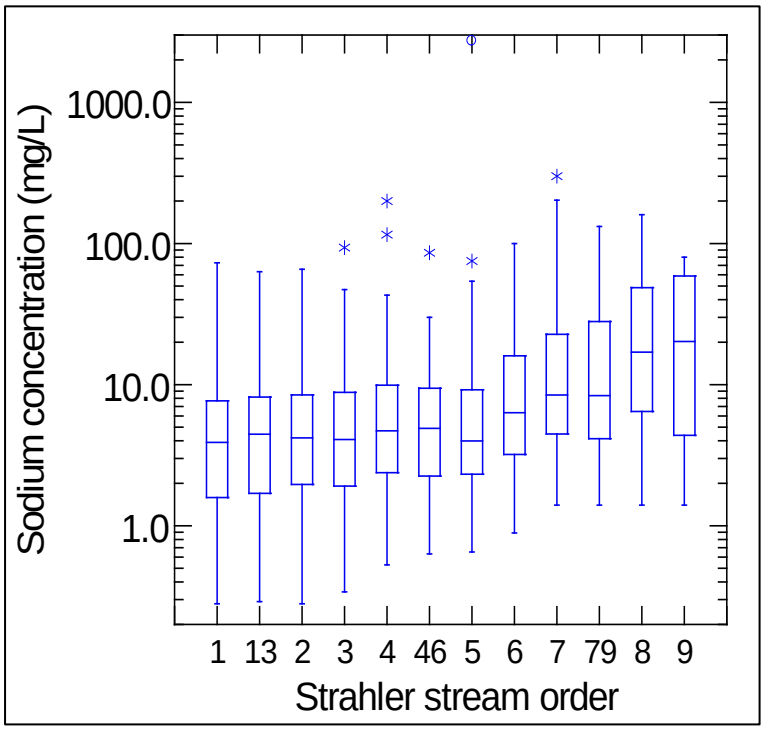

Figure 21. Box plot of estimated ecoregional sodium concentrations as a function of SO

(Refer to note in Figure 17.)

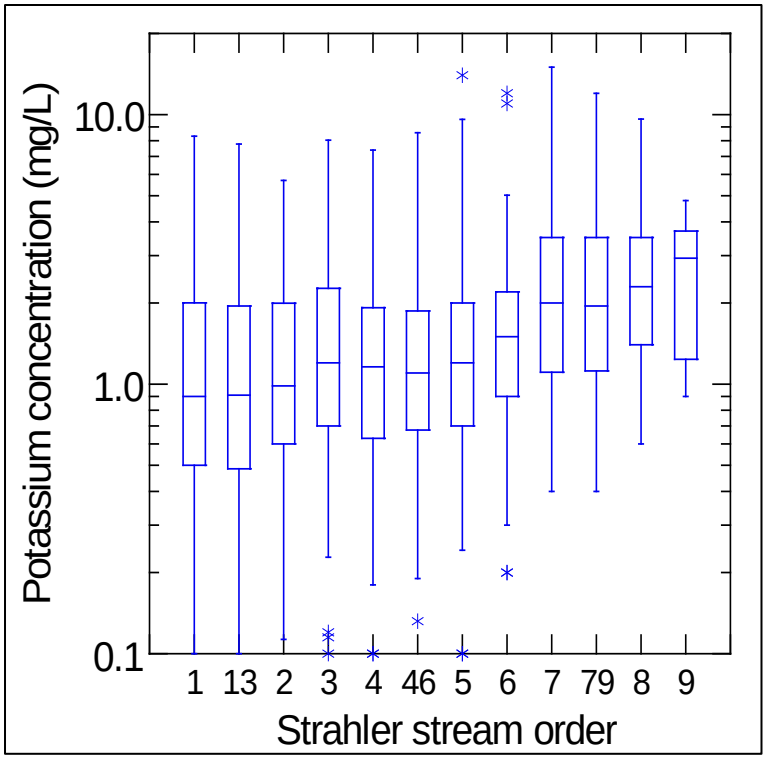

Figure 22. Box plot of estimated ecoregional potassium concentrations as a function of SO

(Refer to note in Figure 17.)

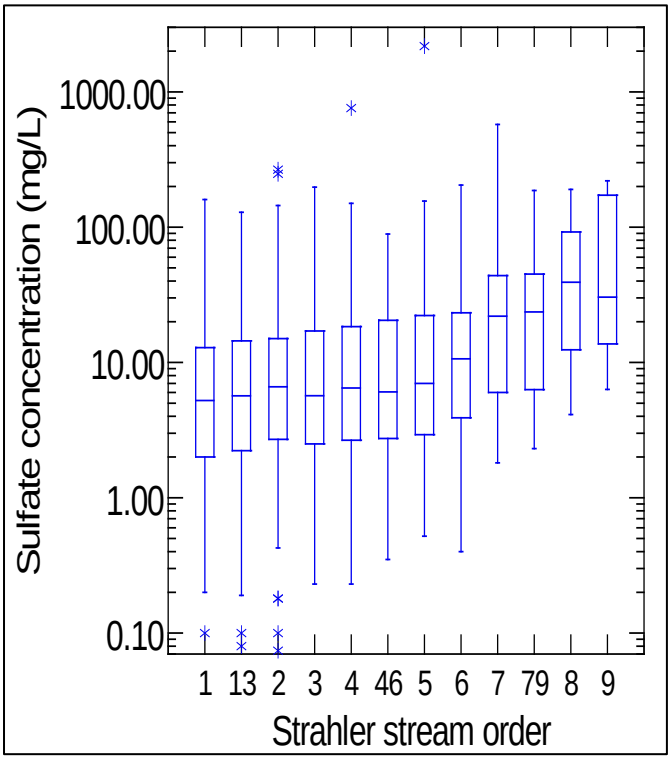

Figure 23. Box plot of estimated ecoregional sulfate concentrations as a function of SO

(Refer to note in Figure 17.)

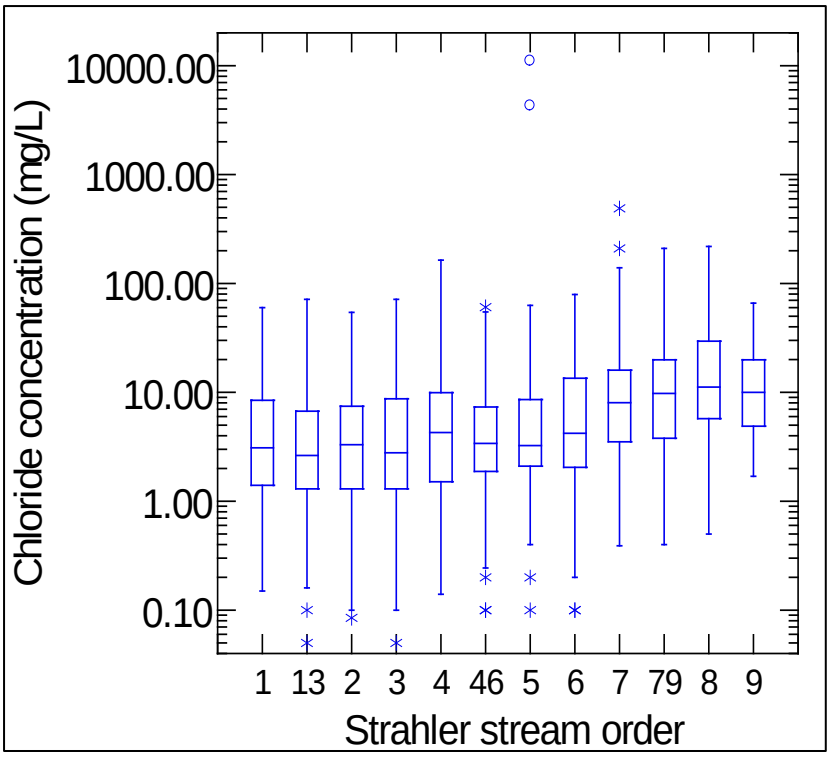

Figure 24. Box plot of estimated ecoregional chloride concentrations as a function of SO

(Refer to note in Figure 17.)

The range of values across SOs overlap greatly due to the inclusion of data across ecoregions. Tenth percentile estimates of conductivity increase with SO group in 58% of ecoregions when comparing low to medium SO groups, and 84% of ecoregions when comparing medium to high SO groups. The same trend was evident for the GIs. For example, 10th percentiles of calcium increased with SO group in 68% of ecoregions for low versus medium SO and 83% of ecoregions for medium versus high SO. In general, parameter estimates (10th percentiles of conductivity and ion concentrations) increased with SO, and the increase was most apparent and consistent for higher SOs (SO ≥7).

3.3.2 SO-Based Parameter Estimates

Tenth percentile parameter estimates for conductivity, GIs and hardness are grouped by SO and Level III ecoregions in Tables 8 through10. Tenth percentile parameter estimates for SOs 1 through 3 are presented in Table 8, for SOs 4 through 6 are presented in Table 9, and SOs 7 through 9 are presented in Table 10. The tables include the sample size for instances in which parameter estimates are highly uncertain due to limited data, i.e., cases where sample size is <10. Water quality data were limited in four ecoregions (11, 16, 49, and 78) for SO group 1 through 3, in Ecoregion 76 for SO group 4 through 6, and in 28 ecoregions for SO group 7 through 9. With the exception of the specific ecoregions and SO classes where data are limited, the parameter estimates in Tables 8 through 10 are recommended as improved default values for use in the BLM when data are not available for a location in a specific Level III ecoregion and SO group.

Table 8. Recommended 10th percentile conductivity, GIs, and hardness estimates for SO Group 1 through 3 (number of stations shown in parentheses if n<10)

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
1	58	6.0	0.8	1.3	0.1	44	0.6	1.1	18.28
2	74	8.8	2.8	3.9	0.5		2.8	3.3	33.48
3	68	9.9	3.8	5.6	1.5		2.3	1.5	40.33
4	16	1.0	0.2	1.8	0.2		0.5	0.2	3.32
5	28	0.6	0.1	0.3	0.1	38	0.1	0.1	1.91
6	279	3.6	8.2	8.4	0.9	73	8.9	7.2	42.62
7	164	19	6.0	14	1.8	120	8.6	6.6	72.1
8	157	29	4.3	10	1.5	70 (7)	2.6	0.4	90.13
9	55	4.4 (8)	0.9 (8)	2.3 (9)	0.4 (9)	35 (2)	0.2	0.2	14.69
10	137	24.0	9.4	10.2	1.4	127	4.6	11	98.54
11	88	8.6 (2)	3.2 (2)			169 (2)			34.62
12	133	13	2.0	6.1	0.8	35	1.4	3.7	40.7
13	109	9.4	1.6	2.7	0.6	45	0.5	3.7	30.06
14	967	15	2.8	6.0	3.0	90 (7)	2.7	6.3	48.98
15	24	3.1	0.8	0.9	0.4	9.0	0.2	1.3	11.03
16	21								
17	93	6.9	1.6	1.5	0.5	31	0.3	3.0	23.81
18	92	22	6.3	4.7	0.9		3.3	7.3	80.83
19	76	59	11	5.1	0.6	96	2.5	44	192.6
20	189	59	12	19	1.4	157	6.7	129	196.7
21	37	3.5	0.7	0.8	0.3	18	0.2	1.9	11.62
22	115	13	1.1	2.3	0.8	55	1.2	7.2	37.01
23	62	6.3	1.8	3.7	0.7	20	0.8	4.3	23.13
24	453	43	7.9	35	3.4	32	20	74	139.89
25	194	43	11	31	3.7	228	7.3	35	152.6
26	199	18	3.0	63	3.4	53	3.6	11	57.3
27	293	21	5.0	9.9	1.5	122	5.1	13	73
28	346	50	8.2	4.4	0.8	125	1.5	22	158.62
29	217	30	4.0	17	2.9	74	8.2	9.9	91.4
30	189	25	1.8	1.6	0.9	99	2.6	4.9	69.88
31	639	48	5.5	47	2.9		71	51	142.55
32	183	26	1.7	5.9	1.9	52	5.1	15	71.97
33	132	24	2.1	5.8	2.3	29	8.1	6.0	68.61
34	141	13	2.3	9.8	2.6	44	13	4.7	41.93
35	25	0.9	0.5	1.9	0.2	5.0	3.0	1.7	4.3
36	19	0.9	0.7	0.9	0.3	4.0	1.3	2.0	5.12
37	107	23	3.5	23	3.3	36	2.5	3.5	71.85
38	51	0.9 (7)	0.6 (7)	0.63 (7)	0.6 (7)	35	1.1	1.8	4.71
39	172	26	1.9	1.3	0.7	62	2.0	4.2	72.79

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
40	223	20	4.6	8.4	2.8	46	5.2	23	68.86
41	93	13	3.8	0.4	0.1	60 (9)	0.1	1.6	48.08
42	256	23	7.3	7.4	1.6	91	1.3	14	87.43
43	327	17	16	26	3.7	144	2.7	119	108.1
44	156	21	3.2	6.5	4.7	80 (2)	0.7	5.5	65.62
45	44	2.7	0.8	2.5	1.2	12	2.4	1.9	10.03
46	400	32	13	15	7.8	94	5.4	60	133.3
47	380	41	11	4.8	1.3	83	12	15	147.6
48	295	31	15	4.9	2.5		4.3	13	139
49	402 (4)					227 (1)			
50	69	5.1	1.0	1.2	0.3	32	0.4	1.2	16.85
51	137	28	11	2.0	0.9	53 (4)	3.3	4.5	115.1
52	432	42	12	2.9	0.7	75	4.8	12	154.2
53	502	27	8.0	11	2.0	42	25	22	100.3
54	574	51	22	8.4	1.2	202	28	44	217.7
55	420	40	11	7.8	1.5	130	19	16	145.1
56	219	28	7.4	3.6	0.9	79	6.3	16	100.34
57	446	35	10	6.8	2.3	84	19	33	128.5
58	25	1.2	0.4	0.3	0.2	1.5	0.4	4.2	4.64
59	69	3.6	1.1	6.2	0.9	3.0	10	5.8	13.51
60	61	4.7	1.2	1.5	0.4	17	1.5	6.6	16.67
61	131	13.0	2.5	8.0	1.1	33	7.8	8.3	42.75
62	32	1.8	0.7	0.8	0.3	2.0	0.8	4.4	7.37
63	108	1.1	0.8	2.8	0.8	5.0	5.5	3.2	6.03
64	134	12.0	4.6	7.0	1.3	19	10	11	48.86
65	25	0.7	0.5	1.5	0.3	3.0	2.6	0.7	3.8
66	12	0.5	0.3	0.7	0.3	3.0	0.7	1.1	2.48
67	81	4.7	2.0	2.0	0.7	3.0	1.8	7.1	19.95
68	32	1.9	1.2	1.3	0.7	3.0	1.4	3.6	9.67
69	36	2.8	0.8	0.4	0.4	4.0	1.0	8.1	10.28
70	125	5.6	4.0	2.7	1.4	12	2.4	16	30.4
71	179	10	1.8	1.1	0.7	54	2.2	2.8	32.38
72	180	17	5.3	5.2	1.7	69	3.5	21	64.23
73	102	6.9	2.7	3.4	2.0	31	2.0	2.9	28.32
74	52	3.2	1.4	2.5	1.5	9.3	2.6	3.9	13.74
75	75	6.6	1.5	4.5	0.5	9.0	9.0	1.5	22.65
76	430	41	1.8	7.8	0.3	116	30	0.1	109.88
77	14	1.6	0.4	0.4	0.2	7.0	0.2	0.7	5.64
78	19						2.1		
79	340	30	6.2	25	2.8	92 (2)	13	23	100.42
80	78	6.3	1.1	4.3	2.2	24	0.2	2.5	20.26
81	203	19	2.4	10	2.8	52	2.6	6.1	57.34

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
82	37	1.5	0.5	4.3	0.2		6.6	1.8	5.8
83	198	16	3.9	5.0	1.0	51	21	29	55.99
84	50	0.8	0.6	2.8	0.6	1.0	5.0	4.4	4.46

Table 9. Recommended 10th percentile conductivity, GIs, and hardness estimates for SO group 4 through 6 (number of stations shown in parentheses if n<10)

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
1	52	3.6	1.0	2.0	0.2	15	1.6	2.2	13.1
2	49	5.3	1.2	1.3	0.1	16	0.8	1.8	18.17
3	62	7.1	2.5	4.3	0.8	29	4.6	2.8	28
4	35	3.5	1.0	2.8	0.4	16	0.8	0.8	12.85
5	18	0.9	0.1	0.6	0.2	5.0	0.4	0.4	2.66
6	316	9.1	4.8	5.4	1.0	32	2.3	4.1	42.43
7	67	6.5	2.5	2.9	0.9	33	1.7	3.2	26.5
8	93	9.0	1.5	8.4	1.0	17	3.2	6.0	28.65
9	52	5.5	0.8	2.4	0.5	22	0.9	2.2	17.03
10	83	8.6	3.2	4.0	0.9	33	1.4	3.1	34.62
11	52	3.7	0.8	1.6	0.7	16	0.3	0.7	12.53
12	95	13	2.5	4.9	1.2	40	2.2	3.8	42.75
13	124	12	3.4	9.6	1.8	68	3.9	8.1	43.94
14	688	58	13	86	7.9	225	55	86	198.3
15	34	3.5	1.1	0.9	0.4	16	0.2	1.2	13.26
16	22	2.4	0.4	1.3	0.4	10	0.2	0.6	7.64
17	123	10	2.5	1.4	0.6	44	0.5	2.2	35.25
18	145	15	4.0	7.1	0.9	57	1.4	16	53.9
19	135	34	9.1	6.7	1.3		7.3	9.5	122.31
20	260	38	9.6	16	1.5	107	4.0	55	134.36
21	74	6.5	1.3	1.9	0.6	18	0.4	3.2	21.58
22	215	25	4.3	9.5	1.3	62	2.4	18	80.13
23	289	31	9.5	5.6	1.1	101	3.5	2.4	116.45
24	240	24	5.1	18	1.5	80	7.0	21	80.91
25	220	14	3.4	8.7	1.6	84	4.1	29	48.94
26	367	39	9.1	29	2.8	79	11	56	134.81
27	351	39	7.1	11	5.1	79	5.6	16	126.61
28	298	68	14	9.6	2.1	150	7.4	39	227.4
29	351	39	6.8	20	2.7	80	20	19	125.38
30	377	44	12	6.5	0.8	140	10	13	159.2
31	447	54	9.4	10	1.8	142	12	29	173.54
32	311	35	2.8	12	2.8	94	12	22	98.98
33	334	33	5.0	16	2.1	45	23	28	103
34	125	10	3.5	13	2.7	40	12.2	3.1	39.35

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
35	67	4.2	1.1	6.5	1.5	10	3.5	5.0	15.01
36	24	1.0	0.8	1.1	0.5	4.0	1.5	2.0	5.78
37	55	3.8	2.3	5.2	1.3	9.6	2.0	2.0	18.93
38	41	8.4	1.0	1.2	0.8	13	1.5	3.2	25.1
39	160	19	1.8	1.6	0.9	73	2.3	3.6	54.88
40	258	33	5.7	6.8	2.7	64	6.7	20	105.87
41	133	19	4.5	0.7	0.2		0.1	1.9	65.95
42	285	22	9.2	12	2.3	141	2.3	34	92.72
43	342	28	8.5	13	2.7	145	2.2	45	104.85
44	232	27	3.8	9.4	7.0	184 (1)	2.5	2.7	83.08
45	44	3.0	1.2	3.0	1.3	13	2.9	2.9	12.42
46	480	32	17	30	8.6	153	9.1	89	149.7
47	390	43	10	7.0	1.7	109	8.5	19	148.5
48	422	40	18	8.5	4.1	170	7.0	32	173.8
49	75	7.9	2.4	2.3	0.7	25	2.3	4.5	29.59
50	78	7.9	2.4	1.2	0.4	37	0.9	2.7	29.59
51	125	19	6.5	2.3	0.9	102	3.0	3.6	74.15
52	221	15	6.3	4.2	1.5	50	6.4	10	63.33
53	389	36	19	9.1	2.2	138	18	20	167.9
54	520	49	22	7.4	1.0	148	22	37	212.7
55	413	43	12	5.6	1.9	162	19	22	156.7
56	389	44	14	10	1.5	133	18	21	167.4
57	489	56	15	10	2.6	108	22	31	201.5
58	38	4.9	0.9	2.9	0.5	5.0	4.3	6.1	15.94
59	81	5.1	1.4	7.8	1.1	10	11	7.3	18.49
60	101	12	2.5	5.3	1.0	20	3.9	8.4	40.25
61	178	20	4.7	8.4	1.8	47	61	11	69.27
62	50	3.9	0.9	4.0	0.6	5.0	2.6	6.1	13.44
63	65	3.6	0.9	3.9	1.2	8.0	6.6	3.2	12.69
64	175	13	4.7	9.9	1.2	48	15	13	51.77
65	43	2.8	0.9	2.2	0.8	6.0	3.4	1.7	10.69
66	14	1.0	0.3	0.8	0.5	4.0	0.4	1.1	3.73
67	89	7.9	2.0	2.9	1.0	14	3.4	8.8	27.95
68	42	4.0	0.9	1.0	0.8	16	1.4	4.4	13.69
69	115	6.8	1.5	1.7	0.6	9.0	1.8	8.8	23.15
70	108	11	3.1	4.2	1.2	11	4.5	22	40.21
71	145	18	2.6	1.4	1.0	53	2.8	3.7	55.66
72	251	25	7.7	8.2	2.1	61	10	30	94.07
73	99	6.4	2.4	3.9	2.2	34	3.0	5.0	25.84
74	46	2.3	1.0	2.2	1.1	11	2.0	1.0	9.85
75	57	2.2	1.0	3.8	0.5	1.0	6.1	1.7	9.6
76									

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
77	44	4.7	1.5	1.1	0.2	34 (3)	0.1	0.9	17.9
78	92	7.9	3.2	4.0	0.6	36	2.1	2.4	32.87
79	371	33	7.1	25	2.0	89	6.5	18	111.61
80	204	15	5.7	4.1	0.8	54	2.0	9.3	60.87
81	146	30	5.7	11	2.0	54	4.5	25	98.37
82	29	3.5	0.7	2.1	0.4	9.2	2.3	5.4	11.62
83	97	11	1.9	3.2	0.7	54	22	25	35.29
84	41	0.8	0.5	2.4	0.7	1.0	4.4	4.5	4.05

Table 10. Recommended 10th percentile conductivity, GIs, and hardness estimates for SO group 7 through 9 (number of stations shown in parentheses if n<10)

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
1	111	12	3.4	4.3	0.8	56	2.3	6.3	43.94
2									
3	58	5.0	1.6	3.4	0.6	20	2.7	2.3	19.06
4	118	13	3.6	3.7	0.9	52	1.7	6.9	47.26
5									
6	101	9.0	4.5	4.9	0.9		1.7	3.1	40.95
7	122	10	5.0	6.2	1.0	56	3.5	4.8	45.5
8									
9									
10	71	5.7	1.5	2.0	0.7	16	0.8	4.2	20.4
11	72	8.5	1.5	3.3	0.7	32	0.8	5.0	27.4
12	310	37	10	13	2.5	122	11	30	133.5
13	430	38	10	32	5.6	175	15	27	136
14	810	64	23	69	3.2	121	55	181	254.3
15	51	5.2	1.5	1.4	0.5	20	0.4	3.1	19.15
16									
17	189	20	5.6	3.7	1.1	69	1.3	13	72.96
18	342	35	11	14	1.3	119	2.5	45	132.6
19	608	55	20	44	2.2	145	13.1	149	219.5
20	373	39	12	25	1.7	102	9.7	85	146.7
21									
22	279	28	4.9	15	1.9	80	4.5	37	90.09
23									
24	554	60	11	76	4.3	107	49	145	195.1
25	830	64	20	60	4.6	127	16	184	242
26	876	56	20	61	3.5	128	24	187	222
27	648	61	16	43	6.2	96	25	112	218.1
28	395	41	8.9	15	6.4	119	10	36	138.99
29	1194	71	19	132	4.7	89	210	130	255.4

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
30									
31	817	59	16	76	4.1	102	88	141	213.1
32	438	45	6.2	28	3.9	107	37	38	137.92
33	428	46	5.3	31	4.7	128	42	28	136.73
34	477	47	6.9	36	4.3	103	49	35	145.79
35	85	6.9	1.6	5.8	1.4	46	5.0	8.0	23.81
36	179	2.4	0.7 (3)	1.4 (3)	1.7 (3)		20	21	8.87
37	355	28	7.0	29	2.9	73	30	28	98.7
38									
39	215	28	7.7	2.8	1.1	96	3.4	6.0	101.57
40	310	34	5.2	6.4	3.0	96	7.3	24	106.32
41									
42	338	49	18	26	3.4	144	8.7	87	196.3
43	394	36	12	24	2.4	122	6.2	74	139.2
44									
45	53	4.1	1.7	6.0	1.5	13	5.0	7.3	17.22
46	642	52	25	49	12	176	22	149	232.5
47	570	48	12	15	3.7	159	11	44	169.2
48	425	44	19	14	5.3	188	9.9	61	187.9
49									
50									
51	353	44	16	7.2	2.1	217	10	13	175.6
52	115	12	4.4	2.9	1.1	40	4.4	5.0	48.04
53	544	53	33	7.9	1.8		19	22	267.8
54	388	41	18	9.7	2.1	131	16	25	176.3
55	502	48	18	20	3.0	182 (4)	32	33	193.8
56									
57	405	43	12	9.5	2.8	104	20	30	156.7
58									
59	65	3.9	0.7	8.5	0.8	6.0	13	6.0	12.62
60									
61									
62									
63	80	3.6	1.4	5.1	2.0	8.5	6.5	7.4	14.74
64	148	14	3.5	4.6	1.3	28	8.0	20	49.35
65	69	4.7	1.2	3.7	1.2	15	4.1	5.7	16.67
66									
67	96	15	3.4	4.7	1.2	28	5.7	12	51.44
68	138	17	3.5	4.1	1.1	57 (8)	5.7	12	56.85
69									
70	225	21	5.4	9.8	1.4	29	10	44	74.64
71	183	23	4.3	3.2	1.4	56	3.8	13	75.13

Ecoregion	Conductivity	Calcium	Magnesium	Sodium	Potassium	Alkalinity	Chloride	Sulfate	Hardness
72	310	34	10	8.3	2.3	88	12	23	126
73	146	14	3.7	3.4	1.7	54	3.5	6.0	50.17
74									
75	70	4.8	1.2	3.2	1.0	15	4.6	3.9	16.92
76									
77									
78									
79									
80	71	8.9 (3)	2.4 (3)	7.7 (3)	2.1 (3)		2.1 (3)	5.1	32.09
81	898	64	23	80	3.8	123	69	160	254.3
82	38	4.0	0.8	1.9	0.4	8.1	2.4	4.5	13.28
83	174	18	3.2	6.1	0.8	41	10	12	58.12
84									

At the level of individual ecoregions, the trends in parameter estimates as a function of SO group often reflect the assessment presented in the previous section. In the majority of ecoregions, most of the parameter estimates increase with SO group, as illustrated in Figure 25 for Ecoregion 46, the Northern Glaciated Plains. However, other trends were observed as well. In Ecoregion 83 (Eastern Great Lakes Lowland), conductivity and cation concentrations were approximately equal in the low and high SO groups and lower in the medium SO group, as shown in Figure 26. Figure 27 illustrates the trends in Ecoregion 54, the Central Corn Belt Plains. In this ecoregion (and several others), most of the parameter estimates decreased with SO group. The explanation for different trends within ecoregions, which may reflect different causes, is beyond the scope of this effort.

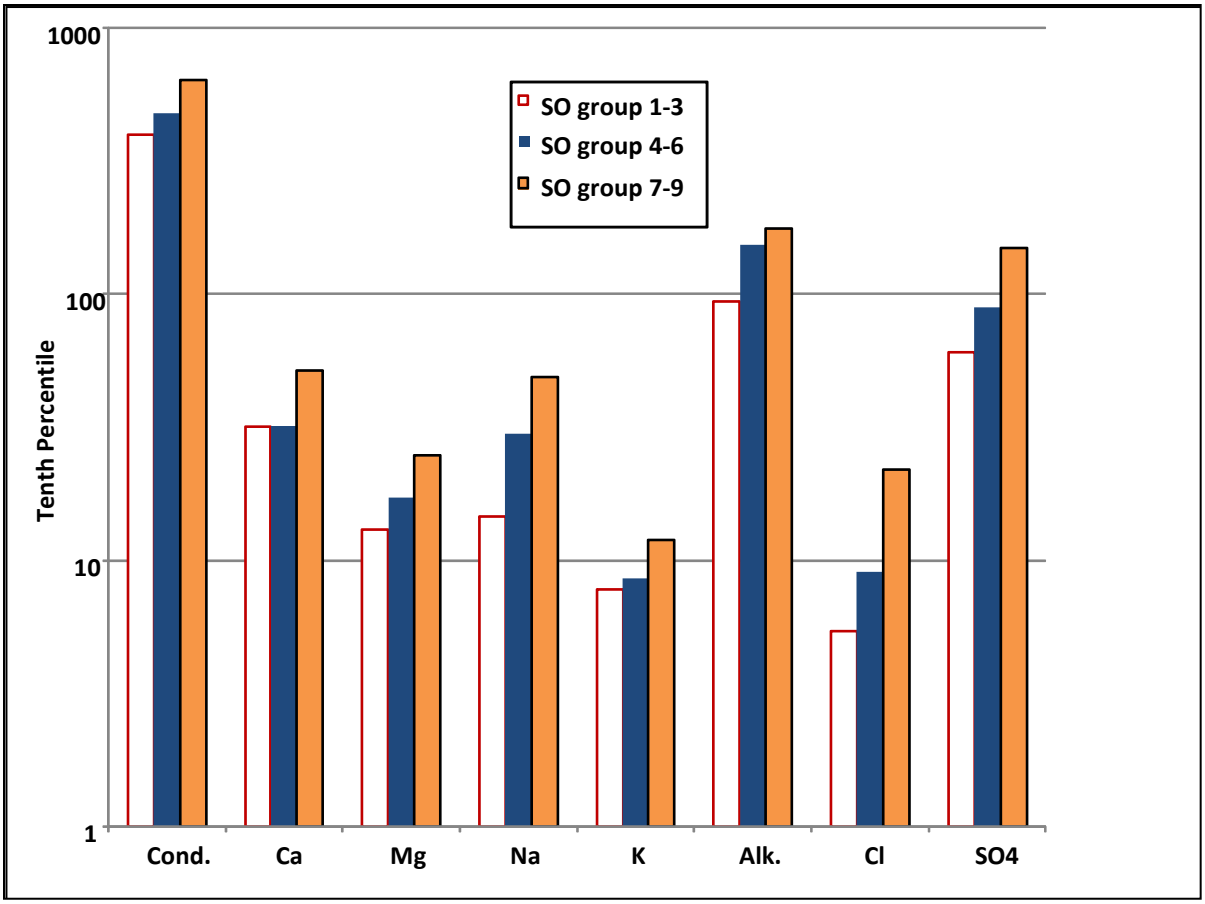

Figure 25. BLM parameter estimates (10th percentile values) for each SO group in Ecoregion 46 (Northern Glaciated Plains)

Key: Stream order: 1-3 are headwater streams, 4-6 are mid-reaches, and 7-9 are rivers.

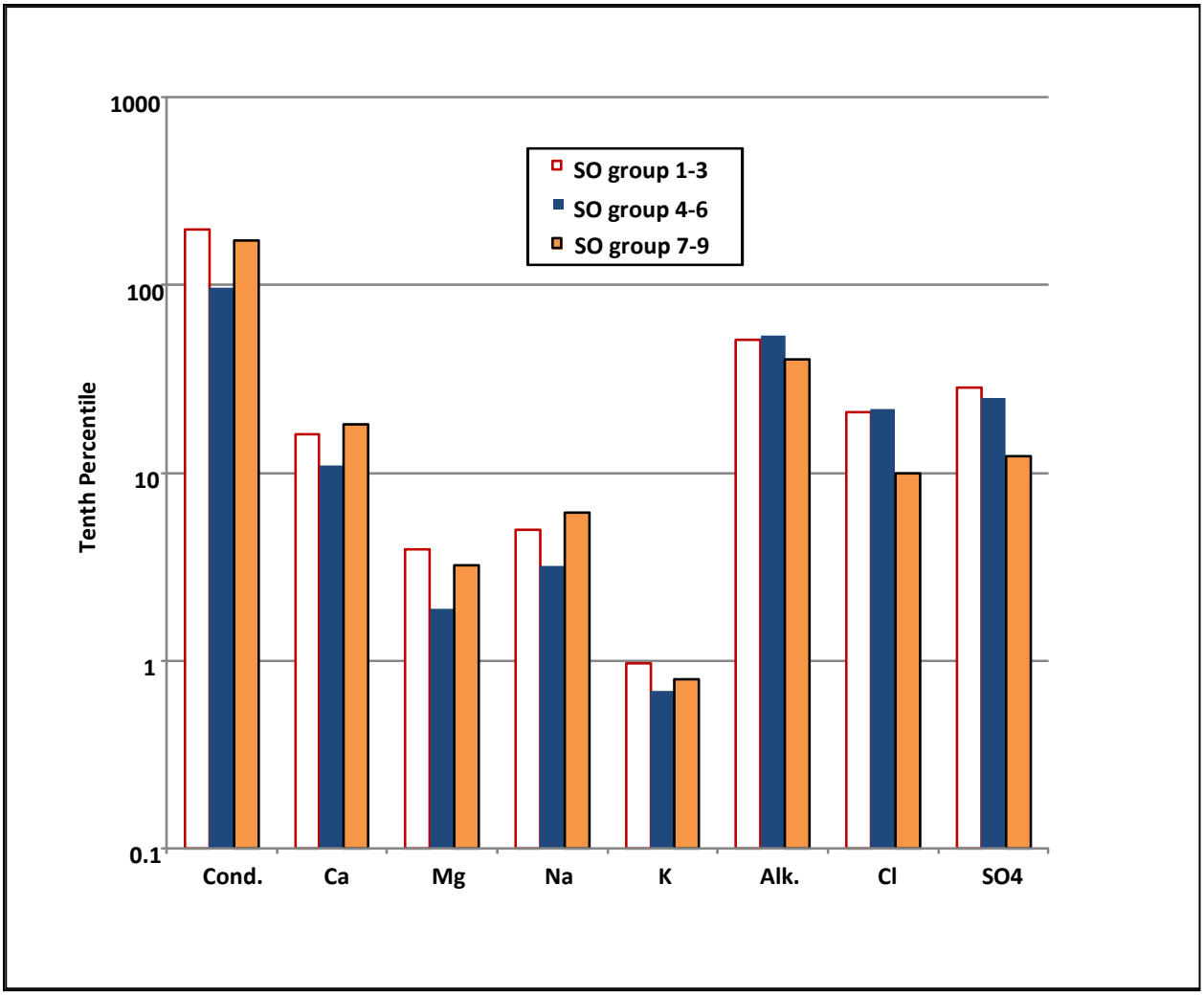

Figure 26. BLM parameter estimates for each SO group in Ecoregion 83 (Eastern Great Lakes Lowland)

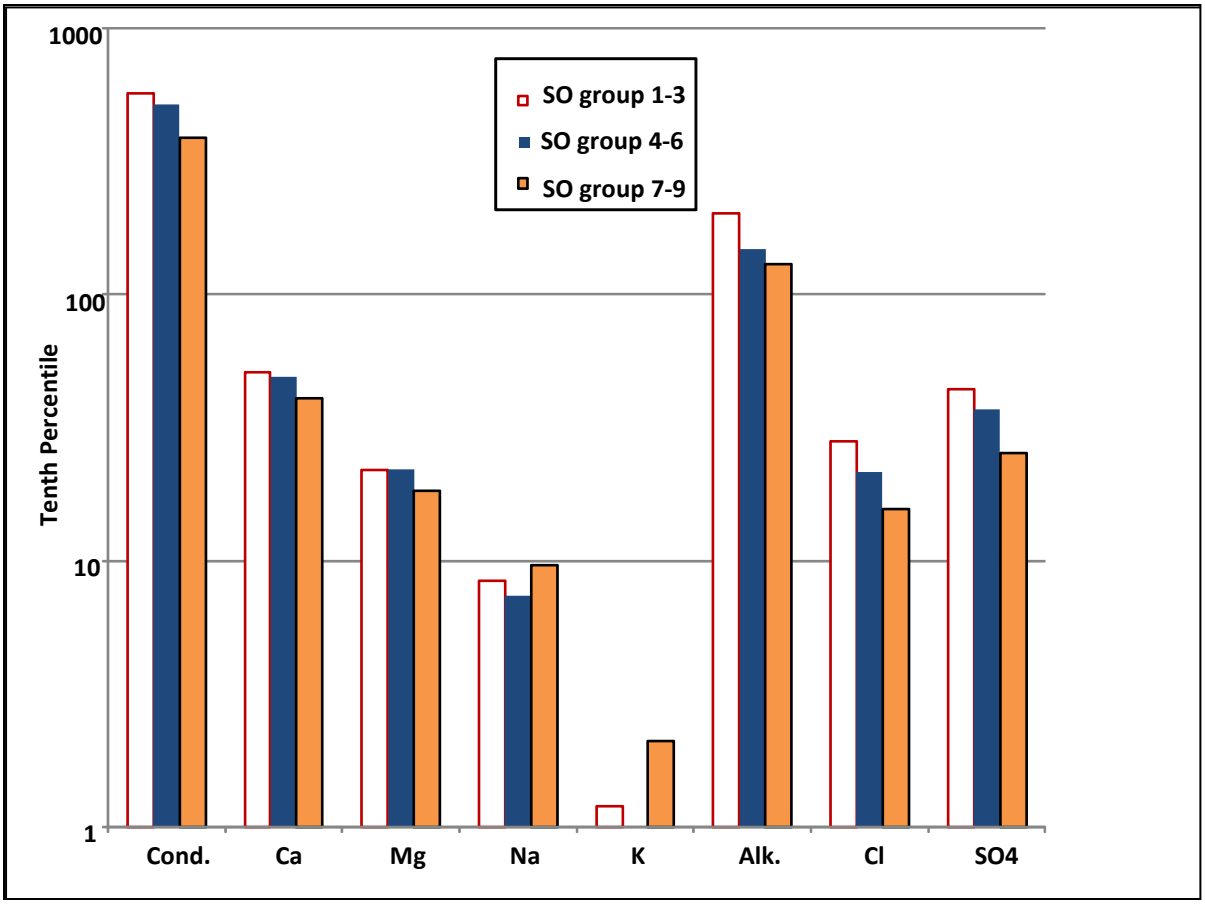

Figure 27. BLM parameter estimates for each SO group in Ecoregion 54 (Central Corn Belt Plains)

3.3.3 Comparison of Parameter Estimates to Results of Probability-Based Surface Water Sampling

There have been relatively few efforts to estimate GI concentrations in surface water at the national scale. Carleton (2006) developed a prototype geostatistical approach to estimate BLM parameters averaged over 8-digit HUC polygons. Carleton examined data from the NWIS and noted several limitations of this dataset in terms of uneven spatial sampling intensity. Carleton's prototype did not incorporate SO in the analysis, nor did it generate BLM parameter estimates.

Griffith (2014) compiled data from probability-based surface water quality sampling surveys conducted by EPA between 1985 and 2009, mostly from wadeable streams (SO group 1 through 4). These surveys included the National Acid Precipitation Assessment Program surveys, EMAP and regional EMAP surveys, WSA, and NRSA. The probability-based sample designs ensured that the results of these surveys represented the character of streams across the continental U.S. The water quality parameters included the same GIs as discussed above, and the results were presented on the same Level III ecoregion-specific basis.

We compared current results to those of Griffith (2014) because the lack of a probability-based sample design is a potential source of bias in the NWIS dataset. Parameter estimates based on the NWIS data for SO group 1 through 3 were compared to the corresponding estimates calculated by Griffith (2014). While Griffith did not tabulate 10th percentiles, he did tabulate first quartile (i.e., 25th percentile)

statistics for each ecoregion in supplemental material published with his article. Accordingly, we calculated 25th percentiles of the ecoregional NWIS data in SO class 1 through 3 (in addition to the 10th percentiles) to facilitate this comparison. The 25th percentiles from the two datasets are compared for conductivity in Figure 28 and calcium in Figure 29. The scatter plots reveal significant log-linear relationships between 25th percentiles for the two datasets; the coefficient of determination (R^2) was 0.668 for conductivity and 0.551 for calcium. For conductivity, the 25th percentiles differed by more than a factor of 2 in 17% of the Level III ecoregions; for calcium, 26% of the ecoregional results differed by more than a factor of 2.

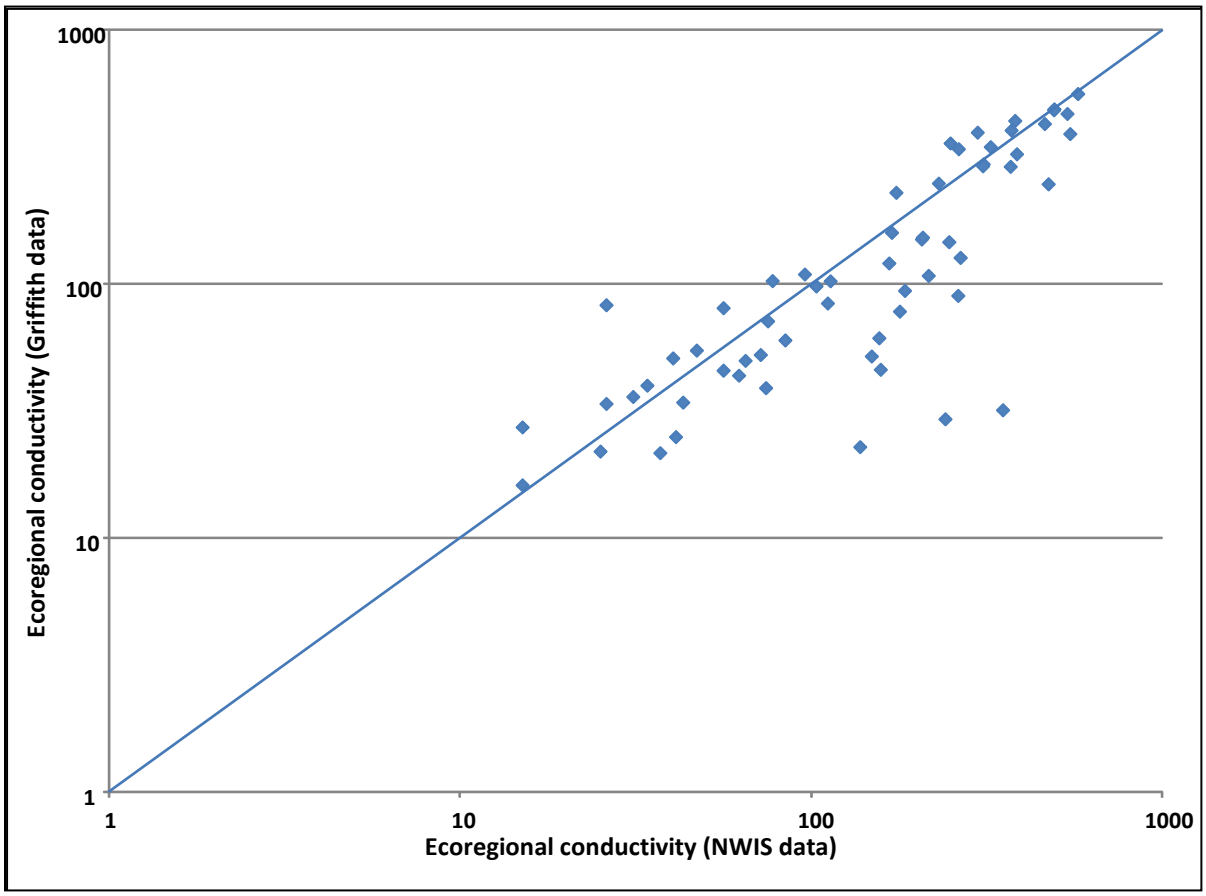

Figure 28. Scatter plot of ecoregional 25th percentile conductivity for NWIS Data (SO Class 1-3) versus ecoregional 25th percentile conductivity for Griffith data (mostly SO 1-4)

Solid diagonal line represents 1:1 agreement.

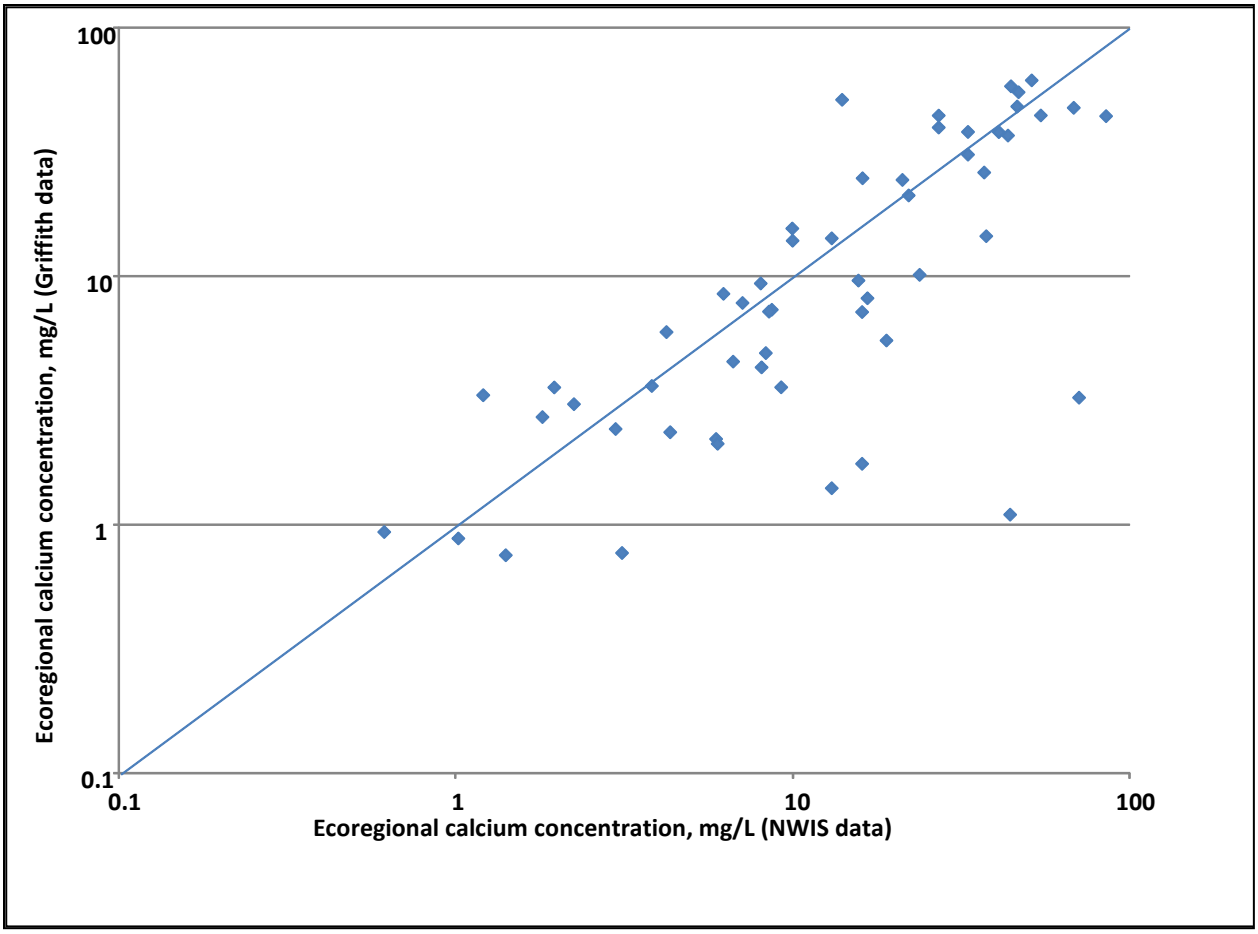

Figure 29. Scatter plot of ecoregional 25th percentile calcium concentration for NWIS data (SO class 1-3) versus ecoregional 25th percentile calcium concentration for Griffith data (mostly SO 1-4)

Solid diagonal line represents 1:1 agreement.

These results suggest reasonable overall consistency between datasets, as well as significant disparity in specific ecoregions. For example, agreement was especially poor in Ecoregions 19 (Wasatch and Uinta Mountains), 37 (Arkansas Valley), 38 (Boston Mountains), 39 (Ozark Highlands), 75 (Southern Coastal Plain), and 78 (Klamath Mountains). NWIS data were examined at the station-specific level to understand why these ecoregional 25th percentiles of conductivity and calcium in the low SO group were so inconsistent with corresponding percentiles presented by Griffith. Table 11 presents salient characteristics of the conductivity data for Ecoregion 19, including the number of stations, samples per station, 25th percentile conductivity, the lowest station-specific 25th percentile conductivity in the NWIS data, and other remarks. In that ecoregion, conductivity data were reported for 62 stations in the NWIS database; the number of observations per station ranged from 1 to 189, with a median of 22 observations per station. In comparison, the EPA data analyzed by Griffith reported conductivity data for 32 stations, with a single observation per station. The 25th percentile conductivity based on NWIS data was 240 versus 22.9 microsiemens per centimeter (µS/cm) based on Griffith's analysis of EPA data. When recalculated for individual stations in Ecoregion 19 low SO group, the 25th percentile conductivities varied widely, from 18.25 to 1590 µS/cm. Griffith reported a median conductivity of 213

µS/cm (nine times larger than the 25th percentile), indicating considerable variability in that data as well.

Table 11. Characteristics of the conductivity data for Ecoregion 19 in the low SO group.		
	NWIS data	Griffith data
Number of stations	62	32
Samples per station	1 – 189	1
Median samples per station	22	1
25th percentile conductivity (µS/cm)	240	22.9
Range of station-specific 25th percentile conductivity (µS/cm)	18.25-1590	
Other remarks	25th percentiles of conductivity reported by Griffith were marginally higher than the *minimum* station-specific 25th percentiles calculated from the NWIS data	Median conductivity = 213 µS/cm (19x the 25th percentile value) indicates high variability

The same information is tabulated for Ecoregions 37, 38, 75, and 78 in Tables 12 through 15. Although the details regarding the data vary in each of these ecoregions, a number of commonalities emerge from these examinations:

- In four ecoregions, the number of NWIS stations in SO group 1 through 3 was small relative to other ecoregions (the median number of ecoregional stations in SO group 1 through 3 was 68).
- In three ecoregions, 10th and 25th percentiles of conductivity and/or calcium decreased with SO group, contrary to the general trend.
- In three ecoregions, 25th percentiles of conductivity reported by Griffith were marginally higher than the minimum station-specific 25th percentiles calculated from the NWIS data for the corresponding ecoregion and low SO group.

Table 12. Characteristics of the conductivity data for Ecoregion 37 in the low SO group.		
	NWIS data	Griffith data
Number of stations	34	45
Samples per station	1 – 129	1
Median samples per station	2	1
25th percentile conductivity (µS/cm)	350	32
Range of station-specific 25th percentile conductivity (µS/cm)	29-846.5	
Other remarks	10th and 25th conductivity percentiles *decrease* with SO group; Small number of NWIS stations; 25th percentiles of conductivity reported by Griffith were marginally higher than the *minimum* station-specific 25th percentiles calculated from the NWIS data	

Table 13. Characteristics of the conductivity data for Ecoregion 38 in the low SO group		
	NWIS data	Griffith data
Number of stations	31	38
Samples per station	1 – 8	1
Median samples per station	3	1
25th percentile conductivity (µS/cm)	137	22.9
Range of station-specific 25th percentile conductivity (µS/cm)	22-384	
Other remarks	10th and 25th conductivity percentiles *decrease* with SO group; Small number of NWIS stations; 25th percentiles of conductivity reported by Griffith were marginally higher than the *minimum* station-specific 25th percentiles calculated from the NWIS data	

Table 14. Characteristics of the calcium data for Ecoregion 75 in the low SO group.		
	NWIS data	Griffith data
Number of stations	360	42
Samples per station	1 – 177	1
Median samples per station	17	1
25th percentile calcium (mg/L)	13	1.41
Range of station-specific 25th percentile calcium (mg/L)	0.02-91	
Other remarks	10th and 25th calcium percentiles *decrease* with SO group	

Table 15. Characteristics of the conductivity data for Ecoregion 78 in the low SO group.		
	NWIS data	Griffith data
Number of stations	15	45
Samples per station	1 – 18	1
Median samples per station	8	1
25th percentile conductivity (µS/cm)	26	98.4
Range of station-specific 25th percentile conductivity (µS/cm)	19-326.5	
Other remarks	Small number of NWIS stations; 6 of 15 stations were Ashland Creek (OR) or tributaries	

It is possible that the disparities noted above arise in part from non-representative sampling in the NWIS data. For example, representativeness of NWIS data is questionable in Ecoregion 78 because 40% of the stations were sampled in a single water body. There was also a difference in the way data for repeated sampling at individual stations were processed, due to differences between the NWIS data and data compiled by Griffith. In the NWIS data, water quality was sampled repeatedly at a significant

number of stations, and we included daily averages of all of these measurements in the calculation of 10th percentile estimates. In the probabilistic EPA surveys analyzed by Griffith, individual stations were usually sampled once. In the case of repeated sampling at a station, Griffith used data from only the first sample reported for the station in the statistics he calculated. This difference implies that our estimates of BLM parameters incorporate temporal as well as spatial variability in water quality, while Griffith's results do not. Thus, it would be unrealistic to expect complete agreement between these results. It should also be reiterated that sampling bias in SO is probably not a factor in these disparities because the estimates from the NWIS data were based on measurements from stream orders 1 through 3, which is generally consistent with the data compiled by Griffith (2014).

It is particularly of concern when percentiles based on NWIS data are higher than those calculated based on Griffith data because this suggests the parameter estimates may result in non-conservative BLM predictions of copper, or others metals, criteria. To evaluate this concern, we ran the BLM (version 2.1.2) using the 25th percentile GI estimates of Griffith and those from the current analysis for NWIS SO group 1 through 3 for each ecoregion in which parameters were available. If the 25th percentile of a GI was not available, the value was projected from the 25th percentile of conductivity using regressions based on NWIS data. If the 25th percentile of conductivity was not available (this occurred in 24 ecoregions), no BLM prediction was made. The inputs for pH and DOC were ecoregional values. There were 60 ecoregions where BLM predictions of copper criteria using the 25th percentile GI estimates from NWIS and Griffith could be compared. The criteria estimated in these ecoregions using GI input parameters from the two sources agree very well, as shown in Figure 30. The R^2 was 0.9897, and relative percent differences (RPDs) ranged from -21 to 39% with an average RPD of 3% and a median of 0.1%.

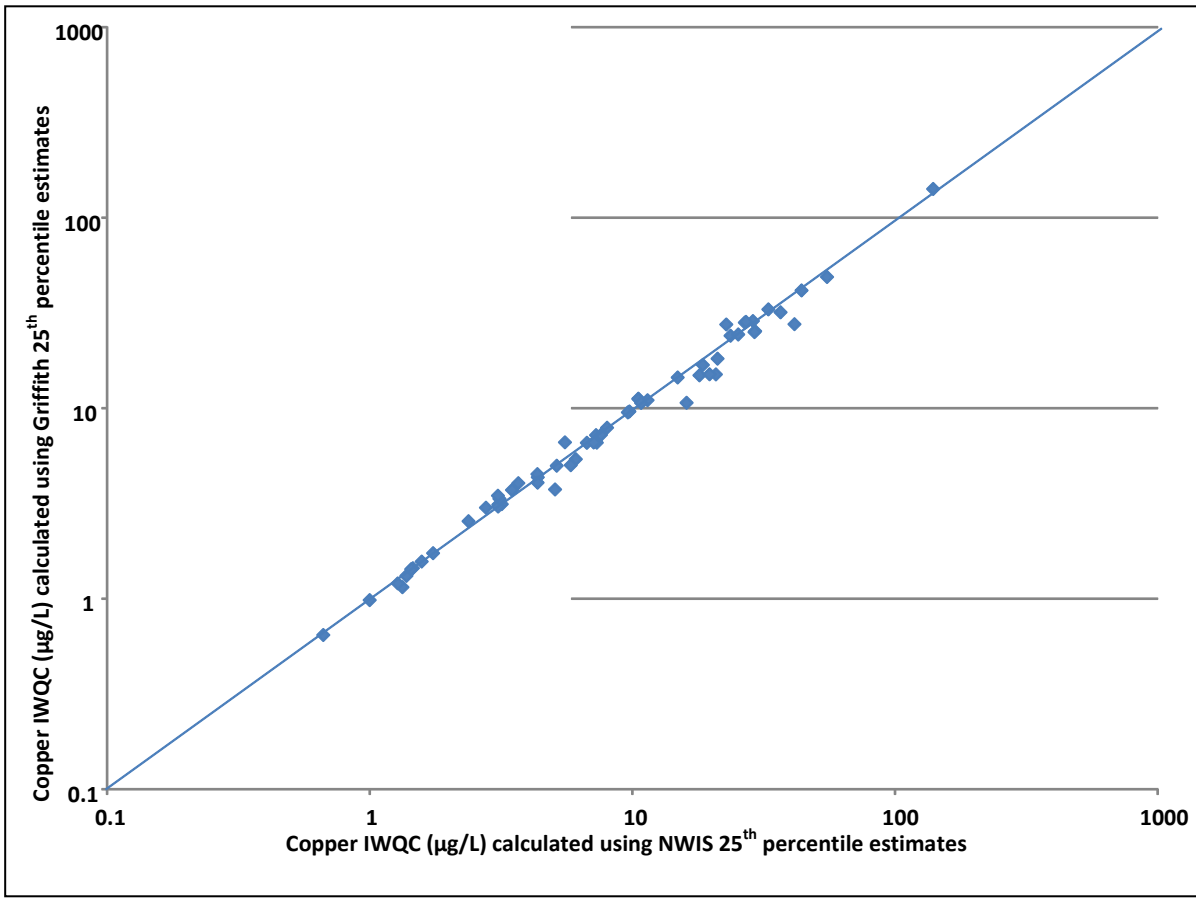

Figure 30. BLM predictions of copper criteria made with GI water quality parameters based on ecoregional 25th percentile from NWIS data (SO class 1-3) versus ecoregional 25th percentile calcium Concentration for Griffith data (mostly SO 1-4).

R^2 = 0.99. Solid diagonal line represents 1:1 agreement

While it is possible that the GI estimates presented here as recommended input values for use in the BLM for copper could be improved, it is not clear that such improvement would result in substantially better predictions of protective copper criteria based on the BLM. This is because BLM predictions for copper are much more sensitive to pH and DOC.

3.4 Summary

In this section we have incorporated SO variation in the GI water quality parameter estimates to refine the default parameters estimates for use in the BLM when data are not available.

EPA found that values of the GI parameter estimates generally increased with SO. This trend was most apparent and consistent for higher order streams (SO≥7). Tenth percentile estimates of conductivity increase with stream order group in 58% of ecoregions (comparing low to medium SO groups) and 84% of ecoregions (comparing medium to high SO groups). The same trend was evident for the GIs that are input parameters to the BLM.

We compared the parameter estimates for SO group 1 through 3 to those calculated by Griffith (2014). This comparison revealed significant log-linear relationships between 25th percentiles for the two datasets; the coefficient of determination (R^2) was 0.668 for conductivity and 0.551 for calcium. For conductivity, the 25th percentiles differed by more than a factor of 2 in 17% of the Level III ecoregions; for calcium, 26% of ecoregional results differed by more than a factor of 2. There is also considerable variability in the relationship between ecoregional statistics based on NWIS versus Griffith's data. Possible causes of these disparities may be due to sampling bias in the NWIS database, limited numbers of samples in some ecoregions, and differences in the degree and treatment of repeated sampling at individual locations.

NWIS percentiles higher than Griffith (2014) suggest that the recommended parameter estimates may result in non-conservative BLM predictions of copper criteria. The BLM was run to predict copper criteria for 60 ecoregions using 25th percentile GI estimates as parameter inputs from NWIS and Griffith's data. The criteria predicted using the two sets of GI parameter inputs agreed favorably. The R^2 was 0.990, and RPDs ranged from -21 to 39% with an average RPD of 3% and a median of 0.1%. These results demonstrated that the recommended default GI parameter estimates are reasonably conservative.

EPA incorporated SO variation in the parameter estimates to refine recommended input values for use in the BLM. EPA found the 10th percentile ecoregion, SO group specific estimates to be reasonably protective inputs and recommends their use where site-specific parameters are not readily available.

4 DOC ESTIMATION USING THE NATIONAL ORGANIC CARBON DATABASE

The following section summarizes our investigation into whether ecoregion and water body-type-specific DOC concentration percentiles tabulated by EPA for the *Development of National Bioaccumulation Factors Technical Support Document* (hereafter referred to as the National Organic Carbon Database (NOCD)) (USEPA, 2003)) offer reasonable estimates of lower-percentile DOC concentrations. A summary of the NOCD's data sources, analysis, and uncertainty associated with ecoregional statistics derived from the NOCD is provided below. This section also discusses how we recalculated ecoregional DOC percentiles for rivers and streams, and then tested for bias in the NOCD. Finally, we compared results based on the NOCD and data from the Wadeable Stream Assessment (WSA) (USEPA, 2006b)) and the National River and Stream Assessment (NRSA) (USEPA, 2013)).

4.1 Description of the NOCD

The NOCD is a compilation of pre-2003 organic carbon data derived from two sources: EPA's Storage and Retrieval Data Warehouse (STORET), recently renamed the STORET Legacy Data Center (LDC), and USGS's National Water Data Storage and Retrieval System (WATSTORE), the predecessor of NWIS. A complete background on the NOCD is available in USEPA 2003.

EPA's LDC database contains water quality monitoring data collected by academia, volunteer groups, and tribes, as well as federal, state, and local agencies. Geographically, the LDC data represent all 50 states and all U.S. territories and jurisdictions, along with portions of Canada and Mexico. The database queried for this investigation is often referred to as the "historical" or "old" STORET database because it contains water quality data dating back to the early part of the 20th century through the end of 1998. Data from 1999 to the present are stored in the "modernized" STORET Data Warehouse.[4] The LDC contains raw biological, chemical, and physical data for both surface water and groundwater. Each sampling result is accompanied by: information on sample collection location (latitude, longitude, state, county, HUC, and a brief site description), date the sample was gathered, the medium sampled, and the name of the organization that sponsored the monitoring.

We retrieved data from LDC and WATSTORE in January 2000. Approximately 800,000 records containing data on particulate organic carbon (POC), dissolved organic carbon (DOC), or total organic carbon (TOC) were obtained for the period beginning in 1970 through the latest year that data were available (1999 for WATSTORE and 1998 for LDC). This initial retrieval was limited to samples taken from ambient surface waters (i.e., samples from wells, springs, effluents, and other non-ambient sources were excluded). Additionally, this retrieval included multiple types of organic carbon measurements to ensure that the data would be sufficiently comprehensive.

WATSTORE was established in 1972 to provide an effective and efficient means for processing and maintaining water data collected through USGS activities and to facilitate release of that data to the public. The WATSTORE database resides on the central computer facilities of the USGS and contains results of approximately two million analyses of both surface and groundwater that provide data on

[4] Refer to http://www.epa.gov/storet/dbtop.html for more information on the STORET Data Warehouse and the LDC.

chemical, physical, biological, and radiological characteristics. EPA queried WATSTORE, the Water Quality File, to retrieve DOC data.

After retrieval, the data from LDC and WATSTORE were combined into a single database. The data were then processed and screened to ensure that only the most appropriate data would be retained. This screening process is outlined below:

- Values that were coded in such a way as to suggest uncertainty in the measurement were deleted from the database.
- The database was restricted to the following water body types: estuaries, lakes, reservoirs, and streams (including rivers).
- "Pseudo-ecoregions" were added for the five Great Lakes.
- The time period for the data was restricted to 1980 through 1999.
- Some values for DOC were reported to be below analytical detection levels. In this situation, the value was assumed to be half of the reported detection level. Values with "high" detection levels (i.e., >1.0 mg/L for DOC) were deleted from the database because of the greater uncertainty involved in estimating definitive values of DOC in these situations.
- A small fraction of the DOC and POC concentrations obtained from the LDC database exceeded concentrations considered to represent upper limits of DOC concentrations reported in U.S. water bodies (i.e., 0.2% exceeded 60 mg/L for DOC). These extreme values were based on a review of organic carbon data by Thurman (1985), who reported extreme values of DOC concentrations as high as 50 mg/L in dystrophic lakes and 60 mg/L in tributaries draining wetland systems. Therefore, values for DOC above 60 mg/L were removed from the database.

The NOCD that resulted from processing and screening data retrieved from the LDC and WATSTORE databases has some limitations, which are described below:

- The WATSTORE and LDC databases do not reflect a random sampling of U.S. surface waters. They contain datasets with a diversity of sampling design and thus data may be biased towards locations and water bodies with known water quality impairments.
- These data also reflect spatial bias due to unequal sampling efforts in different areas. For example, about half of the DOC and POC values in the databases were from samples collected in Maryland, New York, Ohio, Florida, and Delaware. Therefore some states are disproportionally represented, even when one considers the relative surface water area likely to be contained within each state.
- WATSTORE and LDC generally contain more data from sampling sites in larger river and stream systems, and areas subjected to proportionately greater human influence compared with random statistical sampling.

4.2 Recalculation of Ecoregional DOC Percentiles for Rivers and Streams

Lower percentile (1^{st}, 5^{th}, 10^{th}, and 25^{th} percentiles) DOC concentrations were calculated from all data for rivers and streams in each Level III ecoregion (Table 16). Nonparametric (i.e., rank) percentiles were calculated following the recommendations of Dierickx (2008) and Hyndman and Fan (1996). We also calculated confidence limits for the percentiles using the method presented in Berthouex and Brown

(1994). Upper and lower 95% confidence limits (UCLs and LCLs) were calculated if 20 or more DOC concentrations were available for an ecoregion (Berthouex and Brown, 1994).

Table 16. Lower percentile values of DOC in U.S. streams and rivers by ecoregion, including 95% confidence limits for percentile concentrations if n>20.

Level III Ecoregion	Ecoregion Name	n (count)	1%	1% (LCL)	1% (UCL)	5%	5% (LCL)	5% (UCL)	10%	10% (LCL)	10% (UCL)	25%	25% (LCL)	25% (UCL)
1	Coast Range	91	≤0.2	≤0.2	0.6	0.9	≤0.2	1.1	1.12	0.78	1.4	1.8	1.4	2
2	Puget Lowland	835	0.84	0.47	1	1.9	1.7	2	2.5	2.3	2.7	3.9	3.5	4.1
3	Willamette Valley	66	≤0.5	≤0.5	0.73	0.8	≤0.5	1.08	1.07	0.68	1.2	1.48	1.2	2
4	Cascades	101	0	≤0	0.2	0.3	≤0	0.44	0.5	0.3	0.5	0.7	0.5	0.9
5	Sierra Nevada	32	≤1.9	≤1.9	1.9	1.9	≤1.9	2.3	2.09	≤1.9	2.3	2.55	2.13	3.11
6	Southern and Central California Chaparral and Oak Woodlands	480	1.1	≤0	1.5	1.8	1.7	1.9	2.1	1.9	2.2	3	2.67	3.2
7	Central California Valley	180	1.21	≤0.8	1.8	2.11	1.66	2.48	2.71	2.3	3.5	5.3	4.4	6.26
8	Southern California Mountains	6	≤4.4			≤4.4			≤4.4			5.45		
9	Eastern Cascades Slopes and Foothills	13	≤1.3			≤1.3			1.42			1.75		
10	Columbia Plateau	73	≤0.7	≤0.7	1.3	1.67	≤0.7	2.03	2.04	1.29	2.34	2.9	2.3	3.2
11	Blue Mountains	26	≤1	≤1	1.05	1.07	≤1	1.45	1.34	≤1	1.81	1.9	1.28	2.61
12	Snake River Plain	50	≤2	≤2	2	2	≤2	2.2	2.2	≤2	2.43	3.08	2.27	3.68
13	Central Basin and Range	1553	0.8	0.69	0.9	1.2	1.2	1.3	1.5	1.4	1.6	2	2	2.1
14	Mojave Basin and Range	35	≤2.5	≤2.5	2.55	2.58	≤2.5	3	2.84	≤2.5	3.32	3.6	2.99	3.8
15	Northern Rockies	778	0.7	0.6	0.72	0.9	0.9	1	1	1	1.1	1.3	1.3	1.4
16	Idaho Batholith	29	≤1.2	≤1.2	1.2	1.2	≤1.2	1.4	1.4	≤1.2	1.8	1.9	1.39	2.31
17	Middle Rockies	87	≤0	≤0	0	0	≤0	0.14	0.18	0	0.46	1.1	0.42	1.4
18	Wyoming Basin	150	2.05	≤1.9	3.03	3.56	2.26	4.18	4.31	3.59	4.6	5.48	5.1	5.8
19	Wasatch and Uinta Mountains	46	≤1.5	≤1.5	1.5	1.57	≤1.5	1.82	1.8	≤1.5	2.51	2.88	1.9	3.6
20	Colorado Plateaus	798	1.5	0.36	1.6	2.4	2	2.6	3	2.7	3.25	4.3	4.1	4.6
21	Southern Rockies	1129	0.2	0.1	0.3	0.5	0.4	0.5	0.6	0.6	0.6	0.8	0.8	0.9
22	Arizona/New Mexico Plateau	281	1.65	≤1.2	2.01	2.2	2.09	2.4	2.62	2.3	2.91	3.7	3.3	3.9
23	Arizona/New Mexico Mountains	37	≤1.3	≤1.3	1.41	1.48	≤1.3	2.35	2.16	≤1.3	2.64	2.8	2.33	3.57
24	Chihuahuan Deserts	116	0.5	≤0.5	0.64	1	0.5	2.15	2.34	1	3	3.68	3	4.44
25	High Plains	439	0.3	≤0.1	0.4	0.8	0.5	3	4.4	3.2	4.9	7	6.42	7.78
26	Southwestern Tablelands	167	1.94	≤1.8	2.04	2.5	2	2.98	3.28	2.52	3.7	4.4	4	4.9
27	Central Great Plains	228	0.59	≤0.4	2.01	3	1.8	3.2	3.8	3	4	5	4.8	5.71
28	Flint Hills	10	≤4.9			≤4.9			4.9			5.15		
29	Central Oklahoma/Texas Plains	289	1.9	≤1	2.09	3	2.5	3.38	3.8	3	4	4.8	4.21	5
30	Edwards Plateau	200	0.5	≤0.5	0.5	0.5	0.5	0.82	1	0.5	1	1	1	1.6
31	Southern Texas Plains	58	≤0.5	≤0.5	0.5	0.5	≤0.5	1	1	0.5	1.34	2	1	2
32	Texas Blackland Prairies	829	0.5	0.5	0.5	1	1	1	2	2	2	3	3	3.99

Level III Ecoregion	Ecoregion Name	n (count)	1%	1% (LCL)	1% (UCL)	5%	5% (LCL)	5% (UCL)	10%	10% (LCL)	10% (UCL)	25%	25% (LCL)	25% (UCL)
33	East Central Texas Plains	268	1	≤1	1.95	2.6	2	3	3	3	3	3.83	3.1	4
34	Western Gulf Coastal Plain	399	2.8	≤1.7	3	3.7	3	4	4	4	4	5	5	5.9
35	South Central Plains	523	1.45	≤0.4	2.07	3.92	3.2	4	4.6	4.1	4.98	5.7	5.4	6
36	Ouachita Mountains	198	0	≤0	0.47	0.8	0.4	1.1	1.1	0.86	1.6	2.38	2.08	2.8
37	Arkansas Valley	184	0.49	≤0.4	0.65	0.8	0.55	1.4	1.9	0.85	2.39	3.73	3	4.4
38	Boston Mountains	21	≤0.4	≤0.4	0.4	0.4	≤0.4	0.5	0.42	≤0.4	0.5	0.5	0.4	0.74
39	Ozark Highlands	233	1.67	≤1.5	2	2	2	2.12	2.3	2	2.5	3	2.7	3.15
40	Central Irregular Plains	434	2.71	≤1.4	3	3.5	3.1	3.6	4	3.6	4	4.9	4.6	5
41	Canadian Rockies	36	≤0.3	≤0.3	0.35	0.39	≤0.3	0.6	0.57	≤0.3	0.9	0.93	0.6	1.03
42	Northwestern Glaciated Plains	36	≤3	≤3	3.05	3.09	≤3	6.63	5.05	≤3	11.23	13.25	6.12	16.34
43	Northwestern Great Plains	679	2.22	0.72	3.19	4.4	3.8	5.2	6.2	5.6	6.73	9.9	9.19	10
44	Nebraska Sand Hills	4	≤1.4			≤1.4			≤1.4			1.4		
45	Piedmont	309	0.4	≤0	0.5	0.7	0.6	0.9	1	0.8	1.1	1.7	1.4	2.1
46	Northern Glaciated Plains	142	4.03	≤3.3	8.01	9.13	5.11	9.82	9.9	9.16	11	12	11	12.86
47	Western Corn Belt Plains	193	0.44	≤0.2	2.43	2.77	2.28	2.96	3.14	2.8	3.5	3.9	3.6	4.3
48	Lake Agassiz Plain	261	3.1	≤0.8	4.52	6.41	4.72	7.1	7.6	6.87	7.97	9	8.7	9.22
49	Northern Minnesota Wetlands	44	≤7.8	≤7.8	8.25	9.05	≤7.8	11	11	≤7.8	12	13	11	14.88
50	Northern Lakes and Forests	403	2	≤0.1	2.2	2.72	2.36	3.08	3.7	3.06	4.5	5.8	5.5	6.2
51	North Central Hardwood Forests	153	0.54	≤0	2.19	2.67	1.42	2.9	3.1	2.7	3.77	4.95	3.9	5.6
52	Driftless Area	49	≤1.7	≤1.7	2.31	2.4	≤1.7	3.25	3.1	≤1.7	4.31	4.5	3.4	5.8
53	Southeastern Wisconsin Till Plains	439	2	≤0.25	2.1	3.4	2.7	4.3	5.3	4.3	5.6	6.8	6.5	7
54	Central Corn Belt Plains	202	1	≤0.7	1.68	2.12	1.61	2.5	2.73	2.2	3	3.9	3.57	4.8
55	Eastern Corn Belt Plains	1398	0	0	0	0	0	0.2	3.8	3	4	5.2	5	5.3
56	S. Michigan/N. Indiana Drift Plains	287	1.4	≤1.4	1.92	2.7	2.05	3.43	3.8	3.17	4.2	5.4	4.9	5.7
57	Huron/Erie Lake Plains	3825	0	0	0	3.9	3.7	4.07	4.7	4.52	4.8	5.8	5.7	5.9
58	Northeastern Highlands	14044	0.25	0.25	0.25	0.25	0.25	0.5	0.64	0.62	0.65	0.95	0.94	0.97
59	Northeastern Coastal Zone	102	0.05	≤0	1.6	2.02	≤0	2.6	2.63	1.74	2.92	3.28	3	3.63
60	Northern Appalachian Plateau and Uplands	354	0.6	≤0.5	0.8	1	0.87	1.1	1.3	1.1	1.4	2	1.9	2.3
61	Erie Drift Plains	919	0	0	0	4.8	4.6	4.9	5.1	5	5.1	5.5	5.4	5.6
62	North Central Appalachians	106	0.41	≤0.4	0.5	0.5	≤0.4	0.7	0.7	0.5	0.78	0.98	0.8	1.2
63	Middle Atlantic Coastal Plain	16730	1.1	1.01	1.2	1.8	1.8	1.8	2.2	2.2	2.2	2.7	2.7	2.7
64	Northern Piedmont	1525	0.43	0.3	0.5	1	0.99	1	1.3	1.2	1.4	2.2	2.1	2.4
65	Southeastern Plains	3813	1	0.51	1.1	2	1.9	2.03	2.4	2.38	2.5	3.3	3.2	3.3
66	Blue Ridge	699	0	0	0	0.4	0.3	0.4	0.5	0.4	0.5	0.6	0.6	0.7
67	Ridge and Valley	733	0.1	0.1	0.1	0.4	0.3	0.5	0.6	0.5	0.7	1.05	1	1.1
68	Southwestern Appalachians	47	≤0.5	≤0.5	0.5	0.58	≤0.5	0.9	0.9	≤0.5	1	1	0.9	1.38
69	Central Appalachians	864	0.3	0.15	0.4	0.6	0.5	0.6	0.7	0.7	0.8	1.1	1.1	1.2
70	Western Allegheny Plateau	1795	0	0	0	0.78	0.2	1	1.6	1.4	1.7	2.7	2.5	2.9
71	Interior Plateau	559	0.1	≤0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.2	0.3

Level III Ecoregion	Ecoregion Name	n (count)	1%	1% (LCL)	1% (UCL)	5%	5% (LCL)	5% (UCL)	10%	10% (LCL)	10% (UCL)	25%	25% (LCL)	25% (UCL)
72	Interior River Valleys and Hills	328	1.32	≤0.15	1.8	2.1	1.9	2.4	2.69	2.4	3.05	4.2	3.6	4.6
73	Mississippi Alluvial Plain	503	1.7	≤1	2.14	2.82	2.4	3.1	3.4	3.2	3.6	4.3	4.1	4.5
74	Mississippi Valley Loess Plains	21	≤1.4	≤1.4	1.41	1.41	≤1.4	2.61	1.72	≤1.4	2.7	2.7	1.46	4.72
75	Southern Coastal Plain	4223	0.9	0.9	1.09	5.3	4.7	6	8	7.6	8.3	12.1	11.9	12.3
76	Southern Florida Coastal Plain	-	-			-			-			-		
77	North Cascades	50	≤0.2	≤0.2	0.2	0.26	≤0.2	0.4	0.4	≤0.2	0.4	0.48	0.4	0.5
78	Klamath Mountains	8	≤1.7			≤1.7			≤1.7			2.3		
79	Madrean Archipelago	9	≤2.6			≤2.6			2.6			4.05		
80	Northern Basin and Range	16	≤1.6			≤1.6			1.81			2.5		
81	Sonoran Basin and Range	133	1.33	≤1.3	1.7	1.8	1.38	2.1	2.2	1.8	2.6	3.85	2.94	4.4
82	Laurentian Plains and Hills	21	≤4.6	≤4.6	4.7	4.69	≤4.6	5.68	5.52	≤4.6	7.98	8.45	5.15	9.3
83	Eastern Great Lakes and Hudson Lowlands	1430	0	0	0	0	0	0.2	1.9	1	2	5.1	5	5.5
84	Atlantic Coastal Pine Barrens	243	1	≤0.9	1.1	1.22	1.1	1.5	1.6	1.32	2	2.6	2.4	3
	Lake Erie	9	≤1.1			≤1.1			1.18			1.4		
	Lake Michigan	5	≤2.6			≤2.6			≤2.6			2.6		
	Lake Ontario	14	≤0.4			≤0.4			0.4			2.2		
	Lake Superior	7	≤1.2			≤1.2			≤1.2			1.4		

As was the case for the BLM GI input parameters, we consider low percentiles of the ecoregional DOC concentration distributions to be reasonably protective inputs to the BLM for sites where DOC measurements are not available.

4.3 Testing for Bias in the NOCD

EPA conducted an evaluation of bias in the NOCD (USEPA, 2003) using data from EPA's Environmental Monitoring and Assessment Program's (EMAP) 1997 to 1998, which sampled mid-Atlantic streams and rivers (USEPA, 2006b). This effort is described below in Section 4.3.1.

We also evaluated the bias in the NOCD using independent data from EPA's Wadeable Streams Assessment (WSA), which included DOC measurements from a statistically based random sample of approximately 2,000 wadeable, perennial 1st through 5th order streams (USEPA, 2006c). In Section 4.3.2, we compare the WSA data to the ecoregion-specific DOC concentration percentiles calculated from the NOCD.

4.3.1 Previous Efforts Using EMAP Data

Ideally, the data used to generate the distribution of national organic carbon concentration values should originate from a random sampling of U.S. surface waters, and should be appropriately stratified and weighted by spatial and temporal factors that would be expected to influence organic carbon concentrations in aquatic ecosystems (e.g., water body type, hydrologic and watershed characteristics, ecoregion, season). However, these data are not available on a national scale. The strength of this analysis is that the data from USGS's WATSTORE and EPA's LDC databases include a large number of

records (e.g., >110,000 DOC values), a representation of DOC values for all 50 states, and reasonably long period over which data were collected (1980 through 1999 for these analyses).

Data generated by EMAP are based on a stratified, random sampling strategy that was specifically designed to minimize the influence of sampling bias on the data and to enable statistically based extrapolations across geographic regions (Herlihy et al., 2000). At the time the NOCD was developed, the EMAP databases containing DOC measurements were limited to smaller geographic scales and specific water body types.

Previously, to address the question of sampling bias and its impact on the representativeness of the NOCD values, EPA made quantitative comparisons that involved contrasting geographically distinct subsets of the WATSTORE/LDC databases with geographically similar subsets of data produced by EMAP. DOC data from EMAP's 1997 to 1998 sampling of mid-Atlantic streams and rivers were compared with similar geographic subsets from the WATSTORE/LDC databases. The mid-Atlantic EMAP database was chosen because sufficient DOC data were available for rivers and streams to make meaningful comparisons at the state and ecoregion levels. Similar comparisons are made for four mid-Atlantic ecoregions (Piedmont, Ridge and Valley, Central Appalachians, Western Allegheny Plateau) which is well represented in the WATSTORE/LDC databases (USEPA, 2003).

Based on both sets of comparisons, it is apparent that the agreement between the WATSTORE/LDC and EMAP data was best at the middle to lower tails of the distributions, and poorest at the higher end of the distributions. At the lower tails of the distributions (e.g., 10^{th}, 25^{th} percentiles) the WATSTORE/LDC DOC data are generally within 30% of the EMAP data (Ecoregion 70 being the only exception). The median DOC values of the WATSTORE/LDC data show a slightly higher bias compared with median values from the EMAP data, but are usually within a factor of 1.5 (Ecoregions 47 and 70 are about a factor of 2 greater). This result is expected, given the greater focus of the WATSTORE/LDC sampling sites on larger river and stream systems, and on areas subjected to proportionately greater human influence compared with the EMAP sampling sites. Since EPA is interested in supporting the generation of BLM values that are protective of aquatic life, the lack of bias noted for the lower tails of the DOC concentration distributions is noteworthy.

4.3.2 Testing for Bias Using Data from the WSA

A more comprehensive evaluation of the effects of sampling bias on the NOCD can now be made using the results of national statistically-designed water quality sampling surveys. We assembled a database of organic carbon data from 1,313 randomly selected sites throughout the continental U.S. collected for the WSA. GIS procedures were used to associate each site with the Level III ecoregion corresponding to its location.

The 1,392 sites sampled for the WSA were identified using a probability-based sample design, a technique in which every element in the population has a known probability of being selected for sampling (USEPA, 2006). This ensured that the results of the WSA reflect the full range of variation among wadeable streams across the U.S. The target population for the WSA was wadeable, perennial streams in the conterminous U.S. (lower 48 states). The WSA used the National Hydrography Dataset (NHD), a comprehensive set of digital spatial data on surface waters (USGS, 2012), to identify the location of wadeable perennial streams. Rules for site selection included weighting to provide balance in the number of stream sites from each of the 1^{st} through 5^{th} SO size classes (Strahler, 1952, 1957),

and controlled spatial distribution to ensure that sample sites were distributed across the U.S.). The basic sampling design drew 50 sampling sites randomly distributed in each of the EPA Regions and WSA ecoregions. The unbiased site selection of the survey design ensures that assessment results represent the condition of the streams throughout the nation.

4.3.2.1 Selection of Statistical Test to Assess Potential Bias in DOC Data

The most appropriate statistical test for determining bias in the NOCD is the comparison of WSA and organic carbon database DOC data within each ecoregion as independent groups of data to determine if one group tends to contain larger values (Helsel and Hirsch, 2002; USGS, 2002). The WSA and organic carbon database DOC data are independent because there is no natural structure in the order of observations across groups. A nonparametric statistical test is most appropriate since no assumptions regarding normality of the data are required. As noted by Helsel and Hirsch (2002), nonparametric tests are, in general, never worse than their parametric counterparts in their ability to detect departures from the null hypothesis, and may be better. These considerations led us to select the rank-sum test, a nonparametric procedure for determining whether data are significantly different between two independent groups. This test is also known as the Wilcoxon Rank-Sum Test or, alternatively, the Mann-Whitney U-Test.

In its most general form, the rank-sum test is a test to determine whether one group tends to produce larger observations than another group. It has as its null hypothesis:

H_0: Prob [x > y] = 0.5

where x are data from one group and y are from another group (the probability of an x value being higher than any given y value is one-half). The test is typically used to determine whether two groups come from the same population (same median and other percentiles), or alternatively whether they differ only in location (central value or median). If both groups of data are from the same population, about half of the time an observation from either group could be expected to be higher than that from the other, so the above null hypothesis applies. If the groups belong to different populations the null hypothesis does not apply.

In practice, the rank-sum test takes several forms, depending upon the size of the smaller sample (*n* observations) and the larger sample (*m* observations). Walpole and Myers (1978; Section 13.2) present the details of four alternative forms of the rank-sum test, depending on the sizes of *n* and *m*. The exact form of the rank-sum test is the only form appropriate for comparing groups of sample sizes of 10 or smaller per group. When both groups have samples sizes greater than 10 (*n*, *m* > 10), the large-sample approximation may be used.

4.3.2.2 Rank-Sum Test Comparing WSA DOC Data to NOCD

Table 17 presents the results of the rank-sum test comparing Level III ecoregional DOC data from the WSA (USEPA, 2006b) and the NOCD. The left-hand columns present statistics (sample size, median, and Mann-Whitney U_x, and U_y) for the ecoregion-specific DOC data from the two datasets. The next six columns to the right present the test statistics for the appropriate form of the rank-sum test. The right-hand column provides a summary interpretation of the test for each ecoregion indicating whether the null hypothesis (H_0: DOC concentrations from both datasets are not different) should be rejected at the 5% level of significance, in favor of the alternative hypothesis (H_1: DOC concentrations are higher in the

national organic carbon database). In other words, rejection of the null hypothesis implies that DOC concentrations from the National Organic Carbon Database are biased high in that ecoregion relative to the WSA data.

Table 17. Results of rank-sum test comparing Level III ecoregional DOC data from WSA and NOCD

Level III Ecoregion	Ecoregion Name	WSA dataset			NOC database			M-W test (n1>10 and n2>10)		One large sample (n2>20)		Exact test (n2<20)	Exact test (n2<9)	Interpretation of test 1-sided @ 0.05
		n	median DOC	Ux	n	median DOC	Uy	Z	P	Z	P	critical U (0.05 level of signif.)	P	(H0: same mean of distributions)
1	Coast Range	38	1.41	2535.5	91	2.2	922.5	4.17	1.54E-05					reject H0
2	Puget Lowland	3	0.96	2289	835	6.6	216			-2.48	6.63E-03			reject H0
3	Willamette Valley	2	2.47	84	66	2.9	48			-0.65	2.57E-01			do not reject H0
4	Cascades	23	0.75	1673	100	1.4	627	3.39	3.46E-04					reject H0
5	Sierra Nevada	14	0.91	448	32	3.6	0	5.35	5.96E-08					reject H0
6	Southern and Central California Chaparral and Oak Woodlands	32	2.5	12372	479	4.6	2957	5.82	0					reject H0
7	Central California Valley	2	4.2	302	180	13	58			-1.65	4.98E-02			reject H0 (P~5%)
8	Southern California Mountains	41	1.66	245	6	8.9	1			-3.89	5.03E-05			reject H0
9	Eastern Cascades Slopes and Foothills	19	0.97	209	13	2.3	38	3.28	5.18E-04					reject H0
10	Columbia Plateau	8	2.59	477	73	3.6	107			-2.93	1.70E-03			reject H0
11	Blue Mountains	77	1.59	1625	26	3.1	376.5	4.74	1.07E-06					reject H0
12	Snake River Plain													no test

Level III Ecoregion	Ecoregion Name	WSA dataset			NOC database			M-W test (n1>10 and n2>10)		One large sample (n2>20)		Exact test (n2<20)	Exact test (n2<9)	Interpretation of test 1-sided @ 0.05
		n	median DOC	Ux	n	median DOC	Uy	Z	P	Z	P	critical U (0.05 level of signif.)	P	(H0: same mean of distributions)
13	Central Basin and Range	42	1.93	47300	1553	3	17926	4.986	2.98E-07					reject H0
14	Mojave Basin and Range	2	2.65	67	35	5.6	3			-2.15	1.58E-02			reject H0 (P~1.6%)
15	Northern Rockies	19	1.54	9149	778	1.8	5634	1.77	3.81E-02					reject H0 (P~3.8%)
16	Idaho Batholith	19	1.21	495	29	2.4	56	4.63	1.85E-06					reject H0
17	Middle Rockies	70	1.43	3142	81	1.9	2529	1.14	0.126					do not reject H0
18	Wyoming Basin	29	2.29	4006	150	7.3	344	7.17	0					reject H0
19	Wasatch and Uinta Mountains	25	2.11	978	46	4.8	172	4.85	5.96E-07					reject H0
20	Colorado Plateaus	24	2.22	15697	798	6.3	3455	5.34	5.96E-08					reject H0
21	Southern Rockies	43	2.05	18637	1129	1.3	29911	-2.59	4.83E-03					reject H0
22	Arizona/New Mexico Plateau	7	1.83	1723	281	5.6	244			-3.40	3.40E-04			reject H0
23	Arizona/New Mexico Mountains	31	1.94	984	37	5.3	163	5.02	2.38E-07					reject H0
24	Chihuahuan Deserts	1	1.48	110	116	6	6			-1.54	6.18E-02			do not reject H0
25	High Plains	6	3.5	2450	439	11	184.5			-3.62	1.48E-04			reject H0
26	Southwestern Tablelands	17	4.21	2246	167	6.3	593	3.95	3.90E-05					reject H0
27	Central Great Plains	12	4.71	2189	228	7	547	3.50	2.31E-04					reject H0
28	Flint Hills	2	4.91	19	10	8.7	1					1		reject H0

Level III Ecoregion	Ecoregion Name	WSA dataset			NOC database			M-W test (n1>10 and n2>10)		One large sample (n2>20)		Exact test (n2<20)	Exact test (n2<9)	Interpretation of test 1-sided @ 0.05
		n	median DOC	Ux	n	median DOC	Uy	Z	P	Z	P	critical U (0.05 level of signif.)	P	(H0: same mean of distributions)
29	Central Oklahoma/ Texas Plains	6	5.58	1098	289	6.7	636			-1.12	1.32E-01			do not reject H0
30	Edwards Plateau													no test
31	Southern Texas Plains													no test
32	Texas Blackland Prairies	2	7.37	880	829	6	778			-0.15	4.40E-01			do not reject H0
33	East Central Texas Plains	3	15.03	108.5	268	5	695.5			2.17	9.85E-01			do not reject H0
34	Western Gulf Coastal Plain	9	7.47	1860	399	7	1732			-0.183	0.427			do not reject H0
35	South Central Plains	24	7.7	6744	523	7.7	5808	0.62	2.68E-01					do not reject H0
36	Ouachita Mountains	6	2.2	963.5	196	3.7	212.5			-2.66	3.88E-03			reject H0
37	Arkansas Valley	3	4.61	327	184	7	225			-0.55	2.92E-01			do not reject H0
38	Boston Mountains	3	2.18	19	21	0.8	44			1.09	8.62E-01			do not reject H0
39	Ozark Highlands	10	1.8	2248	233	4.1	82			-4.98	3.26E-07			reject H0
40	Central Irregular Plains	8	6.7	1772	434	6.3	1700			-0.10	4.60E-01			do not reject H0
41	Canadian Rockies	5	0.8	127.5	36	1.1	52.5			-1.49	6.76E-02			do not reject H0
42	Northwestern Glaciated Plains	13	9.27	362	36	18	106	2.90	1.87E-03					reject H0
43	Northwestern Great Plains	81	7.45	43205	679	14	11794	8.41	0					reject H0

Level III Ecoregion	Ecoregion Name	WSA dataset			NOC database			M-W test (n1>10 and n2>10)		One large sample (n2>20)		Exact test (n2<20)	Exact test (n2<9)	Interpretation of test 1-sided @ 0.05
		n	median DOC	Ux	n	median DOC	Uy	Z	P	Z	P	critical U (0.05 level of signif.)	P	(H0: same mean of distributions)
44	Nebraska Sand Hills	1	6.46	0	4	3.2	4						0.2	do not reject H0
45	Piedmont	28	2.11	5379	308	3.4	3245	2.17	1.51E-02					reject H0 (P~1.5%)
46	Northern Glaciated Plains	17	14.62	1443	142	15	971	1.315	9.42E-02					do not reject H0
47	Western Corn Belt Plains	42	2.84	6200	193	5.1	1907	5.38	5.96E-08					reject H0
48	Lake Agassiz Plain	13	10.38	1962	261	10	1431	0.95	1.71E-01					reject H0 (P~1.7%)
49	Northern Minnesota Wetlands	1	11.71	37	44	17	7			-1.15	1.24E-01			do not reject H0
50	Northern Lakes and Forests	20	12.28	3099	403	8.1	4961	-1.74	4.05E-02					reject H0 (P~4%)
51	North Central Hardwood Forests	7	8.08	573.5	152	9.2	490.5			-0.35	3.64E-01			do not reject H0
52	Driftless Area	11	2.4	509	49	7.6	30	4.58	2.38E-06					reject H0
53	Southeastern Wisconsin Till Plains	5	3	2067	439	8	128.5			-3.40	3.41E-04			reject H0
54	Central Corn Belt Plains	9	2.69	1498	202	6.1	320			-3.29	5.07E-04			reject H0
55	Eastern Corn Belt Plains	6	2.83	7199	1325	6.8	751			-3.43	3.00E-04			reject H0

Level III Ecoregion	Ecoregion Name	WSA dataset			NOC database			M-W test (n1>10 and n2>10)		One large sample (n2>20)		Exact test (n2<20)	Exact test (n2<9)	Interpretation of test 1-sided @ 0.05
		n	median DOC	Ux	n	median DOC	Uy	Z	P	Z	P	critical U (0.05 level of signif.)	P	(H0: same mean of distributions)
56	Southern Michigan/Northern Indiana Drift Plains	9	4.62	1967	287	7	616			-2.67	3.77E-03			reject H0
57	Huron/Erie Lake Plains	3	5.05	8246	3762	7.1	3040			-1.38	8.33E-02			do not reject H0
58	Northeastern Highlands	23	3.54	79049	14044	1.4854	2E+05	-4.24	1.13E-05					reject H0
59	Northeastern Coastal Zone	10	4.01	561	101	4.2	449			-0.58	2.82E-01			do not reject H0
60	Northern Appalachian Plateau and Uplands	5	3.74	831	354	3.2	939			0.23	5.93E-01			do not reject H0
61	Erie Drift Plain	9	2.99	7834	901	6.2	275.5			-4.82	7.32E-07			reject H0
62	North Central Appalachians	4	3.34	102	106	1.7	322			1.76	9.60E-01			do not reject H0
63	Middle Atlantic Coastal Plain	2	18.54	16060	16726	3.4	17392			0.10	5.39E-01			do not reject H0
64	Northern Piedmont	15	2.18	18010	1524	4	4850	3.84	6.11E-05					reject H0
65	Southeastern Plains	18	2.55	49987	3801	4.3	18432	3.38	3.61E-04					reject H0
66	Blue Ridge	16	1.09	4915	686	0.9	6061	0.71	2.37E-01					do not reject H0
67	Ridge and Valley	27	1.56	10612	733	1.7	9180	0.64	2.61E-01					do not reject H0
68	Southwestern Appalachians	9	1.91	212.5	47	1.7	210.5			-0.02	4.91E-01			do not reject H0

Level III Ecoregion	Ecoregion Name	WSA dataset			NOC database			M-W test (n1>10 and n2>10)		One large sample (n2>20)		Exact test (n2<20)	Exact test (n2<9)	Interpretation of test 1-sided @ 0.05
		n	median DOC	Ux	n	median DOC	Uy	Z	P	Z	P	critical U (0.05 level of signif.)	P	(H0: same mean of distributions)
69	Central Appalachians	10	1.84	3651	864	1.6	4990			0.84	8.01E-01			do not reject H0
70	Western Allegheny Plateau	19	2.17	25890	1735	4	7075	4.28	9.18E-06					reject H0
71	Interior Plateau	14	2.24	2153	559	0.4	5674	-2.88	2.00E-03					reject H0
72	Interior River Valleys and Hills	14	3.86	3143	328	6.2	1449	2.34	9.70E-03					reject H0 (P~0.97%)
73	Mississippi Alluvial Plain	10	8.31	1646	503	5.5	3384			1.87	9.69E-01			do not reject H0
74	Mississippi Valley Loess Plains	1	1.26	21	21	5.4	0			-1.66	4.90E-02			reject H0 (P~4.9%)
75	Southern Coastal Plain	6	6.7	20141	4222	15.5	5191			-2.50	6.18E-03			reject H0
76	Southern Florida Coastal Plain													no test
77	North Cascades	54	0.82	1169	50	0.7	1531	1.18	1.19E-01					do not reject H0
78	Klamath Mountains	43	0.77	343	8	2.6	1			-4.43	4.74E-06			reject H0
79	Madrean Archipelago	3	1	27	9	7.9	0					3		reject H0
80	Northern Basin and Range	26	1.51	372.5	16	3.2	43.5	4.26	1.02E-05					reject H0
81	Sonoran Basin and Range	3	1.72	364	133	5.1	35			-2.44	7.40E-03			reject H0
82	Laurentian Plains and Hills	5	5.68	94	21	9.9	11			-2.70	3.47E-03			reject H0

Level III Ecoregion	Ecoregion Name	WSA dataset			NOC database			M-W test (n1>10 and n2>10)		One large sample (n2>20)		Exact test (n2<20)	Exact test (n2<9)	Interpretation of test 1-sided @ 0.05
		n	median DOC	Ux	n	median DOC	Uy	Z	P	Z	P	critical U (0.05 level of signif.)	P	(H0: same mean of distributions)
83	Eastern Great Lakes and Hudson Lowlands	12	6.72	8228	1346	7	7925	0.11	4.55E-01					do not reject H0
84	Atlantic Coastal Pine Barrens	2	12.62	150	243	5.7	336			0.93	8.24E-01			do not reject H0

4.3.3 Results and Implications of Bias Testing

The results of the rank-sum test indicate that DOC concentrations from the NOCD are biased high (i.e., the null hypothesis was not rejected) in 52 of 81 (64%) of Level III ecoregions in which comparable data were available (no comparison was possible in four ecoregions). For those ecoregions where the null hypothesis was not rejected, BLM users can be confident that the lower percentile DOC concentrations listed in Table 16 are representative for that ecoregion.

For ecoregions where the null hypothesis was rejected, the result suggests that the DOC data from the national organic carbon database are from biased samples. Recall discussion of both database in Sections 4.3, 4.3.1, and 4.3.2 that WSA is a random design sampling that ensures unbiased site selection. Whereas the NOCD is more influenced by locations with known water quality impairments and reflect unequal sampling efforts potentially creating a bias. It is likely that the percentile DOC concentrations tabulated for those ecoregions in Table 16 also reflect this bias towards high concentrations. This was confirmed by comparing the probability distributions of DOC concentrations in the ecoregions where n and m were large (n, m > 30).

In large-sample ecoregions where the null hypothesis was rejected by the rank-sum test (Ecoregions 1, 6, 11, 13, 23, 43, and 47), the probability distributions also show that the DOC concentration percentiles are substantially different, with the NOCD showing higher values. An example of such a comparison of DOC probability distribution is shown in Figure 31. On the other hand, in such large-sample ecoregions where the null hypothesis was not rejected (Ecoregions 17, 21, and 77), the probability distributions show that the DOC concentration percentiles are comparable. An example of such a probability distribution is shown in Figure 32. In all cases where it was possible to compare the DOC probability distributions, the results of the rank-sum test were confirmed.

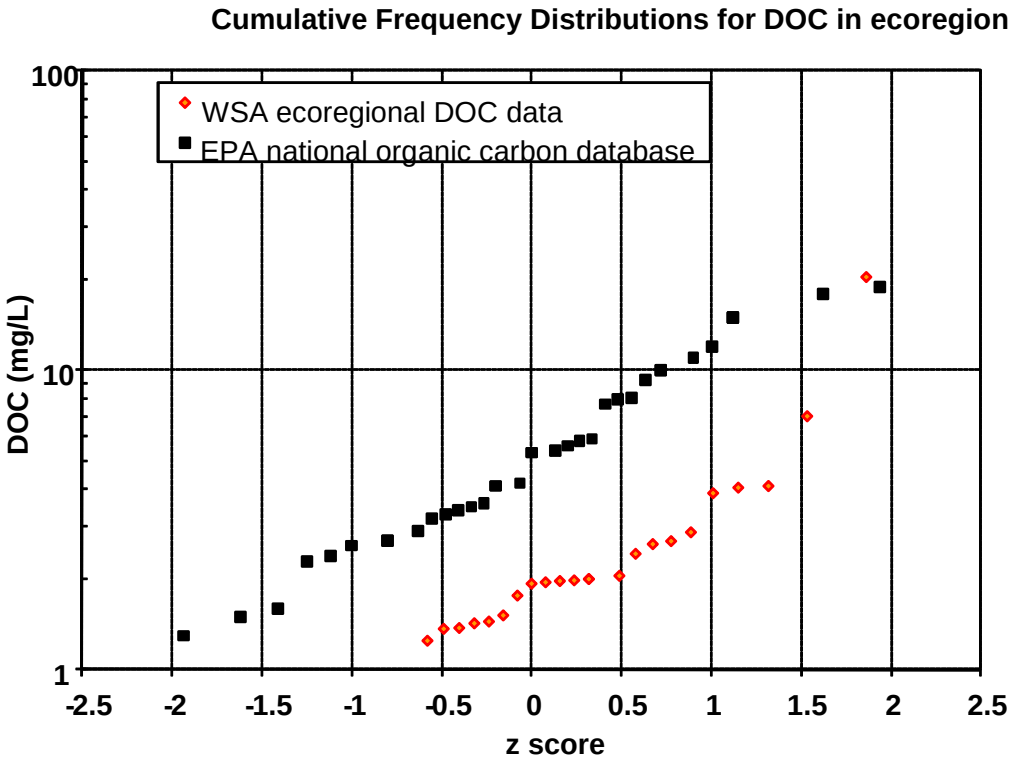

Figure 31. Comparison of probability distributions of DOC concentrations in Ecoregion 23

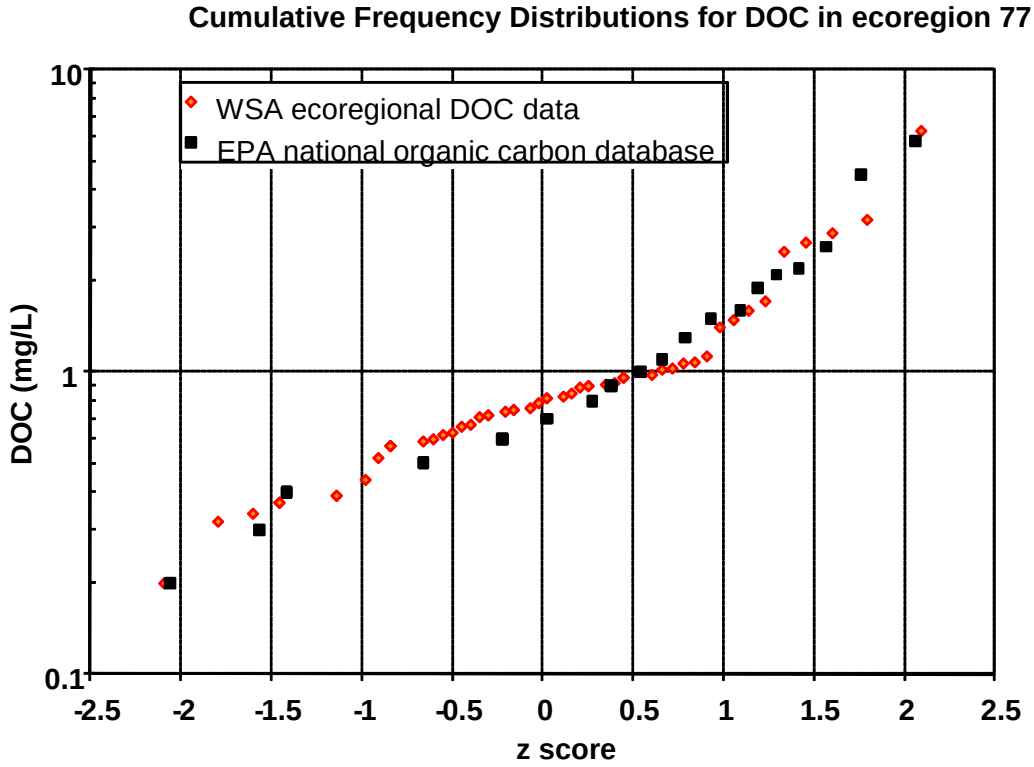

Figure 32. Comparison of probability distributions of DOC concentrations in Ecoregion 77

Because using a DOC concentration that is biased high as input to the BLM may lead to a non-conservative (high) site-specific copper criterion, it would be inappropriate to use the 10th percentile DOC concentrations in Table 16 for ecoregions in which the data from the NOCD come from biased samples.

We have not addressed the issue of whether the streams sampled for the WSA are representative in terms of DOC concentrations for all lotic (flowing) waters. It is possible that larger rivers may have DOC concentrations that are different from streams. For this reason, we recommend that the estimated ecoregional DOC values be compared to data from EPA's NRSA (NRSA, USEPA, 2013b; EPA 841-D-13-001). If necessary, adjustment of the estimated DOC values can be made at that time.

4.4 Comparing NOCD to WSA/NRSA DOC Data

The representativeness of the DOC data in the NOCD was evaluated by statistically comparing the data, at the ecoregional level, with the combined DOC data from two smaller random statistical surveys of rivers and streams. The two smaller surveys were:

(1) the 2004-05 WSA (1,313 sites), and

(2) the 2008-09 NRSA (2,113 sites).

The NRSA was the first nationally consistent survey assessing the ecological condition of the full range of flowing waters in the conterminous U.S. The target population includes the Great Rivers (such as the Mississippi and the Missouri), small perennial streams, and urban and non-urban rivers. Run-of-the-river ponds and pools are included, along with tidally influenced streams and rivers up to the leading edge of dilute sea water.

NRSA sampling locations were selected by random selection. The locations of perennial streams were identified using the EPA-USGS National Hydrography Dataset Plus (NHD-Plus), a comprehensive set of digital spatial data on surface waters at the 1:100,000 scale. Information about stream order was also obtained from the NHD-Plus. The 1,924 sites sampled for the NRSA were identified using a probability-based sample design. Details about the NRSA probabilistic sampling design are described in Section 1.1 of the NRSA: Field Operations Manual (USEPA, 2007; EPA-841-B-07-009). Site selection rules included weighting to provide balance in the number of river and stream sites from each of the size classes. Site selection was also controlled for spatial distribution to make sure sample sites were distributed across the U.S. Among these randomly selected sample sites were 359 of the original 2004 WSA sites. These were revisited as part of the NRSA to examine whether conditions have changed. When sites were selected for sampling, research teams conducted office evaluations and field reconnaissance to determine if the sites were accessible or if a river or stream labeled as perennial in NHD-Plus was, in fact, flowing during the sampling season. If a river or stream was not flowing or was determined to be inaccessible, it was dropped from the sampling effort and replaced with a perennial river or stream from a list of replacement sites within the random design.

The DOC data from these two smaller datasets were combined and described, hereafter, as WSA/NRSA data. GIS was used to determine which sampling sites were in each Level III ecoregion. The statistical test used was the non-parametric Wilcoxon 2-sample test, with a null hypothesis that DOC concentrations in the two different DOC sample datasets were equal. The alternative hypothesis was that DOC concentrations in the NCOD were significantly greater than those in the WSA/NRSA data,

indicating positive bias possibly due to over-representation of impacted sites. The test was applied for each of 84 Level III ecoregions at alpha=0.05.

Table 18 below includes the number of data (n) in each ecoregion, and the 10th percentiles of DOC based upon data from the EPA DOC database (NOCD) and the combined WSA/NRSA data. For each ecoregion, the table also provides the result of the Wilcoxon 2-sample test, in terms of whether the null hypothesis (the two samples are equal) should be accepted or rejected. The null hypothesis was rejected in 59 of 83 ecoregions, indicating bias in DOC concentrations higher in the national organic carbon dataset for the majority of ecoregions. In these 59 ecoregions, low-end percentiles based on DOC concentrations in the WSA/NRSA data were selected as reasonably protective estimates of ecoregional DOC concentrations.

Table 18. DOC concentrations (mg/L) in each Level III ecoregion based upon data from the NOCD and the combined WSA/NRSA data: number of data (n); 10th percentiles; and results of the Wilcoxon 2-sample test

Ecoregion	NOC database		WSA/NRSA database		H_0 (equal means)
	n	10%	n	10%	
1	91	1.1	60	0.7	reject
2	835	2.5	8	0.36	reject
3	66	1.1	12	0.4	reject
4	100	0.5	37	0.3	reject
5	32	2.1	21	0.5	reject
6	479	2.1	42	0.8	reject
7	180	2.7	7	1.1	reject
8	6	4.4	43	0.7	reject
9	13	1.4	25	0.5	reject
10	73	2.0	22	1.0	reject
11	26	1.3	91	0.8	reject
12	50	2.2	6	1.2	reject
13	1553	1.5	82	0.7	reject
14	35	2.8	8	0.8	reject
15	778	1.0	39	0.8	reject
16	29	1.4	34	0.8	reject
17	81	0.2	94	0.7	accept
18	150	4.3	52	1.1	reject
19	46	1.8	41	0.9	reject
20	798	3.0	61	1.2	reject
21	1129	0.6	76	0.8	accept
22	281	2.6	27	0.7	reject
23	37	2.2	48	0.7	reject
24	116	2.3	10	1.4	reject
25	439	4.4	29	1.3	reject

Ecoregion	NOC database		WSA/NRSA database		H_0 (equal means)
	n	10%	n	10%	
26	167	3.3	47	1.9	reject
27	228	3.8	92	2.2	reject
28	10	4.9	8	1.2	reject
29	289	3.8	30	1.7	reject
30	200	1.0	4	1.0	accept
31	58	1.0	4	0.3	accept
32	829	2.0	9	3.1	accept
33	268	3.0	5	5.6	accept
34	399	4.0	18	2.8	accept
35	523	4.6	66	3.3	reject
36	196	1.1	18	0.7	reject
37	184	1.9	18	1.4	reject
38	21	0.4	7	0.5	accept
39	233	2.3	39	0.8	reject
40	434	4.0	32	3.3	reject
41	36	0.6	7	0.6	reject
42	36	5.1	43	2.5	reject
43	679	6.2	234	2.3	reject
44	4	1.4	4	1.0	accept
45	308	1.0	93	1.1	reject
46	142	9.9	28	6.1	reject
47	193	3.1	103	1.7	reject
48	261	7.6	26	5.4	reject
49	44	11.0	9	6.0	accept
50	403	3.7	77	2.7	accept
51	152	3.1	44	3.2	reject
52	49	3.1	49	1.1	reject
53	439	5.3	12	1.9	reject
54	202	2.7	21	1.8	reject
55	1325	3.6	30	2.1	reject
56	287	3.8	38	2.9	reject
57	3762	4.7	14	1.5	accept
58	14044	0.6	92	1.2	accept
59	101	2.6	81	2.7	accept
60	354	1.3	29	1.4	reject
61	901	5.1	26	1.8	reject
62	106	0.7	22	0.9	accept
63	16726	2.2	45	1.7	accept
64	1524	1.3	47	1.0	reject
65	3801	2.4	108	1.4	reject
66	686	0.5	40	0.6	accept

Ecoregion	NOC database		WSA/NRSA database		H_0 (equal means)
	n	10%	n	10%	
67	733	0.6	88	0.9	accept
68	47	0.9	17	0.8	accept
69	864	0.7	31	1.1	accept
70	1735	1.5	67	1.5	reject
71	559	0.1	54	1.1	accept
72	328	2.7	65	2.2	reject
73	503	3.4	107	2.8	reject
74	21	1.7	18	1.2	accept
75	4222	8.0	41	3.6	reject
76	1	na	0	na	na
77	50	0.4	61	0.4	accept
78	8	1.7	56	0.6	reject
79	9	2.6	9	0.8	reject
80	16	1.8	49	1.0	reject
81	133	2.2	13	1.0	reject
82	21	5.5	18	2.8	reject
83	1346	1.0	32	2.6	reject
84	243	1.6	4	3.3	accept

In the 24 ecoregions where the null hypothesis was not rejected (i.e., no significant difference in DOC concentrations was found between datasets), the data were combined and the percentiles of the combined dataset were recalculated (Table 19). In these 24 ecoregions, low-end percentiles based on DOC concentrations in the combined data (NOCD and WSA/NRSA) were selected as reasonably protective estimates of ecoregional DOC concentrations.

Recommended DOC estimated values for 83 of the 84 ecoregions are summarized in Table 20. In the remaining ecoregion (76; Southern Florida Coastal Plain), there were insufficient data in either dataset (NOC database or WSA/NRSA) to calculate DOC concentration percentiles.

Table 19. DOC concentrations (mg/L) in 24 ecoregions where no significant difference in DOC concentrations was found between national organic carbon database (NOCD) and the WSA/NRSA datasets: number of data (n); 10th percentiles from combined NOCD & WSA/NRSA data

Ecoregion	n	DOC (mg/L) 10%
17	175	0.6
21	1205	0.6
30	204	1.0
31	62	1.0
32	838	2.0

Ecoregion	n	DOC (mg/L) 10%
33	273	3.0
34	417	4.0
38	28	0.5
44	8	1.0
49	53	10.4
50	480	3.5
57	3776	4.6
58	14136	0.6
59	182	2.7
62	128	0.7
63	16771	2.2
66	726	0.5
67	821	0.6
68	64	0.9
69	895	0.7
71	613	0.1
74	39	1.5
77	111	0.4
84	247	1.6

Table 20. Recommended ecoregional DOC concentrations (mg/L) based upon combined data from the NOCD and the WSA/NRSA data in 83 Level III ecoregions: number of observations (n); 10th percentiles; and source of data for each ecoregion

Ecoregion	n	DOC (mg/L) 10%	Data Source
1	60	0.7	WSA/NRSA
2	8	0.3	WSA/NRSA
3	12	0.4	WSA/NRSA
4	37	0.3	WSA/NRSA
5	21	0.5	WSA/NRSA
6	42	0.8	WSA/NRSA
7	7	1.1	WSA/NRSA
8	43	0.7	WSA/NRSA
9	25	0.5	WSA/NRSA
10	22	1.0	WSA/NRSA
11	91	0.8	WSA/NRSA
12	6	1.2	WSA/NRSA
13	82	0.7	WSA/NRSA
14	8	0.8	WSA/NRSA
15	39	0.8	WSA/NRSA

Ecoregion	n	DOC (mg/L) 10%	Data Source
16	34	0.8	WSA/NRSA
17	175	0.6	NOCD & WSA/NRSA
18	52	1.1	WSA/NRSA
19	41	0.9	WSA/NRSA
20	61	1.2	WSA/NRSA
21	1205	0.6	NOCD & WSA/NRSA
22	27	0.7	WSA/NRSA
23	48	0.7	WSA/NRSA
24	10	1.4	WSA/NRSA
25	29	1.3	WSA/NRSA
26	47	1.9	WSA/NRSA
27	92	2.2	WSA/NRSA
28	8	1.2	WSA/NRSA
29	30	1.7	WSA/NRSA
30	204	1.0	NOCD & WSA/NRSA
31	62	1.0	NOCD & WSA/NRSA
32	838	2.0	NOCD & WSA/NRSA
33	273	3.0	NOCD & WSA/NRSA
34	417	4.0	NOCD & WSA/NRSA
35	66	3.3	WSA/NRSA
36	18	0.7	WSA/NRSA
37	18	1.4	WSA/NRSA
38	28	0.5	NOCD & WSA/NRSA
39	39	0.8	WSA/NRSA
40	32	3.3	WSA/NRSA
41	7	0.6	WSA/NRSA
42	43	2.5	WSA/NRSA
43	234	2.3	WSA/NRSA
44	8	1.0	NOCD & WSA/NRSA
45	93	1.1	WSA/NRSA
46	28	6.1	WSA/NRSA
47	103	1.7	WSA/NRSA
48	26	5.4	WSA/NRSA
49	53	10.4	NOCD & WSA/NRSA
50	480	3.5	NOCD & WSA/NRSA
51	44	3.2	WSA/NRSA
52	49	1.1	WSA/NRSA
53	12	1.9	WSA/NRSA
54	21	1.8	WSA/NRSA
55	30	2.1	WSA/NRSA
56	38	2.9	WSA/NRSA
57	3776	4.6	NOCD & WSA/NRSA

Ecoregion	n	DOC (mg/L) 10%	Data Source
58	14136	0.6	NOCD & WSA/NRSA
59	182	2.7	NOCD & WSA/NRSA
60	29	1.4	WSA/NRSA
61	26	1.8	WSA/NRSA
62	128	0.7	NOCD & WSA/NRSA
63	16771	2.2	NOCD & WSA/NRSA
64	47	1.0	WSA/NRSA
65	108	1.4	WSA/NRSA
66	726	0.5	NOCD & WSA/NRSA
67	821	0.6	NOCD & WSA/NRSA
68	64	0.9	NOCD & WSA/NRSA
69	895	0.7	NOCD & WSA/NRSA
70	67	1.5	WSA/NRSA
71	613	0.1	NOCD & WSA/NRSA
72	65	2.2	WSA/NRSA
73	107	2.8	WSA/NRSA
74	39	1.5	NOCD & WSA/NRSA
75	41	3.6	WSA/NRSA
77	111	0.4	NOCD & WSA/NRSA
78	56	0.6	WSA/NRSA
79	9	0.8	WSA/NRSA
80	49	1.0	WSA/NRSA
81	13	1.0	WSA/NRSA
82	18	2.8	WSA/NRSA
83	32	2.6	WSA/NRSA
84	247	1.6	NOCD & WSA/NRSA

4.5 Conclusions

EPA tested the 10th percentiles of ecoregional DOC concentrations against data from the Southern Rocky Mountains (Level III Ecoregion 21) as input to the copper BLM. Broad ranges of errors (including some that were larger than an order-of magnitude) were observed in BLM predictions made with the DOC estimates, in comparison to predictions made with actual measured site data. Although the copper criteria values predicted using the parameter estimates for DOC were found to be protective in 90% of the cases, in many of these cases these predictions were overly-protective (e.g., IWQC lower by a factor of 4 to 5). For this reason, BLM users should be cautious when considering lower percentiles of the distribution of DOC as estimates for missing input parameters to the BLM. In general, it is preferable to use site-specific measurements of DOC as BLM input because: (1) copper toxicities (and BLM model predictions) are highly sensitive to DOC concentrations and (2) reasonably protective DOC concentrations can be difficult to estimate at the ecoregional level, when data are limited.

For many ecoregions, the EPA recommended percentiles in Table 20 are based upon a relatively small number of DOC data, which can be a cause for concern in terms of the reliability of these values. For

example, in 47 ecoregions the DOC percentiles were calculated from 50 or fewer concentration values, and in seven ecoregions the DOC percentiles were calculated from fewer than 10 values. In the former case (n≤50), the lower 95% confidence limit of the 10th percentile cannot be calculated (Berthouex and Brown, 1994), while in the latter (n<9) the 10th percentile itself is below the lowest concentration value. Because of these and other limitations on the DOC database and the importance of this parameter in criteria calculation, users are encouraged to sample for DOC as a basis for determining BLM input rather than using default parameters where possible.

5 SUMMARY AND RECOMMENDATIONS

The BLM predicts acute copper toxicity based on site-specific water quality parameters, and calculates aquatic life criteria based on the predicted copper toxicity. The BLM requires 10 input parameters to calculate copper criteria: temperature, pH, DOC, alkalinity, calcium, magnesium, sodium, potassium, sulfate, and chloride, the last seven of which are also referred to as GIs. Given the broad geographical range over which the BLM is likely to be applied, and the limited availability of data for input parameters in many areas, a practical method to estimate missing water quality parameters was developed to support the use of the copper BLM for copper aquatic life criteria.

In this report we described three approaches EPA used to estimate default input parameters for GI and DOC for BLM that could be used where site-specific data are not available. EPA's goal was to provide estimates for these missing input parameters that are reasonably protective. EPA used geostatistics to predict ecoregional input parameters from national water quality databases, and developed correlations between GI parameters and conductivity. These estimates were further refined using stream order.

Our analysis of national data indicates that there is no relationship between conductivity and pH, and geostatistical methods were found to produce similarly ambiguous results. Because pH is one of the most important BLM inputs for predicting criteria for copper, we conclude that site-specific data for pH are needed for successful BLM application. Temperature is a commonly measured parameter and should be easy to obtain by users for input in the BLM.

5.1 Recommendations for BLM inputs for geochemical ions where site-specific data are not available

In Section 2 we used geostatistics to estimate missing GI parameter values based on geography. We supplemented the geostatistical approach by adding conductivity as an additional explanatory variable to generate a more robust spatial estimate of the GI water quality inputs for the BLM because conductivity is one of the most widely monitored water quality indicators in the U.S. and correlates well with GIs. We presented average predicted 10th percentile concentrations for the BLM GI water quality parameters Level III ecoregions. We further refined these estimates by considering the effect of stream order (size) in Section 3. We found that values of the GI estimates generally increased with stream order, a trend that was most apparent and consistent for higher order streams. Tables 8, 9, and 10 present best estimates of GI input parameters for the BLM. Estimated inputs are provided for each GI in each ecoregion categorized by stream order for low, medium, and high order streams, respectively. EPA recommends these 10th percentile Level III ecoregion, stream order group-specific values be used in the BLM where site-specific data are not available.

5.2 Recommendations for BLM inputs for DOC where site-specific data are not available

In Section 4 we determined that the geostatistical and regression-based approaches used to estimate GI input parameters for the BLM do not produce accurate site-specific estimates for DOC. Because previous analyses indicate that DOC is the most important BLM input for estimating criteria for copper, we further refined our approach in Section 4 based on analyses using the NOCD to estimate lower-percentile DOC concentrations. Based on statistical comparisons to an independent probabilistic dataset, we found that DOC concentrations from the NOCD are reasonably protective estimates of DOC for use as input parameters for the BLM for some ecoregions. For other ecoregions, EPA recommends using estimates based on the WSA dataset. Recommended 10^{th} percentile DOC estimated values for 83 of the 84 ecoregions are summarized in Table 20. In the remaining ecoregion (76; Southern Florida Coastal Plain), there were insufficient data in either dataset (NOC database or WSA/NRSA) to calculate DOC concentration percentiles. Because limitations in the DOC database and the importance of this parameter in criteria calculation, users are encouraged to sample for DOC as a basis for determining BLM input rather than using default parameters wherever possible.

5.3 Recommendations for BLM inputs for pH where site-specific data are not available

In Section 2 we determined that geostatistical and regression-based approaches used to estimate GI input parameters for the BLM did not produce accurate site-specific estimates for pH. Our analysis of national data indicates that there is no relationship between conductivity and pH, and geostatistical methods were found to produce similarly ambiguous results. Because pH is one of the most important BLM inputs for predicting criteria for copper, we conclude that site-specific data for pH are needed for successful BLM application. Temperature along with pH is similarly recommended to acquire site-specific data for BLM application with the advantage of both of these been easy parameters to measure.

5.4 Conclusions

The approaches described in this TSD can be used to provide reasonable default values for input parameters in the BLM to derive protective freshwater aquatic life criteria for copper when data are lacking. These data could also be used to provide reasonable default values to fill in missing water quality input parameters in the application of other metal BLM models as well when data are lacking. Default recommended values for GI parameters are 10^{th} percentile ecoregional, stream-order specific values. Default recommended values for DOC are 10^{th} percentile ecoregional values. Both pH and temperature should be measured values when using the BLM. It should be noted that site-specific data are always preferable for use in the BLM and should be used to develop copper criteria via the BLM when possible. Users of the BLM are encouraged to sample their water body of interest, and to analyze the samples for the constituent (parameter) concentrations as a basis for determining BLM inputs where possible.

REFERENCES

Berthouex, P.M. and L.C. Brown. 1994. Statistics for Environmental Engineers. Lewis Publishers. Boca Raton, FL. 335p.

Carleton, J.N. 2006. An Examination of Spatial Trends in Surface Water Chemistry in the Continental United States: Implications for the Use of Default Values as Inputs to the Biotic Ligand Model for Prediction of Acute Metal Toxicity to Aquatic Organisms U.S. Environmental Protection Agency, Office of Water, Office of Science & Technology. (Appendix A to this report).

Clements, W.H., Brooks, M.L., Kashian, D.R. and R.E. Zuellig. 2008. Changes in dissolved organic material determine exposure of stream benthic communities to UV-B radiation and heavy metals: Implications for climate change. Global Change Biology 14:2201-2214.

Clements, W.H., Carlisle, D.M., Lazorchak, J.M. and P.C. Johnson. 2000. Heavy metals structure benthic communities in Colorado mountain streams. Ecological Applications 10:626-638.

Dierickx, T. 2008. Computing percentiles – are your values correct? (http://www.data-for-all.com/documents/computing-percentiles.pdf)

ESRI. 2003. Using ArcGIS Geostatistical Analyst (ArcGIS 9). Environmental Systems Research Institute, Inc. Redlands, California.

FISRWG. 1998. Stream Corridor Restoration: Principles, Processes, and Practices. By the Federal Interagency Stream Restoration Working Group (FISRWG)(15 Federal agencies of the US government). GPO Item No. 0120-A; SuDocs No. A 57.6/2:EN 3/PT.653. ISBN-0-934213-59-3.

Gibbs, R. J. 1970. Mechanisms controlling world water chemistry. Science 170:1088–1090.

Griffith, M.B. 2014. Natural variation and current reference for specific conductivity and major ions in wadeable streams of the conterminous USA. Freshwater Science. 33(1):1-17.

Helsel, D.R. and R.M. Hirsch. 2002. Techniques of Water-Resources Investigations of the United States Geological Survey Book 4, Hydrologic Analysis and Interpretation. Section A3, Statistical Methods. U.S. Geological Survey. September 2002. Publication available at: http://pubs.usgs.gov/twri/twri4a3/ (last accessed February, 2016).

Herlihy, A.T., D.P. Larsen, S.G. Paulsen, N.S. Urquhart, and B.J. Rosenbaum. 2000. Designing a spatially balanced randomized site selection process for regional stream surveys: the EMAP mid-Atlantic pilot study. Environmental Monitoring and Assessment 63:95–113.

HydroQual, Inc. 2001. BLM-Monte User's Guide, Version 2.0. HydroQual, Mahwah, NJ. October, 2001.

Hyndman, R.J., and Y. Fan. 1996. Sample quantiles in statistical packages. American Statistician. 50(4): 361-365.

Linton, T.K., W.H. Clement, W.F. Dimond, G.M. DeGraeve, and G.W. Saalfeld. 2007. Development of a copper criteria adjustment procedure for Michigan Upper Peninsula waters. Proceedings of the 80th Annual Water Environment Federation Technical Exhibition and Conference. San Diego, CA.

MacArthur, R.H. 1972. Geographical Ecology. New York: Harper & Row.

McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., and Rea, A. 2012. "NHDPlus Version 2: User Guide" (http://nhd.usgs.gov/).

National Research Council. 1992. Restoration of Aquatic Ecosystems. Committee on Restoration of Aquatic Ecosystems: Science, Technology, and Public Policy. National Academy Press, Washington, DC.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118-125.

Omernik, J., 2003. The misuse of hydrologic unit maps for extrapolation, reporting, and ecosystem management. Journal of the American Water Resources Association 39(3):563-573.

Omernik, J.M., and G.E. Griffith. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental Management.

Perry, J. and E.L. Vanderklein. 1996. Water Quality: Management of a Natural Resource. Wiley-Blackwell. 656 p.

Smith, RA, Schwarz, GE, Alexander, RB. 1997. Regional Interpretation of Water-Quality Monitoring Data. Water Resources Research. (33):2781–2798.

Stephan, C.E., D.I. Mount, D.J. Hansen, J.H. Gentile, G.A. Chapman and W.A. Brungs. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. PB 85—227049. National Technical Information Service, Springfield, VA.

Strahler, A.N. (1952). Hypsometric (area-altitude) analysis of erosional topology. Geological Society of America Bulletin. 63(1):1117-1142.

Strahler, A.N. (1957). Quantitative analysis of watershed geomorphology. Transactions of the American geophysical Union. 38(6):913-920.

Thurman, E.M. 1985. Organic geochemistry of natural waters. Martinus Nijhoff / DR W. Junk Publishers. 489pp.

USEPA. 2002. Development of Methodologies for Incorporating the Copper Biotic Ligand Model into Aquatic Life Criteria: Application of BLM to Calculate Site-Specific Fixed Criteria. Prepared by Great Lakes Environmental Center (GLEC) for U.S. Environmental Protection Agency, Office of Science and Technology, Health and Ecological Criteria Division, Work Assignment 3-38, Contract No. 68-C-98-134. 63p plus figures.

USEPA. 2003. Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health (2000), Technical Support Document Volume 2: Development of National Bioaccumulation Factors (EPA-822-R-03-030). December 2003.

USEPA. 2006a. Approaches for Estimating Missing BLM Input Parameters: Projections of Total Organic Carbon as a Function of Biochemical Oxygen Demand. Prepared by Great Lakes Environmental Center (GLEC) for U.S. Environmental Protection Agency, Office of Science and Technology, Health and Ecological Criteria Division, Contract No. 68-C-04-006, Work Assignment 2-34, Task 1, Subtask 1-7. Report: December 7, 2006. (Appendix B to this report).

USEPA. 2006b. Mid-Atlantic Integrated Assessment (MAIA). State of the Flowing Waters Report. United States Environmental Protection Agency, Office of Research and Development. Washington, DC 20460. EPA/620/R-06/001. February 2006.

USEPA. 2006c. Wadeable Streams Assessment, A Collaborative Survey of the Nation's Streams. United States Environmental Protection Agency, Office of Research and Development and Office of Water. Washington, DC 20460. EPA 841-B-06-002, December 2006.

USEPA. 2007. Approaches for Estimating Missing BLM Input Parameters: Correlation approaches to estimate BLM input parameters using conductivity and discharge as explanatory variables. Prepared by Great Lakes Environmental Center (GLEC) for U.S. Environmental Protection Agency, Office of Science and Technology, Health and Ecological Criteria Division, Contract No. 68-C-04-006, Work Assignment 2-34, Task 1, Subtask 1-7. (Appendix C to this report).

USEPA. 2007. National Rivers and Streams Assessment: Field Operations Manual. EPA-841-B-07-009. U.S. Environmental Protection Agency, Washington, DC.

USEPA. 2008. Copper Biotic Ligand Model (BLM) Software and Supporting Documents Preparation, Task 3c: Development of Tools to Estimate BLM Parameters. Prepared by Great Lakes Environmental Center (GLEC) for U.S. Environmental Protection Agency, Office of Science and Technology, Health and Ecological Criteria Division. Contract No. 68-C-04-006, Work Assignment 4-18, Task 3 Progress Report: May 22, 2008. (Appendix D to this report).

USEPA. 2013a. U.S. Environmental Protection Agency, 2013, Level III ecoregions of the continental United States: Corvallis, Oregon, U.S. EPA - National Health and Environmental Effects Research Laboratory, map scale 1:7,500,000. ftp://ftp.epa.gov/wed/ecoregions/us/Eco_Level_III_US.pdf

USEPA. 2013b. National Rivers and Streams Assessment 2008–2009. A Collaborative Survey (Draft). U.S. Environmental Protection Agency. Office of Wetlands, Oceans and Watersheds. Office of Research and Development. Washington, DC. 20460 (EPA/841/D-13/001). February 28, 2013.

USEPA. 2015. Connectivity of Streams and Wetlands to Downstream Waters: A Review and Synthesis of the Scientific Evidence (Final Report). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-14/475F.

USGS. 2012. National Hydrography Geodatabase: The National Map viewer available on the World Wide Web (http://viewer.nationalmap.gov/viewer/nhd.html?p=nhd), accessed, 2012.

Walpole, R.E. and R.M. Myers. 1978. Probability and Statistics for Engineers and Scientists. MacMillan Publishing Co., New York. 580 p.

Ward, J. V. 1992. A mountain river. Pages 793-510 in Calow, P., and G. E. Petts (Editors). The Rivers Handbook. Blackwell, Oxford, U.K.

Appendix A: An Examination of Spatial Trends in Surface Water Chemistry in the Continental United States: Implications for the Use of Default Values as Inputs to the Biotic Ligand Model for Prediction of Acute Metal Toxicity to Aquatic Organisms

Internal EPA Report (2006)
James N. Carleton
EPA, Office of Water, Office of Science & Technology.

A.1 Abstract

A large database of surface water chemistry monitoring data was examined to look for spatial trends in five chemical constituents that are key inputs to a model for predicting metal toxicity to aquatic organisms. Continuous prediction maps of concentrations were generated using various kriging techniques to interpolate between site-median values measured at several thousand separate locations throughout the continental United States (U.S.). Continuous concentration surfaces were then averaged over 8-digit Hydrologic Unit Code (HUC) polygons to produce block-averaged mean estimates of site-median concentrations. Pairwise comparisons indicated distinct trends between various HUC-averaged predicted constituents. The same analyses performed on data from 772 locations where all five constituents had been measured revealed similar relationships between monitored constituents. Principal components analyses performed on these data sets showed that 80 to 90% of the variance in both cases could be explained by a single component with loadings on three of the five constituents. The use of kriging to produce appropriate quantile maps for block-averaging is suggested as a possible approach for developing regional values to use as default model inputs, when site-specific monitoring data are lacking.

A.2 Background

The U.S. Environmental Protection Agency is planning in the near future to release proposed water quality criteria for copper (*Note: EPA's BLM-based Freshwater Copper Aquatic Life Ambient Water Quality Criteria document was released in 2007, EPA-822-R-07-001*). These criteria are unlike most water quality criteria in that acceptable (safe) concentrations for aquatic life support, rather than being defined as simple numerical values that apply everywhere, will be addressed through the use of a chemical speciation model – the Biotic Ligand Model (BLM) (EPA, 2003). The BLM calculates metal toxicity to aquatic organisms as a function of simultaneous concentrations of additional chemical constituents of water, for example other ions that can either complex with copper and render it biologically unavailable, or compete with copper for binding sites at the point of entry into a vulnerable organism (i.e. at the fish gill). While the BLM has the potential to improve the accuracy of metal ecotoxicity predictions, its use requires input concentrations of nine separate chemical constituents and water temperature. Of these nine chemical constituents (Alkalinity (alk), calcium (Ca), magnesium (Mg), sodium (Na), sulfate (SO42-), potassium (K), chloride (Cl), dissolved organic carbon (DOC), and pH), model-predicted toxicity is most sensitive to five: Ca, alk, pH, Na, and DOC. States or other entities wishing to use the BLM to assess compliance with the proposed criteria in specific waters, or to develop effluent permit limits, will therefore require monitoring information on a suite of chemical

constituents – information that is not always available. One possible way to deal with such missing information is to develop reasonably protective default values for these various model inputs, especially the five to which the BLM is most sensitive. Given that ambient surface water chemistry reflects, among other things, the influences of local soil types and land uses, it may make sense that any such defaults be developed on some kind of regional or local basis.

The exercise described in this report comprises a geospatial examination of a large amount of water chemistry monitoring data collected in recent years by the U.S. Geological Survey, and recorded in their National Water Information System (NWIS) database. The data includes monitoring information from several thousand separate surface water sampling locations throughout the U.S. (Figure A-1). The latitudes and longitudes of each sampling location are part of the data record. The primary objective of this analysis is to look for any obvious spatial trends in typical concentrations of the five most sensitive constituents, and to suggest procedures for making use of these trends to define regional default values for use as inputs to the BLM. For purposes of expediency, the geographic extent of this analysis is limited to the continental U.S.

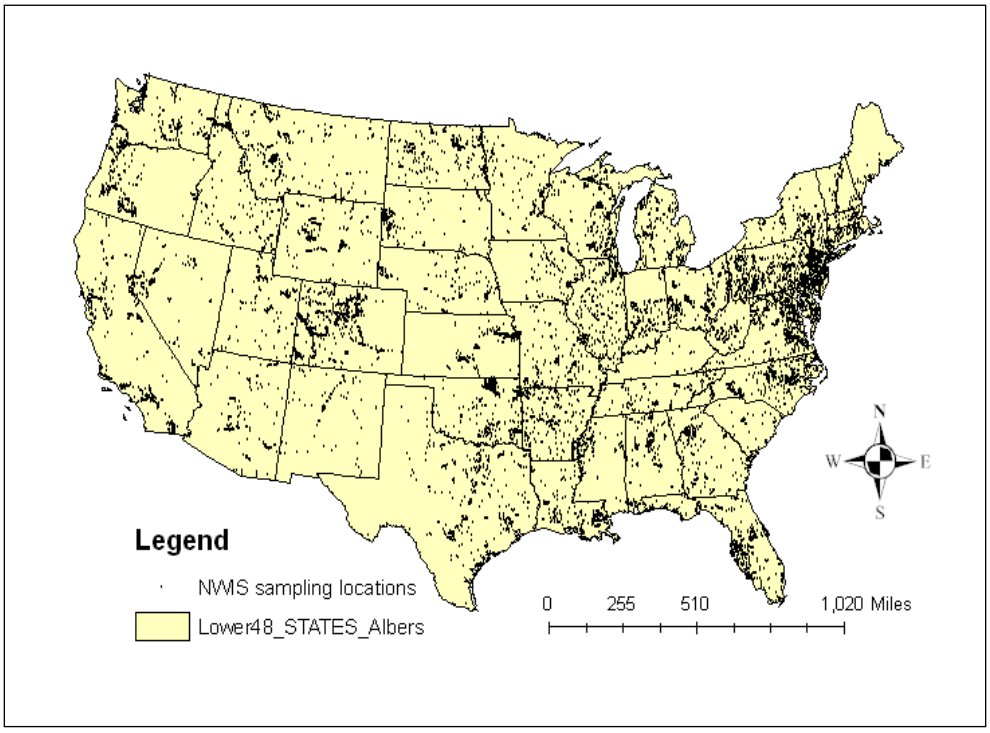

Figure A-1. NWIS sample collection locations in the continental U.S.

A.3 Description of Data

Although NWIS contained data from 207,153 sampling events at 13,824 individual sampling locations in the continental U.S. (Figure A-1), all 10 constituents of relevance to the BLM were not monitored at each location. For the five constituents of interest, the numbers of discrete sampling locations were as follows: alk, 5,900; Ca, 10,940; DOC, 3,726; Na, 10,424; pH, 11,780. Numbers of sampling events at individual locations ranged from 1 to 2,605, with a mean of 15, and a mode of one (i.e. most sites were

only sampled once). Examination of the spatial distribution of numbers of sampling events per site reveals that the most intensive sampling tended to occur in Midwestern and western states (Figure A-2). Because environmental sampling data tend to be lognormally distributed, disparities in numbers of samples may tend to produce higher mean and median values at more-frequently-sampled locations. As spatial distributions of representative (e.g., median) concentrations are examined, it should be kept in mind that apparent geographic trends in concentration may be in part simply the result of uneven sampling intensity.

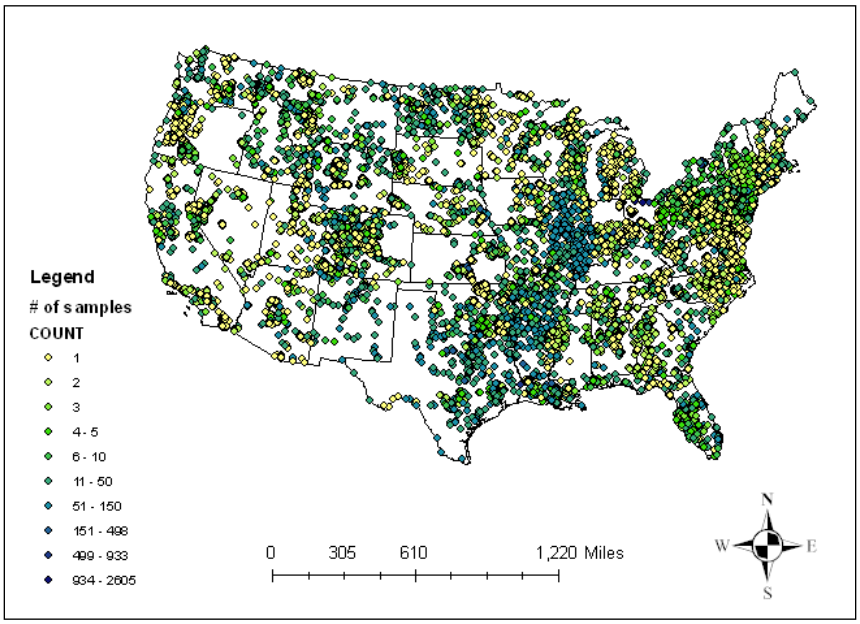

Figure A-2. Intensity of sampling (number of separate sampling dates) at each NWIS site

A.4 Data Analysis

Because environmental data tend to be positively skewed, the median statistic was chosen as providing the best central-tendency representation of each location's concentration. For the purpose of looking for general spatial trends in the five constituents, the first step involved simply mapping the sampling locations as points, color-graded by median concentration. Figure A-3, for example, shows some apparent trends in alkalinity across the country, with lower concentrations along the eastern seaboard, and higher concentrations in parts of the Midwest. Similar kinds of trends at the national scale were also seen with the other constituents.

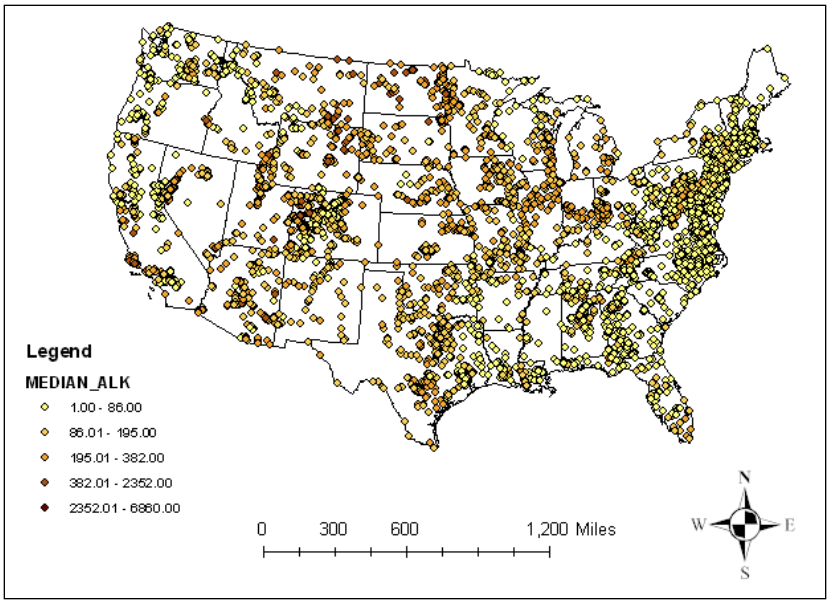

Figure A-3. Median measured alkalinity (mg/L as CaCO3) at NWIS locations

The next step in data visualization involved the calculation of median concentrations averaged over each 8-digit HUC containing sampling locations. These display essentially the same information as the point displays (Figure A-3), but with a degree of smoothing and summarization provided by the spatial averaging process, to make visual interpretation of general trends easier (Figure A-4).

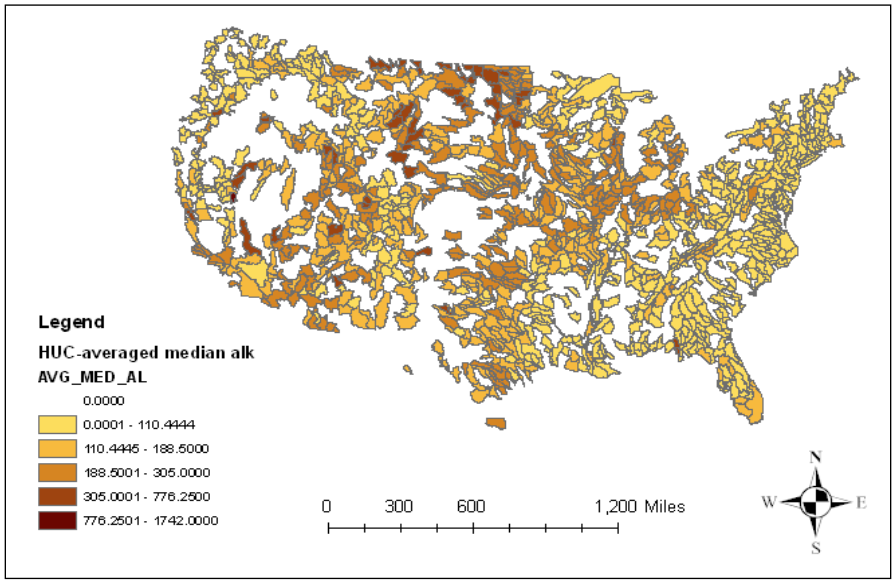

Figure A-4. HUC-averaged mean median observed alkalinity in the continental U.S.

The use of 8-digit HUCs as the areal units over which to calculate representative concentrations for default BLM inputs makes some physical sense: HUCs are areas that are defined by some degree of

interconnection between associated surface water features. HUCs may be either watersheds in their own right or downstream sections of larger watersheds (Omernik, 2003). In either case, all flowing surface water that passes through a HUC eventually (in theory) passes through the same downstream "pour point". One advantage of using HUCs is that they divide the land area into roughly equally sized areas at a level of resolution roughly consistent with gross variations in median concentration (Figure A-3). One problem with using HUCs for spatial aggregation is that not all HUCs contain NWIS sampling locations, as the blank areas in Figure A-4 make clear. The third step in this analysis therefore involved the use of kriging to create continuous surfaces of interpolated concentrations that cover the entire area of interest. Spatial averaging of the results over each HUC was then used to provide estimates of expected concentrations for all HUCs, including those lacking NWIS samples.

For each of the five key constituents, the Geostatistical Analyst extension in ArcGIS was used to explore the data, and to look for sets of kriging model options that provided the best fit to the data. The criteria used to evaluate goodness of fit were as follows:

1. Mean Standardized Error as small as possible
2. RMSE as small as possible
3. Root-Mean-Square Standardized Error close to 1.0
4. RMSE and Average Standard Error close together

Trial and error parameter selection was used to search for a set of model options that best attained each of these four goals simultaneously. For each constituent, 10 to 20 combinations were tried, until a best option for each emerged, as determined by judgment of the author. The results are as follows:

Alk: Universal kriging, log transformation, constant trend, 50% global, 50% local, spherical semivariogram, no anisotropy.

Ca: Ordinary kriging, log transformation, constant trend, 50% global, 50% local, exponential semivariogram, anisotropy.

DOC: Universal kriging, log transformation, constant trend, 50% global, 50% local, hole-effect semivariogram, anisotropy.

Na: Universal kriging, log transformation, constant trend, 50% global, 50% local, hole-effect semivariogram, anisotropy.

pH: Ordinary kriging, no transformation, constant trend, 50% global, 50% local, spherical semivariogram, no anisotropy.

Prediction surface maps were generated for each constituent using the above sets of kriging options. Figure A-5, which displays the results for alkalinity, shows patterns that are generally consistent with those in the data (Figures A-3 and A-4). Figure A-6 shows the predicted alkalinities projected into three dimensions using ArcScene.

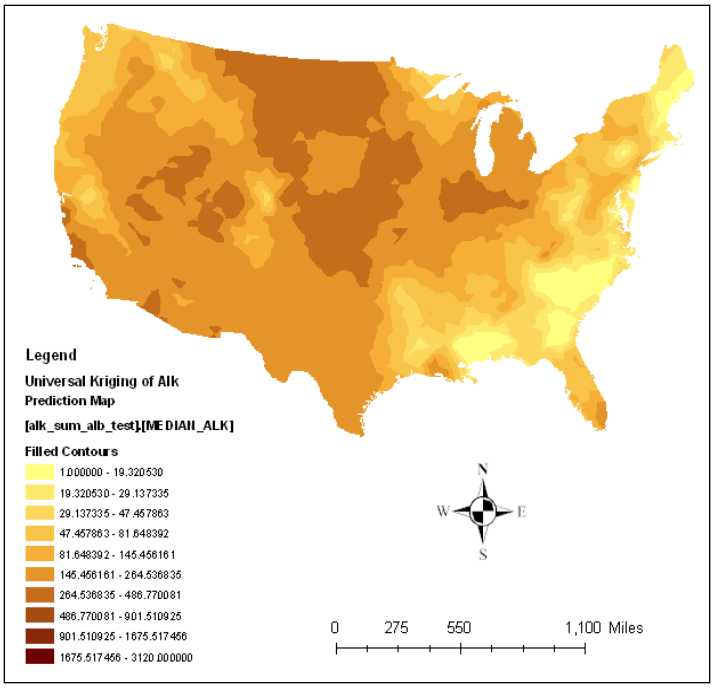

Figure A-5. Kriging prediction map of median alkalinity

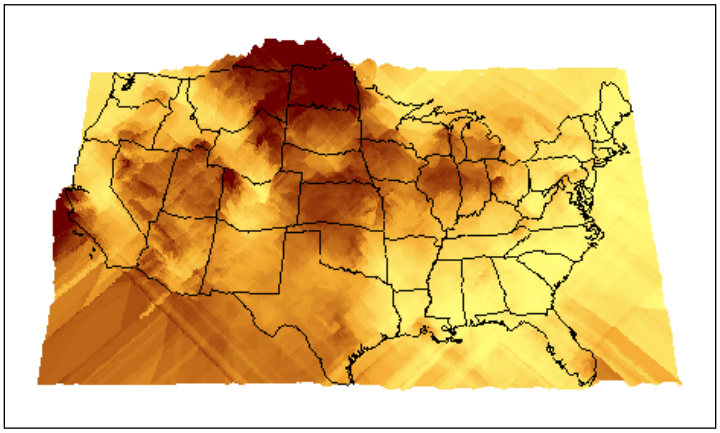

Figure A-6. Kriging map of alkalinity, projected into vertical dimension

This technique demonstrates broad geographic trends most dramatically, for example emphasizing the fact that the highest alkalinities are apparently found in northern North Dakota and Montana. Figure A-7 shows the predicted values averaged over HUC polygons by using the Zonal Statistics function of ArcGIS Spatial Analyst.

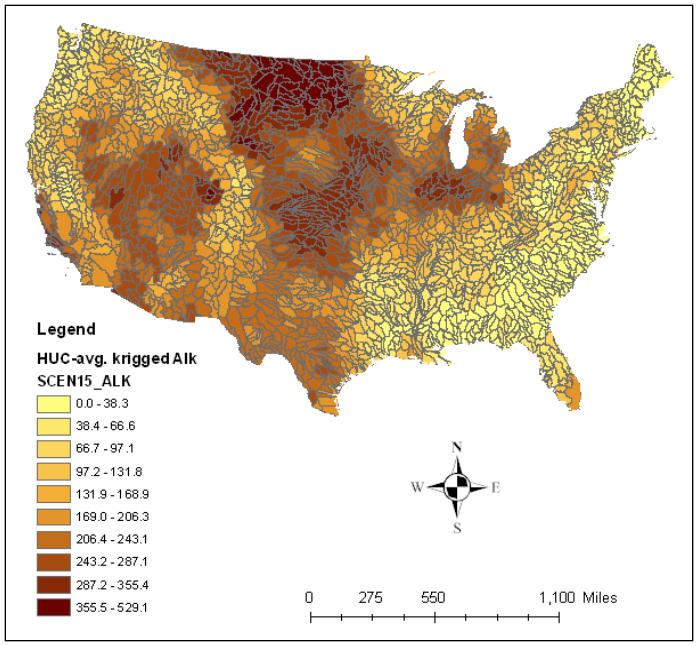

Figure A-7. Kriging-based alkalinity predictions, averaged over 8-digit HUC polygons

For HUCs containing NWIS sampling locations, linear regression plots of predicted versus measured concentrations (Figure A-8) provided a check on the accuracy of the kriging predictions. R-squared values for the five constituents were: 0.537 (alk), 0.238 (Ca), 0.686 (DOC), 0.351 (Na), and 0.139 (pH). In most cases, a handful of outliers appeared to be responsible for smaller-than-expected correlation coefficients.

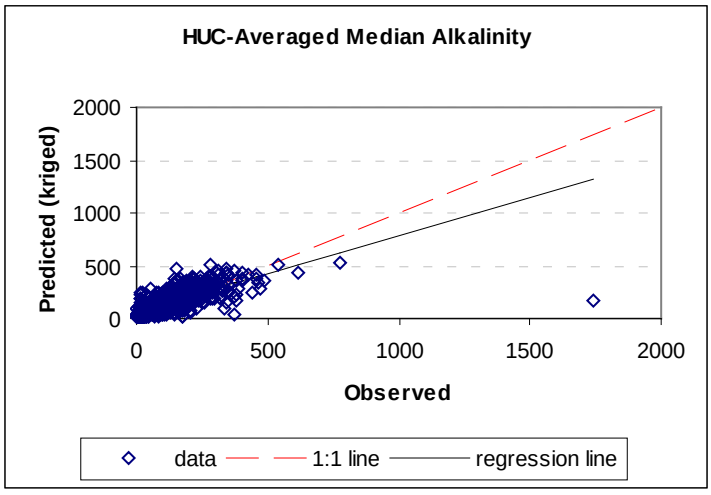

Figure A-8. Kriging-predicted vs. calculated HUC-averaged alkalinity; r^2=0.537

A scatter plot matrix of cross-constituent comparisons revealed some interesting, non-random relationships between HUC-averaged concentrations (Figure A-9). For comparative purposes, a subset of 772 sampling locations was also identified, at which sampling for all five of the constituents had taken place. Coincident concentrations of all constituents allowed a scatter plot matrix of this data (Figure A-10) to also be constructed. Similarities between the kinds of relationships in Figure A-9 and A-10 suggest that the predicted HUC mean median values are reasonable.

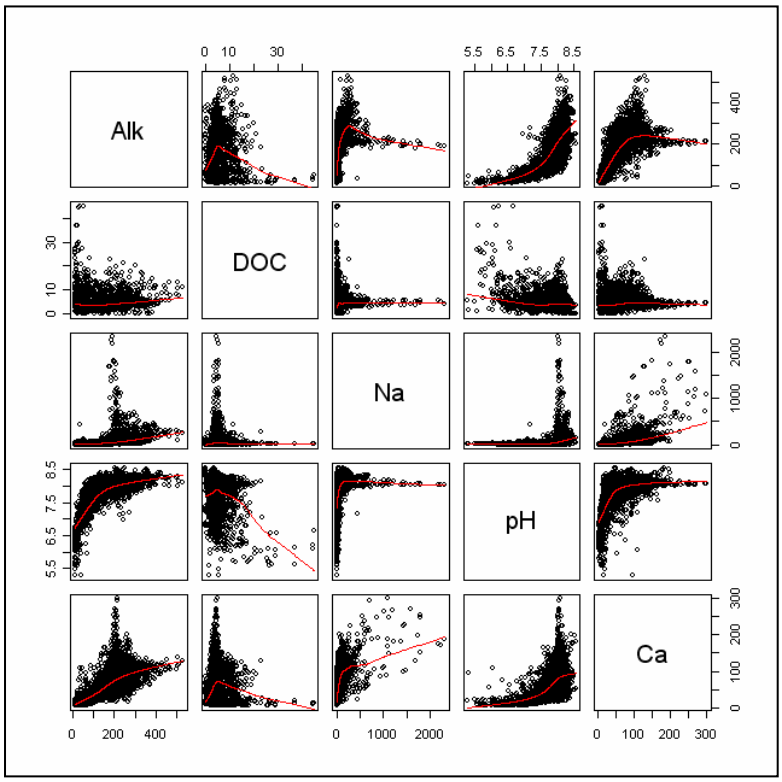

Figure A-9. Scatter plot matrix of median concentration kriged predictions, averaged over 8-digit HUCs regions covering the continental U.S.

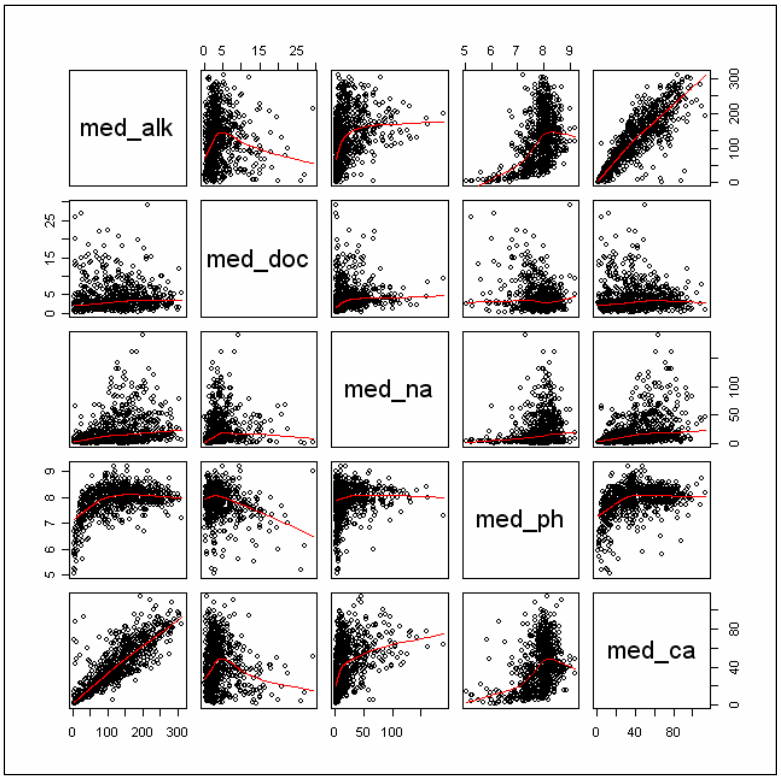

Figure A-10. Scatter plot matrix of median concentrations from 772 monitoring locations in the continental U.S.

In addition to scatter plots, correlation coefficient matrices between constituents in each of the two data sets (HUC-mean kriged median values and site median values for 772 locations) were generated (Table A-1). Although not identical, the coefficients were generally similar between the two datasets, again suggesting that the kriging predictions are reasonable.

Table A-1. Matrices of correlation coefficients between constituent concentrations

2096 HUC-averaged predicted median values					
	Alk	DOC	Na	pH	Ca
Alk	1				
DOC	-0.01456	1			
Na	0.327599	-0.02661	1		
pH	0.761675	-0.27746	0.286512	1	
Ca	0.698379	-0.02585	0.531727	0.58514	1
772 site-median values					
	Alk	DOC	Na	pH	Ca
Alk	1				
DOC	0.019145	1			
Na	0.327028	0.165445	1		
pH	0.453161	-0.24067	0.169238	1	
Ca	0.842484	-0.05097	0.387617	0.374592	1

Principal components analyses (PCA) were also run on both the HUC-averaged predictions and the 772 sets of monitored constituent concentrations to look for linear combinations of variables that might explain most of the observed variation. Figures A-11 and A-12 show the resulting plots of the variance explained by each component, and Table A-2 lists the loadings of the components onto the original variables. The first component comprised 80 and 88% of the variance in the HUC-based and site-based analyses, respectively. As Table A-2 indicates, this component loaded entirely onto alk, Na, and Ca in both cases. For the HUCs, component 1 was primarily loaded on Na, while for the sites, it primarily loaded on alk.

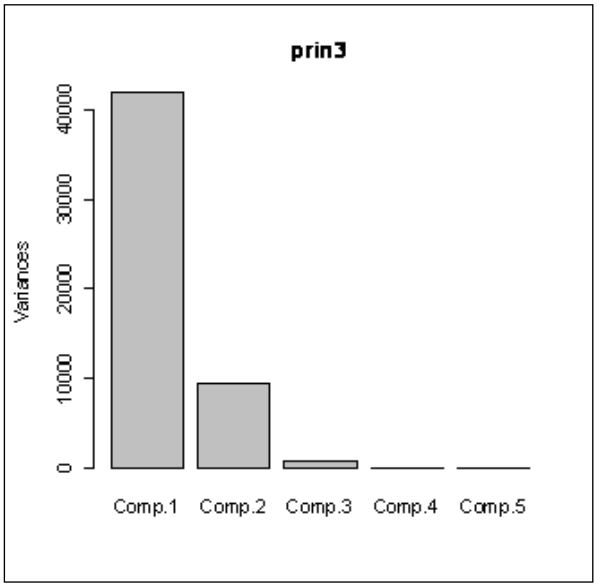

Figure A-11. Variance plot from PCA of HUC-average kriging-predicted concentrations

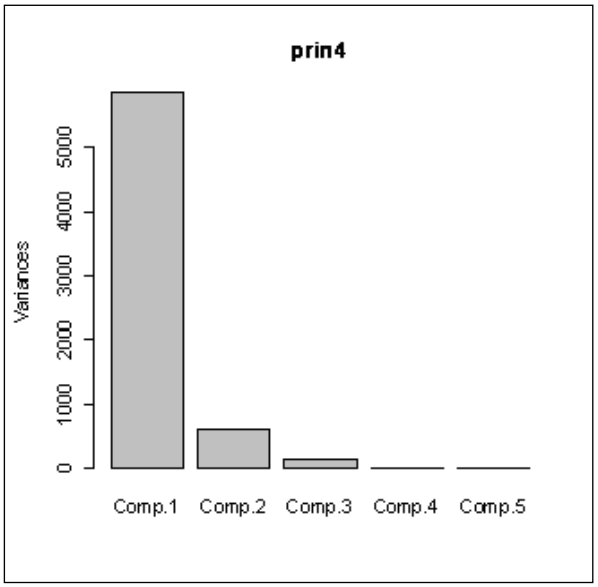

Figure A-12. Variance plot from PCA of site-median measured concentrations

Table A-2. Loadings onto original variables from PCA on HUC-averaged predictions and site-median concentrations

HUCs:					
	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Alk	0.219	0.932	-0.289		
DOC				0.999	
Na	0.965	-0.25			
pH					-0.999
Ca	0.142	0.263	0.954		
Sites:					
	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Alk	0.952	0.162	0.259		
DOC				-0.997	
Na	0.13	-0.982	0.133		
pH					0.999
Ca	0.277		-0.955		

A.5 Developing Regional Defaults

Besides prediction maps of best-estimate median concentrations, the Geostatistical Analyst can be used, with the same sets of kriging parameters listed previously, to generate quantile surface maps that represent reasonably protective inputs to the BLM than standard kriging predicted values. The five key inputs examined in this paper are all positively associated with BLM-predicted LC50s. Thus, lower values of all of them tend to result in lower (i.e., more protective) site-specific criteria. Lower quantile predictions can be used to produce protective regional default inputs. As an example, Figure A-13 displays the 25th percentile prediction map for alkalinity. When these values are block-averaged over the HUC polygons, the resulting alkalinities are lower than 67% of the site-minimum alkalinities (Figure A-14) measured inside the same areas.

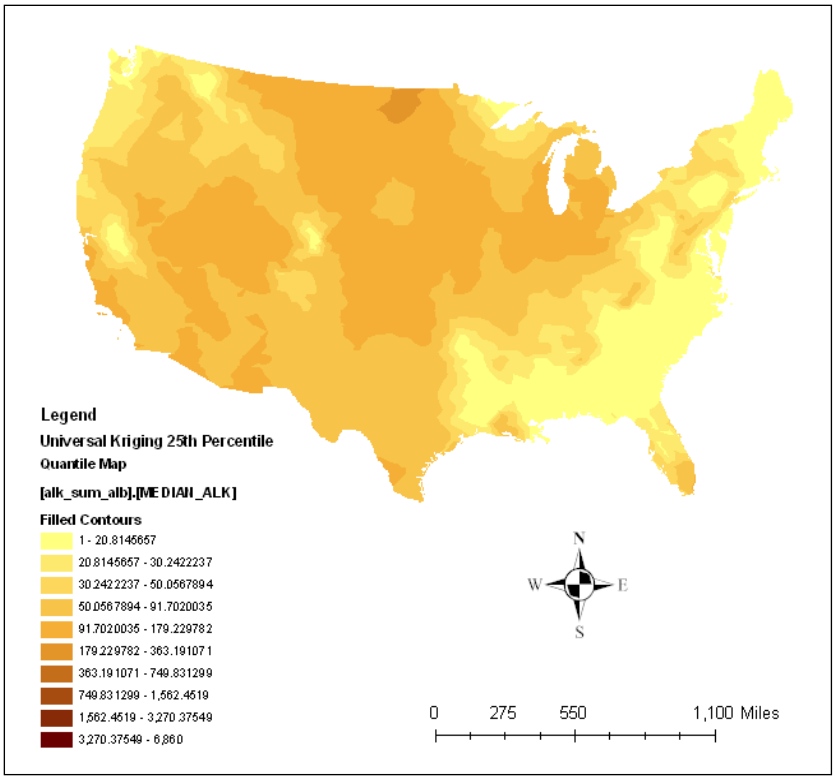

Figure A-13. Kriging 25th percentile map of median alkalinity

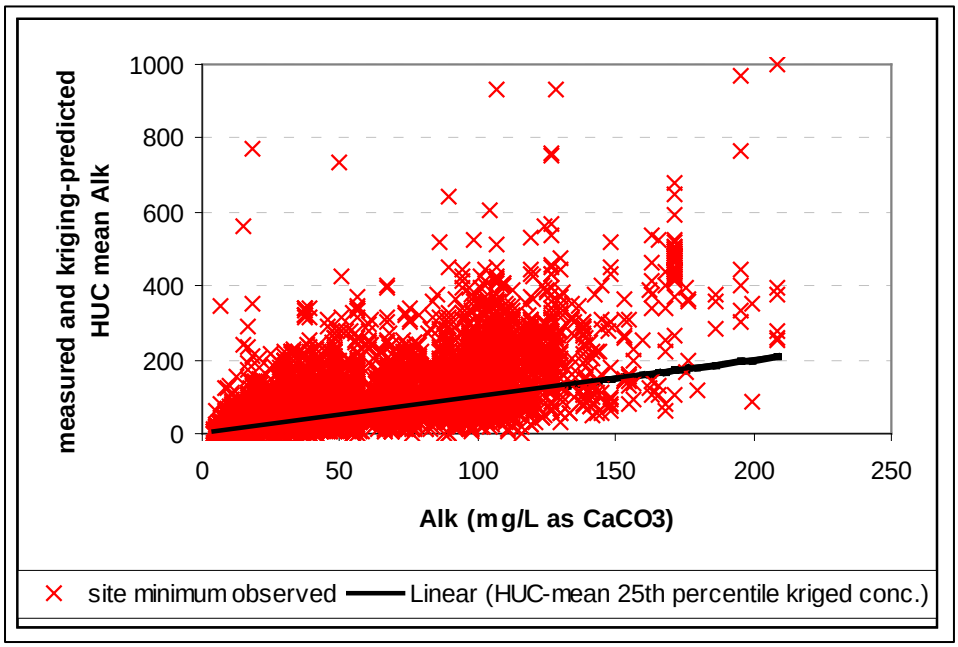

Figure A-14. Comparison of observed site-minimum alkalinities with HUC-mean 25th percentile kriging-predicted values

A.6 Discussion

The use of HUCs for spatial averaging of surface water concentrations is not without conceptual difficulties. First, only about 45% of HUCs are actual watersheds (Omernik, 2003); the rest receive drainage from additional upgradient areas. Concentrations measured in flowing waters reflect the soil, vegetation and land use properties of the aggregate upstream drainage areas, rather than of the sampling locations themselves (Smith et al., 1997). Assignment of measured concentrations to a HUC through block averaging may understate the spatial relevance of the samples for HUCs that are only parts of watersheds. One way to address this concern might be to use, as the aggregation polygons, only samples from watersheds that are entirely contained within single ecoregions (Omernik, 1987). However, this would have the unacceptable consequence of excluding large areas, and perhaps much of the data, from analysis. Another critical problem with this idea is that watershed boundaries for all of the NWIS sampling locations are not readily available, so there is currently no basis for deciding which points should be included or excluded. One advantage provided by the use of HUCs is that they divide the entire land mass of interest in this case into roughly equal sized polygons, at a level of resolution that appears to be roughly compatible with that of observed concentration trends. Block averaging using other sets of similarly sized polygons, such as counties, might serve equally well for empirically capturing broad spatial variability in concentrations. However the resulting concentrations would be less useful because they would lack even the incomplete degree of organization by connected hydrology that HUCs provide.

A.7 Conclusions

Kriging-predicted median concentrations of five water quality constituents, averaged over 8-digit HUCs, showed similar inter-constituent relationships as median concentrations from 772 specific sampling locations. PCA analyses revealed that in both cases, most of the observed variability was related to variations in three of the five constituents: alk, Na, and Ca. Results suggest that block averaging of kriging predictions over irregularly spaced sampling points can provide estimates that preserve much of the interrelationships between different measured entities. The use of suitable low-quantile kriging predictions is suggested as a way to estimate reasonably protective concentrations to serve as regional default inputs to the BLM.

A.8 References

Omernik, J., 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118-125.

Omernik, J., 2003. The misuse of hydrologic unit maps for extrapolation, reporting, and ecosystem management. Journal of the American Water Resources Association 39(3):563-573.

Smith, R.A., Schwartz, G.E., and R.B. Alexander, 1997. Regional interpretation of water-quality monitoring data. Water Resources Research 33912):2781-2798.

U.S. Environmental Protection Agency, 2003. The Biotic Ligand Model: Technical Support Document for its Application to the Evaluation of Water Quality Criteria for Copper, EPA 822-R-03-027.

Appendix B: Approaches for Estimating Missing Biotic Ligand Model Input Parameters. Correlation approaches to estimate Biotic Ligand Model input parameters using conductivity and discharge as explanatory variables

B.1 Introduction

Derivation of water quality criteria for copper and other metals from predictions of bioavailability generated by the Biotic Ligand Model (BLM) introduces a number of issues. For example, obtaining the data needed to apply the BLM may be problematic for many dischargers and receiving waters. The BLM requires 10 input parameters to characterize water quality at a particular site; the most important ones for predicting copper bioavailability and toxicity include pH, dissolved organic carbon (DOC), calcium, magnesium, sodium, alkalinity, and temperature. In stream segments with only small dischargers, or possibly no dischargers at all, the data needed to apply the BLM may not be available. Water quality criteria that rely upon BLM predictions would be greatly facilitated by the development of practical approaches to estimate values for BLM water quality parameters, which could be applied when data for one or more of these parameters are missing at a site.

Given the broad geographical range over which the BLM is likely to be applied, potentially over the entire Nation, and the limited information that is available for many areas, a practical method to estimate missing water quality parameters is needed. The geostatistical methods employed by the U.S. Environmental Protection Agency (EPA) (Carleton, 2006) presented a viable system to estimate missing water quality parameters required by the BLM. The prototype work developed by Carleton applied kriging to predict average concentrations of alkalinity, DOC, sodium, pH and calcium over hydrologic units (8-digit Hydologgic Unit Codes [HUCs]), using the U.S. Geological Service (USGS) National Water Information Service (NWIS) as the source of spatial data. Comparison of measured concentrations with kriging predictions were encouraging for several of the BLM water quality parameters, although the errors and uncertainties associated with these predictions were not fully explored.

The geostatistical approach utilizes knowledge of spatial correlation to project values of a water quality parameter at sites where it has not been measured. The accuracy of these projections depends upon the availability of sufficient and spatially-proximate data for the specific parameter of interest. In addition, the seasonal and annual temporal variation in water quality must also be addressed in order to apply the BLM at a site. Water quality parameters often experience large changes during periods of snowmelt or intense rainfall. In many rivers and streams, the chemical composition and physical properties of water are following trends associated with increased land use in watersheds, water diversion for irrigation, regulation of river flow by dams, and other anthropogenic disturbances.

The acute BLM predicts an *instantaneous* acute copper criterion (i.e., a maximum short-term, non-toxic concentration of copper), which will vary according to changes in the water quality parameters. An appropriately protective copper criterion must therefore reflect the variability of water quality parameters at the site. In previous analyses we found that protective water quality criteria for copper

generally corresponded to approximately the 2.5th percentile of the distribution of instantaneous water quality criteria (IWQC) predicted by the BLM.[5] BLM criteria predictions made for a site using the corresponding percentiles (i.e., 2.5%) of the water quality parameter distributions will be a conservative approximation of this protective criterion. The sensitivity of criteria predictions to the most important BLM water quality inputs is proportional (sensitivity to DOC is ~100%[6], [H^+] is ~50%, calcium, magnesium and sodium is ~20%). Relevant site-specific water quality parameters will be values from the lower "tail" of the measured or estimated distributions.

There may be great value in supplementing the geostatistical approach with classical estimation methods, such as regression and correlation. Examination of the NWIS data used to develop the geostatistical approach suggests that two variables, discharge (flow rate) and conductivity, may be useful for estimating BLM input water quality parameters. The USGS maintains the most comprehensive routine water flow and water quality data for streams and rivers in the Nation. Discharge may be a relevant explanatory variable because the USGS measures or estimates flow on a daily basis for a large number of stream and river segments. Among water quality parameters, the data for conductivity are the most complete and cover the longest time period (Wang and Yin, 1997). The literature also indicates that conductivity is one of the most widely monitored water quality indicators in the U.S. In part, this is because conductivity measurements are usually included in automated multiparameter systems for monitoring changes in the quality of surface waters (Allen and Mancy, 1972).

Conductivity is useful as a general measure of stream water quality. Each stream tends to have a relatively constant range of conductivity that, once established, can be used as a baseline for comparison with regular conductivity measurements (USEPA, 1997). Conductivity in streams and rivers is affected primarily by the geology of the area through which the water flows. Streams that run through areas with granite bedrock tend to have lower conductivity because granite is composed of more inert materials that do not ionize (dissolve into ionic components) when washed into the water. On the other hand, streams that run through areas with clay soils tend to have higher conductivity because of the presence of materials that ionize when washed into the water. Ground water inflows can have the same effects depending on the bedrock they flow through.

Conductivity reflects the strength of major ions in water and is a good estimator of total dissolved solids (TDS). Linear relationships between conductivity and TDS have been developed for many USGS monitoring sites. Conductivity is also linearly related to the sum of cations (McCutcheon et al., 1993). In addition, conductivity measurements provide information about the total concentration of ionic species in a water sample (Tyson, 1988). Figure B-1 illustrates how conductivity relates to hardness and anion concentrations in a river that has a rather saline base flow maintained by irrigation drainage and groundwater inflows. The chemical characteristics of the base flow are generally constant but they are subject to seasonal dilution by runoff. Relationships between conductivity and chloride and sulfate concentrations are well defined. A similarly good association with hardness (calcium+magnesium) is

[5] This was the median for 17 sites; the range was 1 to 36%.

[6] 100% sensitivity implies that a model prediction (in this case, the criteria predicted by the BLM) varies in direct proportion to the change in the value of a specified input parameter.

indicated. Lines drawn by eye through the points for chloride and sulfate show slight curvature, but the departure from linearity is insignificant. It seems evident that a record of conductivity at this station could be used to compute the other chemical characteristics of the water with a good level of accuracy for major ions, except at high flow when the relationships would not be as well defined (Hem, 1985).

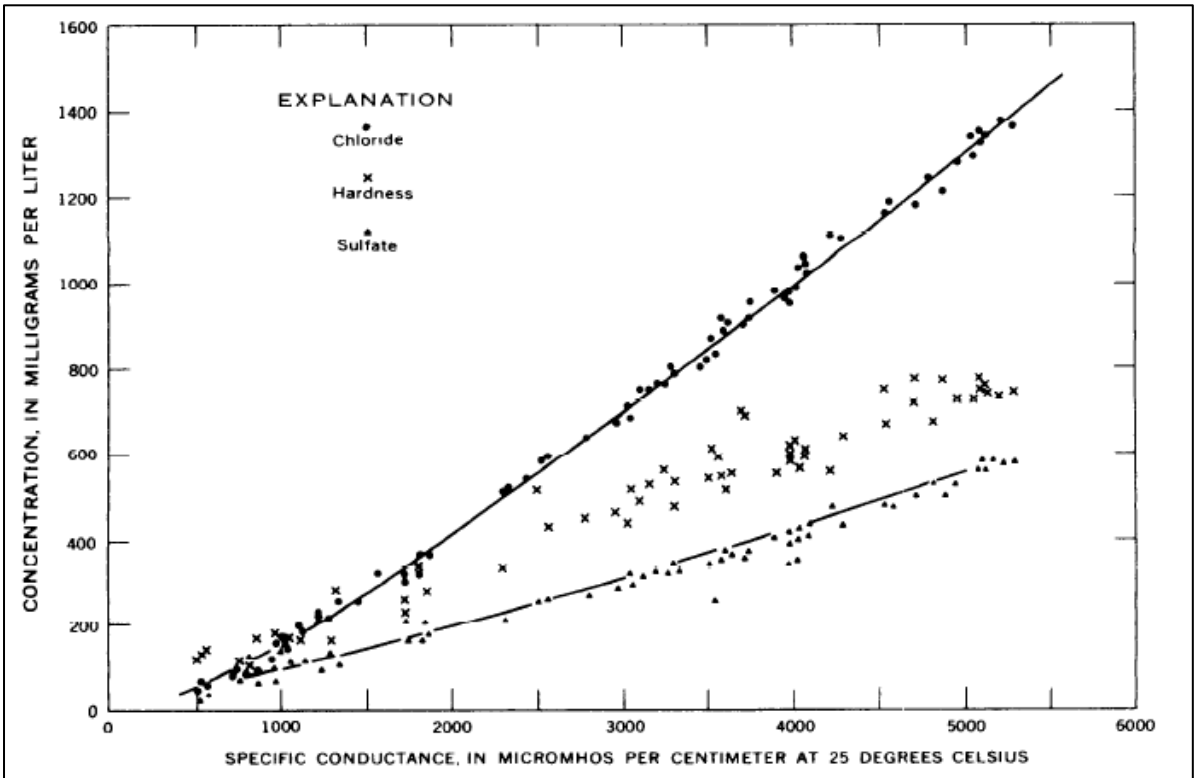

Figure 11. Relation of conductance to chloride, hardness, and sulfate concentrations, Gila River at Bylas, Ariz., October 1, 1943, to September 30, 1944.

Figure B-1. Relation of conductivity to chloride, hardness and sulfate concentrations in the Gila River at Bylas, Arizona

(reprinted from Hem, 1985)

Wang and Yin (1997) established conductivity as a general water quality indicator based on spatial data. The concentration of major base metal cations in water explained the positive correlation between conductivity and hardness. This also explained a rather weak correlation between conductivity and the pH value. Its relationship with other materials, however, most likely resulted from the dilution effect of stream flow. Conductivity was negatively correlated with discharge (ρ=-0.729), and the same was found for most water quality variables that were positively correlated with conductivity. With increasing stream flow, the concentration of the dissolved material decreased, as did the conductivity. Wang and Yin's analysis suggests that conductivity could be used as a general indicator of water quality, which is positively related to dissolved materials and soluble metals. As it is

widely monitored and has relatively long records, conductivity has the potential to be a very useful variable for estimating missing water quality input parameters for the BLM.

We explored this possibility by assessing the degree of correlation between conductivity and each of the BLM water quality parameters. We used NWIS data from three contiguous states in the Western U.S. (Colorado, Utah and Wyoming) for this analysis. These states were selected because of the large spatial and temporal variability observed in BLM water quality parameters, and because they provided us a tractable dataset for analysis.

Discharge was included as one of the parameters in this correlation analysis. However, discharge is most often used to explain water quality variation at a particular site (Hem, 1985). The concentration of dissolved solids in the water of a stream is related to many factors, but it seems obvious that one of the more direct and important factors is the volume of water from rainfall available for dilution and transport of weathering products. Presumably, therefore, the concentrations of dissolved solids should be an inverse function of the rate of discharge of water over all or at least most of the recorded range (Hem, 1985). Regressing water quality parameter measurements against discharge is a common practice in environmental engineering, and many references on this subject are available (McDiffett et al.,1989; Chanat and Hornberger, 2002; Christensen et al., 2005; Godsey and Kirchner, 2005). We should also point out that correlating the variation in water quality parameters to streamflow is also necessary for effluent dilution calculations associated with use of the BLM (for example, the probabilistic dilution framework incorporated in the BLM-Monte software [HydroQual, 2001]).

B.2 Data

Data for discharge, conductivity, and BLM water quality parameters (temperature, pH, DOC, alkalinity calcium, magnesium, sodium, potassium, sulfate, and chloride,) were retrieved from the USGS NWIS web interface (http://nwis.waterdata.usgs.gov/usa/nwis/qwdata). Data were selected for 790 stream and river stations in Colorado, Utah, and Wyoming reporting 100 or more water quality observations. This latter constraint was imposed to eliminate the large number of stations reporting very few (often one) water quality observations. Even when the analysis was restricted to sites with more than 100 water quality observations, there were frequently a marginal number of data for the multiple parameters needed to measure between-parameter correlations. We also restricted the analysis to observations made since 1975 to avoid the possible influence of pre-Clean Water Act discharges on water quality.

Natural logarithms of the discharge data were used in the analysis, because discharge was clearly lognormally distributed at the majority of sites. In cases where a parameter was measured simultaneously by more than one method (field pH versus laboratory pH, for example), the reported results were averaged for analysis. We did not consider other approaches for selecting data based on preference for a particular analytic method (Roberson et al., 1963).

B.3 Results

Table B-1 provides an inventory of the number of observations, and number of sites with data, for several of the parameters in the state of Colorado (these numbers reflect the full NWIS dataset, uncensored for minimum number of observations or date). From this table, it is apparent that a vast amount of conductivity data exists, both in terms of the total number of observations and the number of sites reporting this parameter in comparison to the BLM water quality parameters. For example,

there are almost four times as many observations of conductivity as there are for calcium, and they are measured at more than twice the number of sites. Discharge data is similarly abundant.

Table B-1. Number of observations and sites reported in NWIS for streams and rivers in Colorado

Parameter	Number of Observations	Number of Sites
pH	62,005	3668
alkalinity	8136	839
calcium	45,490	2708
conductivity	168,110	6101
discharge	127,275	3340

To quantify the relationship between conductivity levels and values of water quality parameters required by the BLM, we performed correlation analyses on the NWIS water quality data for the three states. We estimated correlations for several statistics that summarized the distribution of conductivity and water quality values at each station. These included median levels, as well as the first quartile and fifth quantile. The last two statistics represent the lower end of the distribution of parameter values at a site, and are appropriate statistics for calculation of BLM instantaneous criteria. A non-parametric correlation (Spearman's rank correlation) was employed to avoid the problems of unknown data distributions and possibly non-linear relationships. To determine the statistical significance of the rank correlation coefficient (ρ), the significance level (P) was also calculated. The Spearman's rank correlation was also used to examine the relationship between stream discharge and the water quality variables to reveal the effect of dilution.

For the median site concentrations, we found that six BLM water quality parameters, two-thirds of the nine variables examined in this study, had non-zero rank correlation coefficients at the 0.001 significance level (Table B-2). As expected, strong positive correlations between conductivity and salt concentrations were found. For example, the correlation coefficients between conductivity and the concentration of salt cations and anions (sodium, potassium, magnesium, calcium, sulfate and chloride) were all higher than 0.80. However, median site conductivity was not significantly correlated to several other important BLM parameters including pH, DOC, and alkalinity. In terms of the site medians, there appears to be limited correlation between conductivity and the BLM water quality parameters. Furthermore, for the median site concentrations neither conductivity nor any of the BLM water quality parameters were significantly correlated to discharge.

Table B-2. Results of Spearman rank tests for correlation (ρ) between median values of variables at each site.

Probability values (P) are not exact due to the presence of ties in the data

	Conductivity	Discharge
Conductivity		ρ: 0.012 P: 0.892
pH	ρ: 0.175 P: 0.019	ρ: 0.441 P: 0.008
DOC	ρ: 0.866 P: 0.333	ρ: P:
Ca	ρ: 0.867 P: <0.001	ρ: -0.371 P: 0.068
Mg	ρ: 0.882 P: <0.001	ρ: -0.516 P: 0.008
Na	ρ: 0.921 P: <0.001	ρ: 0.139 P: 0.695
K	ρ: 0.846 P: <0.001	ρ: -0.128 P: 0.551
SO_4	ρ: 0.905 P: <0.001	ρ: -0.514 P: 0.010
Alkalinity	ρ: -0.600 P: 0.350	ρ: P:
Cl	ρ: 0.827 P: <0.001	ρ: -0.866 P: 0.333

We then repeated the correlation analysis for the site first quartiles (Table B-3) and fifth quantiles (Table B-4). For both of these low-end distribution statistics, *all* of the BLM water quality parameters were significantly correlated to conductivity, having non-zero rank correlation coefficients at the 0.001 significance level, as listed in Tables B-3 and B-4. The correlation coefficients are lower for pH and DOC than for the salts and alkalinity, but are nevertheless significant. Apparently, the correlation structure between conductivity and the BLM water quality parameters is much stronger at the lower end of the site distributions. Ambiguity in correlations between conductivity and BLM water quality parameters disappears when low-end distribution statistics are analyzed. As was the case for the median site concentrations; neither conductivity nor any of the BLM water quality parameters were correlated with discharge for the low-end distribution statistics.

Table B-3. Results of Spearman rank tests for correlation (ρ) between the first quartile of values at each site.

Probability values (P) are not exact due to the presence of ties in the data

	Conductivity	Discharge
Conductivity		ρ: 0.057 P: 0.144
pH	ρ: 0.287 P: <0.001	ρ: 0.070 P: 0.168
DOC	ρ: 0.618 P: <0.001	ρ: -0.149 P: 0.031
Ca	ρ: 0.920 P: <0.001	ρ: -0.060 P: 0.305
Mg	ρ: 0.935 P: <0.001	ρ: -0.107 P: 0.066
Na	ρ: 0.910 P: <0.001	ρ: -0.075 P: 0.129
K	ρ: 0.773 P: <0.001	ρ: -0.109 P: 0.075
SO_4	ρ: 0.941 P: <0.001	ρ: -0.068 P: 0.247
Alkalinity	ρ: 0.829 P: <0.001	ρ: 0.099 P: 0.381
Cl	ρ: 0.752 P: <0.001	ρ: 0.004 P: 0.958

Table B-4. Results of Spearman rank tests for correlation (ρ) between the fifth quantile of values at each site.

Probability values (P) are not exact due to the presence of ties in the data

	Conductivity	Discharge
Conductivity		ρ: 0.056 P: 0.213
pH	ρ: 0.382 P: <0.001	ρ: 0.032 P: 0.579
DOC	ρ: 0.558 P: <0.001	ρ: -0.107 P: 0.134
Ca	ρ: 0.920 P: <0.001	ρ: 0.017 P: 0.791
Mg	ρ: 0.929 P: <0.001	ρ: -0.056 P: 0.383
Na	ρ: 0.845 P: <0.001	ρ: -0.089 P: 0.078

K	ρ: 0.694 P: <0.001	ρ: -0.065 P: 0.353
SO_4	ρ: 0.908 P: <0.001	ρ: 0.017 P: 0.790
Alkalinity	ρ: 0.784 P: <0.001	ρ: 0.184 P: 0.102
Cl	ρ: 0.706 P: <0.001	ρ: 0.034 P: 0.671

To further illustrate these correlations, scatter plot matrices (or SPLOMs) were prepared for the first quartiles (Figure B-2) and fifth quantiles (Figure B-3). SPLOMs show scatter plots for each combination of parameters, arrayed as a matrix, with parameters labeled along the borders of the plot. Histograms for each parameter are plotted on the main diagonal. The correlations between conductivity and each of the BLM water quality parameters are apparent by examining the second row (from the top) of scatter plots in Figures B-2 and B-3. Likewise, the lack of correlation between these parameters and discharge is apparent in the top row of the scatter plots in these same figures.

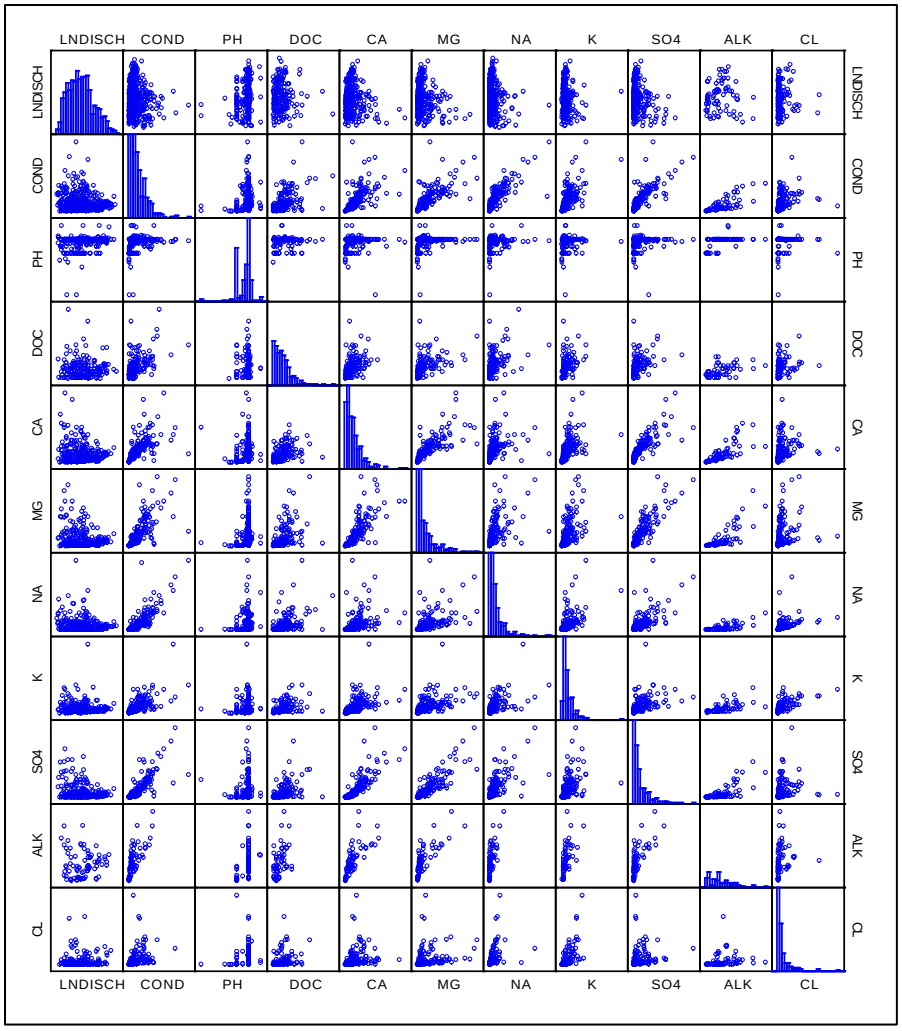

Figure B-2. Scatter plot matrix for first quartile of site-specific data for discharge (LNDISCH), conductivity (COND), and BLM water quality parameters

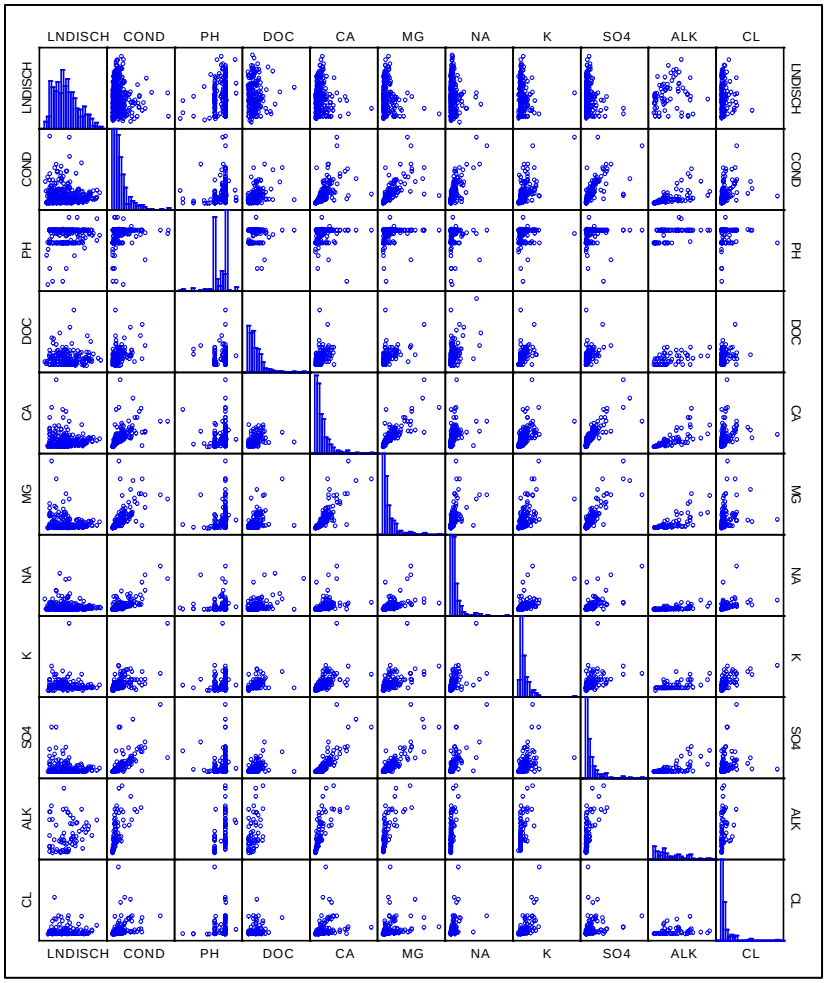

Figure B-3. Scatter plot matrix for fifth quantile of site-specific data for discharge (LNDISCH), conductivity (COND), and BLM water quality parameters

To understand why the correlations between conductivity and the other BLM water quality parameters are so much stronger for the low-end distribution statistics than for the medians, it is necessary to examine the site-specific data itself. Figure B-4 is a SPLOM of the conductivity, discharge, and BLM water quality parameter data for a representative USGS station in Colorado. The histograms for conductivity, salts, and alkalinity are remarkable in that the distribution of each is clearly bimodal (i.e., two separate peaks are evident in the histograms). This was observed for many of the sites in this dataset (not shown). In Figure B-5, the conductivity and discharge data for this site are plotted as a time series, which reveals why the water quality data are bimodal: high values of conductivity (> 5,000 micromhos/cm) occur when streamflow discharge is low, and low values of conductivity (< 2,000 micromhos/cm) occur when the discharge is high. At this station (and many others in this region), streamflow discharge is high in the May-June period coinciding with snowmelt at higher elevations. Feth and others (1964) reported conductivities of melted snow in the Western US ranging from about 2 to 42 picomhos/cm. Thus, the low values of conductivity (as well as concentrations of the salts and alkalinity) are the result of annual dilution from snow melt. At most other times, conductivity and salt and alkalinity concentration values are much higher. Depending upon how the water quality samples

are allocated at a site, the median concentration of these parameters may fall in either mode of the bimodal distribution, resulting in quite different values that appear almost random. Fortunately, the low-end distribution statistics avoid this seeming randomness because they consistently reflect sampling from the lower mode of the concentration distribution.

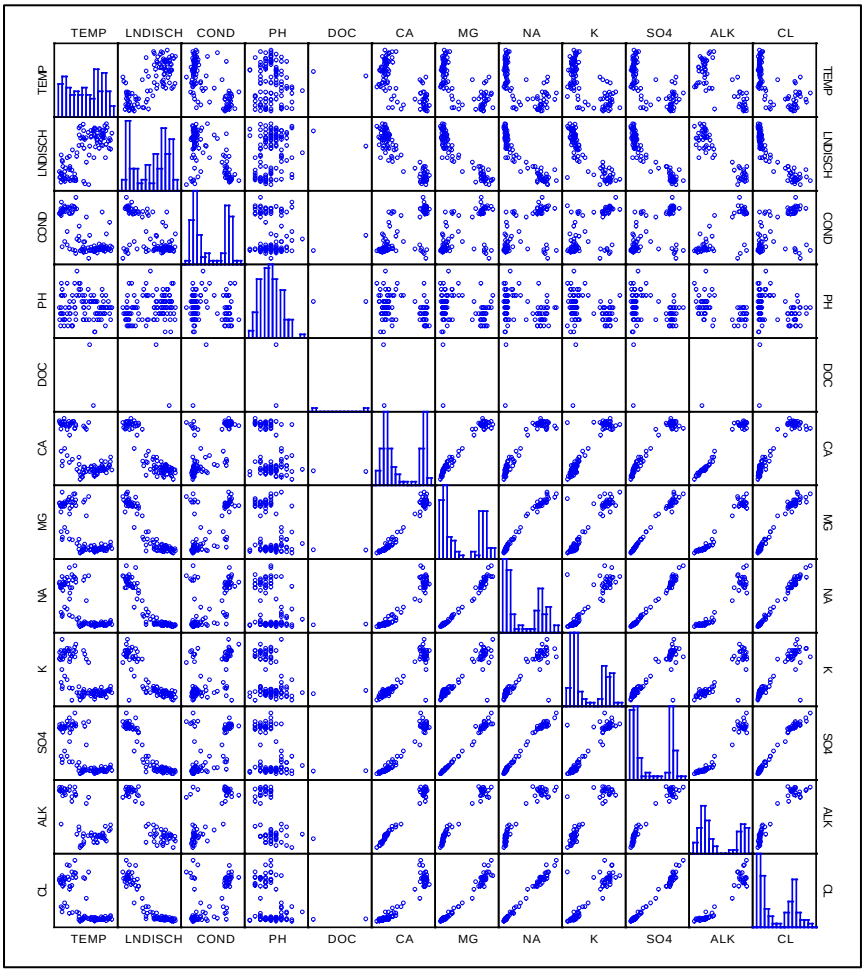

Figure B-4. Scatter plot matrix of BLM water quality parameter data from NWIS Station 384551107591901 (Sunflower Drain at Highway 92, near Read, Delta County, Colorado)

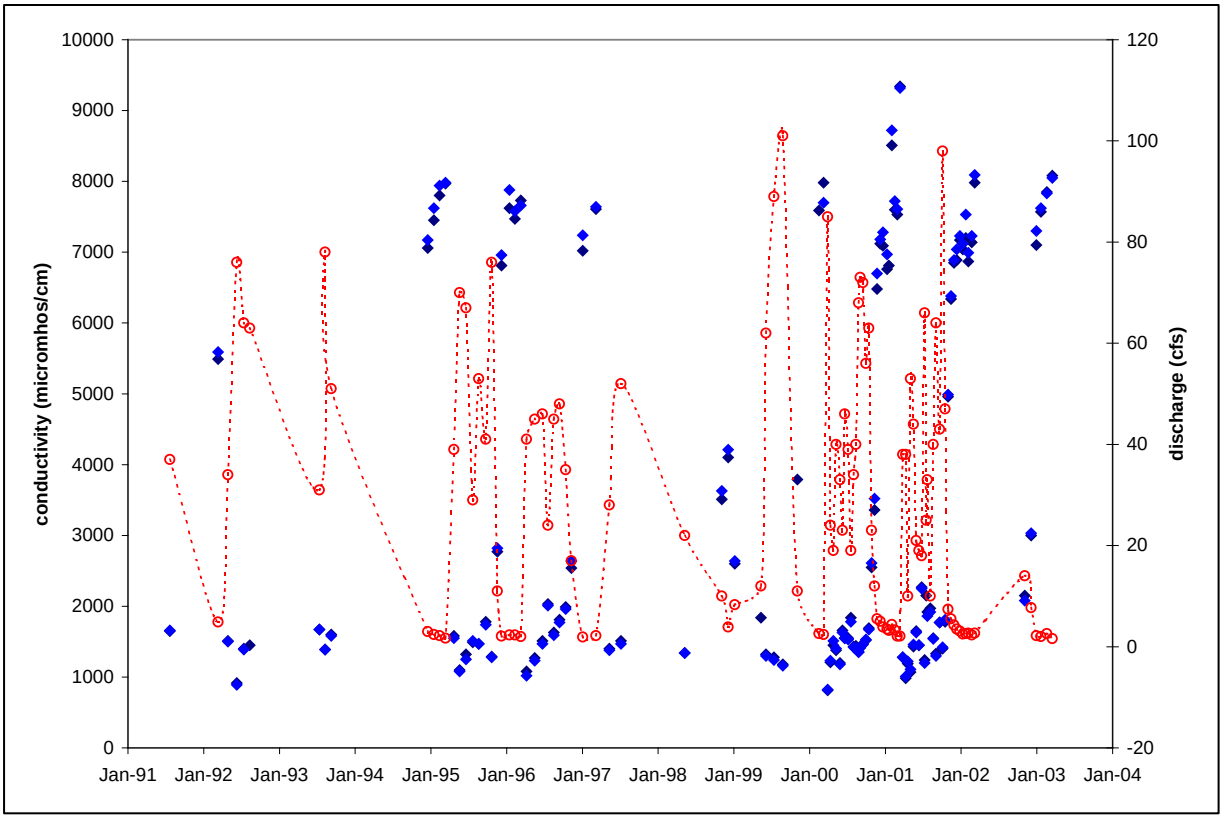

Figure B-5. Time series plot of conductivity (diamond symbols) and discharge (open circles connected by dashed line) at Station 384551107591901 (Sunflower Drain at Highway 92, near Read, Delta County, Colorado)

Figure B-4 also illustrates that, in terms of explaining site-specific variability, discharge is a much better predictive variable for a number of the BLM water quality parameters than conductivity. Each of these parameters (alkalinity calcium, magnesium, sodium, potassium, sulfate, and chloride) is clearly correlated to discharge, but not to conductivity. Discharge correlations are observed at many locations, and are commonly used to project water quality for various applications (Hem, 1985).

B.4 Discussion

Incorporating classical water quality correlation approaches, using conductivity and discharge as explanatory variables, within the geostatistical approach prototyped by EPA, appears promising. Conductivity, but not discharge, is significantly correlated to BLM water quality parameters between sites, especially for the low-end distribution statistics of interest for criteria calculations. Since conductivity data is abundant and it correlates well to BLM water quality parameters, it is reasonable to incorporate conductivity in spatial projections of BLM parameters. This may simplify the geostatistical approach and allow more robust spatial extrapolation of BLM water quality parameters.

Conversely, discharge is correlated to concentrations of a number of BLM parameters (salts and alkalinity) within many sites. Streamflow is a good explanatory variable for a number of the BLM water quality parameters (the salts and alkalinity) because their variabilities largely reflect dilution at high flow rates. Discharge data are also plentiful, so we believe that incorporating classical methods of

correlating concentration to discharge may be a useful means to address within-station variability for BLM water quality parameters.

It should also be recognized that geostatistical and/or correlation approaches appear to most often fail for those water quality parameters which are the most sensitive and important to the BLM, namely DOC and pH. Additional sampling effort will likely be required to address these deficiencies. In the case of pH, it is worth noting that many surface water sampling crews carry electronic multiparameter instruments which measure pH, conductivity, and temperature simultaneously in the field. Therefore, data collection strategies which incorporate these three measurements may be especially effective.

Measurement of DOC is considerably more difficult and expensive. It may be worthwhile to investigate whether ultraviolet (UV) absorption spectroscopy could be used as a surrogate measurement technique for DOC. The organic ligands that bind metals are humic and fulvic compounds (HydroQual, 2005). At least some of these compounds can be measured by UV absorption spectroscopy or related methods (Kalbitz et al., 2000; Wang and Hsieh, 2001), which may be easier and less expensive than DOC analysis.

B.5 References

Allen, H.E. and K.H. Mancy. 1972. Design of measurement systems for water analysis. In: Ciaccio, L.L., ed. Water and water pollution handbook. Marcel Dekker, Inc. New York, N.Y.

Carleton, J.N. 2006. An Examination of Spatial Trends in Surface Water Chemistry in the Continental United States: Implications for the Use of Default Values as Inputs to the Biotic Ligand Model for Prediction of Acute Metal Toxicity to Aquatic Organisms U.S. Environmental Protection Agency, Office of Water, Office of Science & Technology.

Chanat, J.G., Rice, K.C. and G.M. Hornberger. 2002. Consistency of patterns in concentration-discharge plots. Water Resources Research, 38 (8): 22-1.

Christensen, V.G., Jain, X. and A.C. Ziegler. 2005. Regression Analysis and Real-Time Water-Quality Monitoring to Estimate Constituent Concentrations, Loads, and Yields in the Little Arkansas River, South-Central Kansas, 1995-99. U.S. Geological Survey, Water-Resources Investigations Report 00-4126

USEPA. 2002. Development of Methodologies for Incorporating the Copper Biotic Ligand Model (BLM) into Aquatic Life Criteria: Application of BLM to Calculate Site-Specific Fixed Criteria. Prepared by Great Lakes Environmental Center (GLEC) for U.S. Environmental Protection Agency, Office of Science and Technology, Health and Ecological Criteria Division, Contract No.68-C-980134 Work Assignment 3-38. 63 pp.

Feth, J. H., Roberson, C. E., and W.L. Polzer. 1964. Sources of mineral constituents in water from granitic rocks, Sierra Nevada, California and Nevada: U.S. Geological Survey Water-Supply Paper 1535-L 70 p.

Godsey, S.E. and J.W. Kirchner. 2005. Concentration-Discharge Relationships Across Temporal and Spatial Scales, Eos Trans. AGU, 86(52), Fall Meet. Suppl., Abstract H22B-04.

Hem, J.D. 1985. Study and Interpretation of the Chemical Characteristics of Natural Water. U.S. Geological Survey, Water Supply Paper 2254.

HydroQual, Inc. 2001. BLM-Monte User's Guide, Version 2.0. HydroQual, Mahwah, NJ. October, 2001.

HydroQual, Inc. 2005. Biotic Ligand Model Windows Interface, Version 2.1.2, User's Guide and Reference Manual. HydroQual, Mahwah, NJ. June, 2005.

Kalbitz, K., Geyer, S. and W. Geyer. 2000. A comparative characterization of dissolved organic matter by means of original aqueous samples and isolated humic substances. Chemosphere. 40(12):1305-12, June 2000.

McCutcheon, S.C., Martin, J.L. and T.O. Barnwell, Jr. 1993. Water Quality, Section 11, Handbook of Hydrology, David Maidment, Ed., McGraw-Hill, New York.

McDiffett, W.F., Beidler, A.W., Dominick, T.F. and K.D. McCrea. 1989. Nutrient concentration-stream discharge relationships during storm events in a first-order stream. Hydrobiologia. 179 (2) July, 1989.

Roberson, C.E., Feth, J.H., Seaber, P.R., and P. Anderson. 1963. Differences between field and laboratory determinations of pH, alkalinity, and specific conductance of natural water. U.S. Geological Survey Professional Paper 475– C, p. C212–C215.

Tyson, J. 1988. Analysis: What analytical chemists do. London: The Royal Society of Chemistry.

USEPA. 1997. Volunteer Stream Monitoring: A Methods Manual. United States Environmental Protection Agency, Office of Water (4503F), EPA 841-B-97-003, November, 1997.

Wang, G.S. and S.T. Hsieh. 2001. Monitoring natural organic matter in water with scanning spectrophotometer. Environ. Int. 26(4):205-12.

Wang, X. and Z.Y. Yin. 1997. Using GIS to assess the relationship between land use and water quality at a watershed level. Environment International. 23(1): 103-114.

Appendix C: Development of Tools to Estimate Biotic Ligand Model Parameters

C.1 Introduction

The U.S. Environmental Protection Agency (EPA) explored using regression models that project BLM water quality parameters from conductivity data for sites where there may be few or no data available to characterize water. We demonstrated previously (USEPA, 2007) that conductivity (specific conductance) is significantly correlated to Biotic Ligand Model (BLM) water quality parameters between a large number of monitoring sites in three western states (Colorado, Utah, and Wyoming), especially for the low-end distribution statistics of interest for site-specific fixed water quality criteria calculations. Since conductivity data are also abundant, it is reasonable to incorporate conductivity in spatial projections of BLM parameters.

C.2 Regression Analysis

Water quality data were retrieved from the U.S. Geological Survey (USGS) National Water Information System (NWIS; http://waterdata.usgs.gov/nwis/qw). We focused our efforts on data collected from rivers and streams in the western states of Colorado, Utah, and Wyoming between 1984 and 2005. Data from these three states were selected because conductivity was known to vary substantially, and the legacy of past mining in the region made the contamination of waterbodies by trace metals a possibility. Data collected prior to 1984 was excluded because a number of the analytical methods used by USGS prior to that date have been replaced by methods with improved precision and lower detection limits. Furthermore, only sites with 40 or more samples were included in the analysis. Data were retrieved for all BLM water quality input parameters including pH, dissolved organic carbon (DOC) (or total organic carbon (TOC), if no DOC data were available) and the geochemical ions (GIs). We also retrieved discharge measurements and filtered (dissolved) copper concentration data, although these data were not included in the regression analysis.

In work described in Appendix B, we found that the correlation structure between conductivity and the BLM water quality parameters was much stronger at the lower end of the concentration distributions. For various low-end distribution statistics, all of the BLM water quality parameters were significantly correlated to conductivity, having non-zero rank correlation coefficients at the 0.001 significance level. The correlation coefficients for pH and DOC were lower than for the GIs, but were nevertheless significant. We exploited this feature of the data in our current work. For each site, we estimated the 10^{th} percentile (i.e., the value exceeded by 90% of the data) of conductivities and the 10^{th} percentile of BLM water quality parameter values. We then fit regression models to project 10^{th} percentiles of BLM parameter values as a function of 10^{th} percentiles of conductivities.

We also fit regression models to the full NWIS dataset (data for all rivers and streams sampled in Colorado, Utah, and Wyoming between 1984 and 2005). This was done out of concern that the lower percentile data might be skewed due to sampling bias, censoring, fewer sites, etc. The results of both approaches are presented below.

C.2.1 pH

The following regression model appeared to be optimum for projecting the 10^{th} percentile of pH from the 10^{th} percentile of conductivity at the sites for which appropriate data were available:

$$\ln(pH) = 1.85 + 0.0352 \cdot \ln(EC)$$

We did not fit a regression model to the full NWIS dataset, because no trend was evident between conductivity and pH.

C.2.2 DOC

The following regression model appeared to be optimum for projecting the 10th percentile of DOC concentrations from the 10th percentile of conductivity at the sites for which appropriate data were available:

$$\ln(DOC) = 0.671 \cdot \ln(EC) - 1.60$$

As with pH, we did not fit a regression model to the full NWIS dataset, since no trend was evident between conductivity and DOC.

C.2.3 Alkalinity

The following regression model appeared to be optimum for projecting the 10th percentile of alkalinity concentrations from the 10th percentile of conductivity at the sites for which appropriate data were available:

$$\ln(alkalinity) = 1.14 \cdot \ln(EC) - 4.68$$

For the full NWIS dataset, the following regression model was developed to project alkalinity concentrations from conductivity:

$$\ln(alkalinity) = 0.652 \cdot \ln(EC) + 0.530$$

C.2.4 Calcium

The following regression model appeared to be optimum for projecting the 10th percentile of calcium concentrations from the 10th percentile of conductivity at the sites for which appropriate data were available:

$$\ln(Ca) = 1.14 \cdot \ln(EC) - 4.35$$

For the full NWIS dataset, the following regression model was developed to project calcium concentrations from conductivity:

$$\ln(Ca) = 0.866 \cdot \ln(EC) - 1.51$$

C.2.5 Magnesium

The following regression model appeared to be optimum for projecting the 10th percentile of magnesium concentrations from the 10th percentile of conductivity at the sites for which appropriate data were available:

$$\ln(Mg) = 1.27 \cdot \ln(EC) - 4.81$$

For the full NWIS dataset, the following regression model was developed to project magnesium concentrations from conductivity:

$$\ln(Mg) = 0.986 \cdot \ln(EC) - 3.48$$

C.2.6 Sodium

The following regression model appeared to be optimum for projecting the 10th percentile of sodium concentrations from the 10th percentile of conductivity at the sites for which appropriate data were available:

$$\ln(Na) = 0.578 \cdot \ln(EC) - 2.62$$

For the full NWIS dataset, the following regression model was developed to project sodium concentrations from conductivity:

$$\ln(Na) = 1.32 \cdot \ln(EC) - 4.96$$

C.2.7 Potassium

The following regression model appeared to be optimum for projecting the 10th percentile of potassium concentrations from the 10th percentile of conductivity at the sites for which appropriate data were available:

$$\ln(K) = 0.882 \cdot \ln(EC) - 3.29$$

For the full NWIS dataset, the following regression model was developed to project potassium concentrations from conductivity:

$$\ln(K) = 0.647 \cdot \ln(EC) - 3.04$$

C.2.8 Sulfate

The following regression model appeared to be optimum for projecting the 10th percentile of sulfate concentrations from the 10th percentile of conductivity at the sites for which appropriate data were available:

$$\ln(SO_4) = 1.16 \cdot \ln(EC) - 4.85$$

For the full NWIS dataset, the following regression model was developed to project sulfate concentrations from conductivity:

$$\ln(SO_4) = 1.43 \cdot \ln(EC) - 4.47$$

C.2.9 Chloride

For the full NWIS dataset, the following regression model was developed to project chloride concentrations from conductivity:

$$\ln(\text{chloride}) = 1.39 \cdot \ln(EC) - 6.15$$

Unfortunately, there were an insufficient number of sites reporting chloride data for a regression model to be developed for the 10th percentile of chloride.

C.3 Application of Conductivity Regressions

There are a number of ways in which the conductivity regressions could be used to project BLM water quality inputs. However, the most important situation may be when a fixed copper criteria value must be calculated for a site where there may be little data available to characterize water quality. In such cases, the regressions allow some or all of the BLM water quality inputs to be projected from either (1)

a limited number of conductivity measurements or (2) a low-end conductivity value estimated by geostatistical or other methods. The first approach, projecting BLM water quality inputs from conductivity measurements, will be demonstrated in this section for a limited number of test sites. The second approach, projecting the BLM water quality inputs based on conductivities estimated by geostatistical methods, is demonstrated in the following section (Section C4).

The regression models presented above for projecting BLM water quality inputs from conductivity were tested using data and BLM predictions from a number of sites. For each site, a fixed copper criteria value was calculated using the Monte Carlo method described in EPA 2002. The BLM version 2.2.3 was used for all BLM calculations. Fixed copper criteria values were determined by the Monte Carlo method, utilizing site-specific data for parameter distributions and variance-covariance structure of all BLM water quality parameter inputs as well as filtered copper concentrations. The test sites (below) were selected on the basis of convenience, number of water quality observations, and geographic location.

The BLM water quality inputs projected from the conductivity regressions are low-end percentiles appropriate for predicting the instantaneous criterion (IC) predicted by the BLM to estimate the fixed site criteria (FSC) value. We suggested this approximation previously, based on the observation that protective FSC for copper generally corresponded to approximately the 2.5th percentile of the distribution of IC predicted by the BLM. BLM estimates made for a site using the corresponding percentiles of the water quality parameter distributions will be a conservative approximation of this protective criteria values. For the present work, we are using this approach to test the 10th percentile water quality parameter values projected from the conductivity regressions.

Previously we noted that filtered copper concentrations were correlated to BLM input water quality parameters at many sites. Furthermore, we found that the degree of correlation between copper concentrations and BLM input parameters appeared to be an important site-specific factor in determining the relationship between the FSC and the IC. Copper concentrations are not required to run the BLM in its toxicity prediction mode, but they are used in the Monte Carlo method to determine the FSC. Because of this, we calculated the FSC both with and without (neglecting) the correlation between copper concentrations and BLM input parameters at each test site.

C.3.1 Naugatuck River, Connecticut

The USGS has sampled the Naugatuck River near Waterville, Connecticut (Station 01208049) since 1967. Ninety-one water samples collected since 1984 provided near-concurrent measurements of all BLM water quality inputs and filtered copper concentrations. The water is low in hardness and alkalinity, slightly acidic (mean pH = 7.32), and fairly low in conductivity (10th percentile = 134 µS/cm at 25º C). Organic carbon concentrations are representative for rivers and streams in this region and nationwide (logmean TOC = 4.02 mg/L), and the filtered copper concentrations are low (logmean filtered copper = 3.62 µg/L). The FSC for copper at this site was calculated to be 11.4 µg/L when the correlation between copper concentrations and BLM parameters was considered, and 7.0 µg/L when this correlation was neglected. Test results at this site are show in Table C-1.

Table C-1. Copper Fixed Site Criterion predictions for the Naugatuck River, Connecticut using various calculation methods

Calculation Method	DOC (mg/L)	pH	Geochemical Ions	Copper Fixed Site Criterion (µg/L)
Monte Carlo FSC with [copper] correlated to inputs (r =0.7)	Data	Data	Data	11.4
Monte Carlo FSC with no [copper] correlation	Data	Data	Data	7.0
IC calculated with 10th % of input data	2.8 (10th % of data)	7.1 (10th % of data)	(10th % of data)	6.4
IC calculated with input from 10th % of conductivity and correlations	5.49 (projected from correlations)	7.55 (projected from correlations)	(projected from correlations)	21.5
IC calculated with input from 10th % of conductivity and correlations except DOC	2.7 (10th % from L3 ecoregion)	7.55 (projected from correlations)	(projected from correlations)	10.5
IC calculated with input from 10th % of conductivity and correlations except pH & DOC	2.7 (10th % from L3 ecoregion)	7.1 (10th % of data)	(projected from correlations)	5.7
IC calculated with input from 10th % of conductivity and correlations except pH & DOC	2.8 (10th % of data)	7.1 (10th % of data)	(projected from correlations)	5.9

The BLM was then applied to predict IC for copper, using low-end percentiles of the measured BLM water quality inputs. Using the 10th percentile values of all measured input data, an IC of 6.4 µg/L was predicted. This IC is 43% smaller than the FSC calculated considering the copper correlation, but only 21% smaller than the FSC neglecting this correlation. When the 10th percentiles of all of the BLM water quality inputs were instead projected from conductivity using the regression models, the predicted ICs were 21.7 µg/L (using the regressions based on 10th percentiles of the three-state data) and 21.5 µg/L (using the regressions based on all of the data). These results illustrate two important points. First, the BLM predictions based on water quality inputs all projected from conductivity correlations are quite different from BLM predictions based on site data; this will be further considered below. Secondly, however, the BLM predictions based on projected water quality inputs do not really depend on which correlations are used.

Clearly, this result shows that the regression models are unable to accurately project all BLM water quality inputs at this site. However, this was almost entirely due to inaccuracy in the pH and organic carbon projections. To demonstrate this, we recalculated the IC several times, using better estimates of the organic carbon and/or pH data, but all other BLM water quality inputs projected from conductivity using the regression models. The first recalculation was made using the 10th percentile of DOC from rivers and streams in the Northeastern Coastal Zone, the Level III ecoregion where the

Naugatuck River is located.[7] In this case the predicted IC was 10.5 µg/L, a value much closer to the FSCs as well as the IC calculated using the 10th percentile values of all the measured input data. A second recalculation was made using the 10th percentile of the pH data, together with the ecoregional 10th percentile of DOC and all other BLM water quality inputs projected from conductivity using the regression models. Finally, a third recalculation was made in which the 10th percentiles of both pH and DOC data were input, with the remaining BLM water quality inputs projected from conductivity using the regression models. For both of these cases, the ICs were within about 10% of the prediction made using the 10th percentile values of all the measured input data. In summary, if BLM predictions are made for copper IC using measured values of pH and organic carbon, minimal error results from projecting the other BLM water quality inputs using conductivity and the regression models. As will be shown in the following sections, the same result was found for the other test sites.

The correlation between filtered copper concentrations and BLM parameters and output was quite strong at this location (r = 0.70 between filtered copper and IC predictions). As a result, the FSC corresponds to an elevated percentile (40%) of the IC predictions. If this correlation is neglected in the Monte Carlo method, the FSC corresponds to only the 14th percentile of the IC predictions. This suggests that the relationship between FSC and IC (in terms of the percentile of the IC distribution corresponding to the FSC) may be somewhat site-specific. Regardless of this complication, the conductivity regressions appear to project reliable low-end percentile estimates of the BLM water quality inputs other than pH and organic carbon. This was demonstrated by repeating the analysis described above using 5th, 2.5th, and 1st percentile input values and projections, each of which produced comparable results (not shown).

C.3.2 San Joaquin River, California

The USGS has sampled the San Joaquin River near Vernalis, California (Station 1130500) since 1950. Water samples collected since 1984 provided 283 near-concurrent measurements of all BLM water quality inputs and 77 filtered copper concentrations. The water has moderate values of hardness and alkalinity, neutral pH, and moderately high conductivity (10th percentile = 307 µS/cm at 25º C). DOC concentrations are representative for rivers and streams in this region and nationwide (logmean DOC = 5.35 mg/L), and the filtered copper concentrations are low (logmean filtered copper = 1.75 µg/L). The FSC for copper at this site was calculated to be 39.1 µg/L, and the correlation between copper concentrations and BLM parameters was strong (r = 0.624 between filtered copper and IC predictions). This FSC value corresponds to the 46th percentile of the distribution of IC. When the FSC for copper was recalculated assuming no correlation between copper concentrations and BLM parameters, the value decreased to 11.1 µg/L (corresponding to the 4.5th percentile of the IC distribution). Test results at this site are tabulated Table C-2.

[7] Ecoregion and water body-type specific DOC concentration percentiles were tabulated for the *Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health (2000), Technical Support Document Volume 2: Development of National Bioaccumulation Factors* (EPA-822-R-03-030).

Table C-2. Copper Fixed Site Criterion predictions for the San Joaquín River, California using various calculation methods

Calculation Method	DOC (mg/L)	pH	Geochemical Ions	Fixed Site Criterion (µg/L)
Monte Carlo FSC with [copper] correlated to inputs (r =0.6)	Data	Data	Data	39.1
Monte Carlo FSC with no [copper] correlation	Data	Data	Data	11.1
IC calculated with 10th % of input data	2.7 (10th % of data)	7.5 (10th % of data)	(10th % of data)	11.9
IC calculated with input from 10th % of conductivity and correlations	9.38 (projected from correlations)	7.77 (projected from correlations)	(projected from correlations)	54.0
IC calculated with input from 10th % of conductivity and correlations except DOC	2.79 (10th % from L3 ecoregion)	7.77 (projected from correlations)	(projected from correlations)	16.0
IC calculated with input from 10th % of conductivity and correlations except pH & DOC	2.79 (10th % from L3 ecoregion)	7.5 (10th % of data)	(projected from correlations)	11.6
IC calculated with input from 10th % of conductivity and correlations except pH & DOC	2.7 (10th % of data)	7.5 (10th % of data)	(projected from correlations)	11.2

The BLM was then applied to predict IC for copper, using low-end percentiles of the BLM water quality inputs. Using the 10th percentile values of all measured input data, an IC of 11.9 µg/L was predicted, which is 70% smaller than the FSC calculated considering the copper concentration correlation but 7% higher than the FSC neglecting this correlation. When the 10th percentiles of all of the BLM water quality inputs were instead projected from conductivity using the regression models, the predicted IC were 50.0 µg/L (using the regressions based on 10th percentiles of the three-state data) and 54.0 µg/L (using the regressions based on all of the data). Again, the BLM predictions based on projected water quality inputs do not really depend on which correlations are used. And, as was the case at the Naugatuck River site, the regression models were unable to accurately project all BLM water quality inputs at this site, although the error is again almost entirely due to inaccuracy in the pH and organic carbon projections. As in the previous case, we demonstrated this by recalculating the IC several times, using better estimates of the organic carbon and/or pH data, but all other BLM water quality inputs projected from conductivity using the regression models. The first recalculation was made using the 10th percentile of DOC from rivers and streams in the Central California Valley, the Level III ecoregion where the San Joaquin River is located. In this case the predicted IC was 16.0 µg/L, a value much closer to the uncorrelated FSCs as well as the IC calculated using the 10th percentile values of all the measured input data. A second recalculation was made using the 10th percentile of the pH data, together with the ecoregional 10th percentile of DOC and all other BLM water quality inputs projected from conductivity using the regression models. Finally, a third recalculation was made in which the 10th percentiles of both pH and DOC data were input, with the remaining BLM water quality inputs projected from conductivity using the regression models. For both of these cases, the ICs were within

about 5% of the prediction made using the 10th percentile values of all the measured input data. BLM predictions made for copper IC at this site using measured values of pH and organic carbon, but all other BLM water quality inputs projected using conductivity regressions, were found to be accurate in comparison to model predictions made using all measured input data.

C.3.3 South Platte River, Colorado

The South Platte River has been sampled by the USGS at Denver, Colorado (Station 06714000) since 1972. Water samples collected since 1984 provided 93 near-concurrent measurements of all BLM water quality inputs and 10 filtered copper concentrations. The water is moderately high in hardness and alkalinity, neutral pH, and moderate conductivity (10th percentile = 229 µS/cm at 25° C). Organic carbon concentrations are representative for rivers and streams in this region (logmean DOC = 5.50 mg/L), and the filtered copper concentrations are low (logmean filtered copper = 3.27 µg/L). The FSC for copper at this site was calculated to be 35.4 µg/L. This FSC value corresponds to the 32nd percentile of the distribution of IC. Moderate correlation between copper concentrations and BLM parameters was observed at this site (r = 0.50 between filtered copper and IC predictions). When the FSC for copper was recalculated assuming no correlation between copper concentrations and BLM parameters, the value decreased to 20 µg/L (corresponding to the 4.3rd percentile of the IC distribution). Test results at this site are shown in Table C-3.

Table C-3. Copper Fixed Site Criterion predictions for the South Platte River, Colorado using various calculation methods

Calculation Method	DOC (mg/L)	pH	Geochemical Ions	Copper Fixed Site Criterion (µg/L)
Monte Carlo FSC with [copper] correlated to inputs (r =0.5)	Data	Data	Data	35.4
Monte Carlo FSC with no [copper] correlation	Data	Data	Data	20.0
IC calculated with 10th % of input data	4.1 (10th % of data)	7.5 (10th % of data)	(10th % of data)	17.3
IC calculated with input from 10th % of conductivity and correlations	7.7 (projected from correlations)	7.7 (projected from correlations)	(projected from correlations)	37.5
IC calculated with input from 10th % of conductivity and correlations except DOC	4.5 (10th % from L3 ecoregion)	7.7 (projected from correlations)	(projected from correlations)	21.6
IC calculated with input from 10th % of conductivity and correlations except pH & DOC	4.5 (10th % from L3 ecoregion)	7.5 (10th % of data)	(projected from correlations)	17.3
IC calculated with input from 10th % of conductivity and correlations except pH & DOC	4.1 (10th % of data)	7.5 (10th % of data)	(projected from correlations)	15.9

The BLM was applied to predict IC for copper, using low-end percentiles of the BLM water quality inputs. Using the 10th percentile values of all measured input data, an IC of 17.3 µg/L was predicted, which is 51% smaller than the FSC calculated considering the copper concentration correlation but only 14% smaller than the FSC neglecting this correlation. When the 10th percentiles of all of the BLM water quality inputs were instead projected from conductivity using the regression models, the predicted IC was 37.5 µg/L. As was the case at the previous sites, the regression models were again unable to accurately project pH and organic carbon concentrations for input to the BLM. We demonstrated this by recalculating the IC several times, using better estimates of the organic carbon and/or pH data, but all other BLM water quality inputs projected from conductivity using the regression models. The first recalculation was made using the 10th percentile of DOC from rivers and streams in the Western High Plains, the Level III ecoregion where the South Platte River is located. In this case the predicted IC was 21.6 µg/L, a value much closer to the uncorrelated FSCs as well as the IC calculated using the 10th percentile values of all the measured input data. A second recalculation was made using the 10th percentile of the pH data, together with the ecoregional 10th percentile of DOC and all other BLM water quality inputs projected from conductivity using the regression models. Finally, a third recalculation was made in which the 10th percentiles of both pH and DOC data were input, with the remaining BLM water quality inputs projected from conductivity using the regression models. For both of these cases, the ICs were within 10% of the prediction made using the 10th percentile values of all the measured input data. As with the previous cases, the BLM predictions made for copper IC at this site using measured values of pH and organic carbon, but where all other BLM water quality inputs were projected using conductivity regressions, were found to be accurate in comparison to model predictions made using all measured input data.

C.3.4 Halfmoon Creek, Colorado

The USGS has sampled Halfmoon Creek near Malta, Colorado (Station 07083000) since 1959. Seventy-three water samples collected since 1984 provided near-concurrent measurements of all BLM water quality inputs and 18 filtered copper concentrations. The water is very low in hardness and alkalinity, slightly acidic (mean pH = 7.76), and low in conductivity (10th percentile = 50.1 µS/cm at 25° C). Organic carbon concentrations are low (logmean DOC = 0.92 mg/L), as are the filtered copper concentrations (logmean filtered copper = 1.75 µg/L). The FSC for copper at this site was calculated to be 1.56 µg/L, corresponding to the 6th percentile of the distribution of IC. The correlation between copper concentrations and BLM parameters was negligible at this site, so the Monte Carlo FSC were not calculated twice (i.e., with and without the copper correlation) as was done at the other sites. Test results at this site are shown in Table C-4.

Table C-4. Copper Fixed Site Criterion predictions for the Halfmoon Creek, Colorado using various calculation methods

Calculation Method	DOC (mg/L)	pH	Geochemical Ions	Copper Fixed Site Criterion (µg/L)
Monte Carlo FSC with [copper] correlated to inputs (r =0.01)	Data	Data	Data	1.56
IC calculated with 10th % of input data	0.6 (10th % of data)	7.2 (10th % of data)	(10th % of data)	1.42
IC calculated with input from 10th % of conductivity and correlations	2.8 (projected from correlations)	7.3 (projected from correlations)	(projected from correlations)	7.43
IC calculated with input from 10th % of conductivity and correlations except DOC	0.6 (10th % from L3 ecoregion)	7.3 (projected from correlations)	(projected from correlations)	1.58
IC calculated with input from 10th % of conductivity and correlations except pH & DOC	0.6 (10th % from L3 ecoregion)	7.2 (10th % of data)	(projected from correlations)	1.39
IC calculated with input from 10th % of conductivity and correlations except pH & DOC	0.6 (10th % of data)	7.2 (10th % of data)	(projected from correlations)	1.39

The BLM was then applied to predict IC for copper, using low-end percentiles of the BLM water quality inputs. Using the 10th percentile values of all measured input data, an IC of 1.42 µg/L was predicted, only 9% smaller than the FSC. When the 10th percentiles of all of the BLM water quality inputs were instead projected from conductivity using the regression models, the predicted IC was 7.43 µg/L. Again, this result clearly shows that the regression models are unable to accurately project all BLM water quality inputs at this site. As in the previous examples, this was almost entirely due to inaccuracy in the pH and organic carbon projections. As in the previous cases, we demonstrated this by recalculating the IC several times, using better estimates of the organic carbon and/or pH data, but all other BLM water quality inputs projected from conductivity using the regression models. The first recalculation was made using the 10th percentile of DOC from rivers and streams in the Southern Rockies, the Level III ecoregion where Halfmoon Creek is located. In this case the predicted IC was 1.58 µg/L, a value within about 10% of the FSC as well as the IC calculated using the 10th percentile values of all the measured input data. A second recalculation was made using the 10th percentile of the pH data, together with the ecoregional 10th percentile of DOC and all other BLM water quality inputs projected from conductivity using the regression models. Finally, a third recalculation was made in which the 10th percentiles of both pH and DOC data were input, with the remaining BLM water quality inputs projected from conductivity using the regression models. For both of these cases, the ICs were within about 2% of the prediction made using the 10th percentile values of all the measured input data. As in the previous examples, if BLM predictions are made for copper IC using measured values of pH and organic carbon, minimal error results from projecting the other BLM water quality inputs using conductivity and the regression models.

C.3.5 Summary of Site-Specific Test Results

The results of this work can be summarized as follows:

Regression models were developed to project 10th percentiles of BLM water quality parameters from the 10th percentile of conductivity distributions at sites in Colorado, Utah, and Wyoming. The regression models were tested using data and copper BLM predictions for four sites, and produced highly consistent results. The regression models for pH and DOC, the most sensitive of BLM water quality parameters, were not sufficiently accurate to make reliable BLM predictions. However, regression models for the GI parameters (alkalinity, calcium, magnesium, sodium, potassium, sulfate, and chloride,) were reasonably accurate, as judged by comparison of model predictions made using projected values of the GI BLM input parameters to model predictions made using all measured input data. The regression models used to project GI parameters from conductivity were calculated two different ways; however, the BLM predictions of IC were not sensitive to this difference.

We were unable to find an estimate for site-specific pH that was superior to the (admittedly poor) conductivity regression. To improve upon this estimate it was necessary to use actual site-specific pH data. This appears to be the general case for reliable site-specific BLM application.

For DOC, the ecoregion and water body-type specific DOC concentration percentiles tabulated by EPA for the National Bioaccumulation Factors Technical Support Document appear to be far better estimates of lower-percentile DOC concentrations than the projections made using the conductivity regression. These tabulations are based on an organic carbon database compiled prior to 2003 from a number of sources including EPA's STOrage and RETrieval Data Warehouse (STORET) and the USGS NWIS. The utility of these tabulations could be improved by updating them to incorporate newer information. For example, EPA recently released data from the Wadeable Stream Assessment, which included DOC measurements from a statistically based random sample of ~2,000 streams. Other statistically-based national water quality surveys, including national assessments of lakes and large rivers, will also be providing additional data in future years.

The Monte Carlo method developed to calculate FSC for copper was applied at each of the four sites, both with and without the correlation between filtered copper concentrations and the BLM water quality parameters that were found to be significant at three of the sites. We also approximated the FSC using the 10th percentile of the distribution of IC predicted by the BLM at each site. When copper concentration correlations were considered in the FSC calculations, the 10th percentile of the IC distributions was found to be highly conservative approximations of the FSC, underestimating the FSC by 44 to 70%. This is illustrated in Figure C-1, which also shows the good agreement between IC predicted with the BLM using site-specific data and IC predicted using measured pH and organic carbon but projected values of the GI BLM input parameters. Ecoregion and water body-type specific DOC concentration percentiles ("L3-DOC" in the figure below) were also an improvement over the projections based on conductivity regressions.

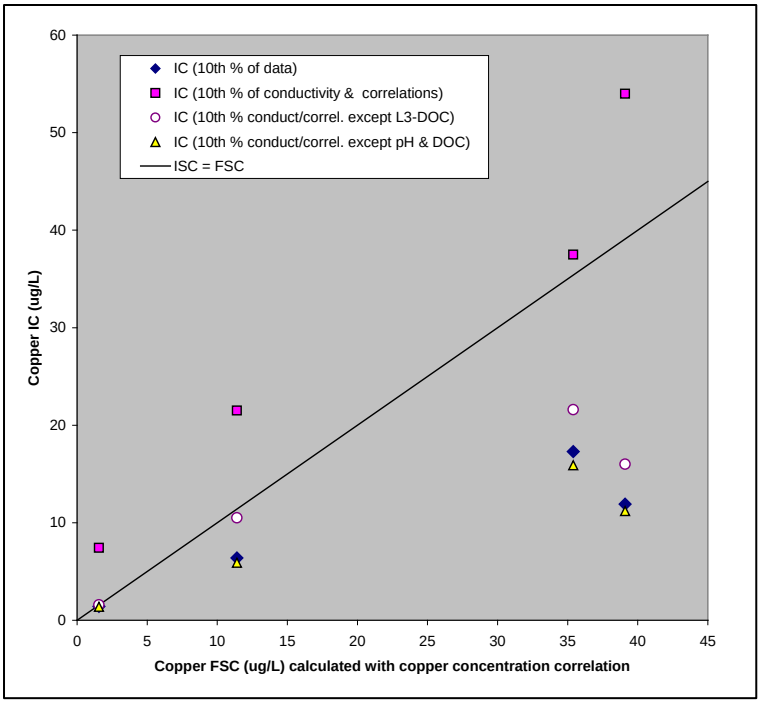

Figure C-1. Instantaneous Criteria (IC) predicted with the BLM using site-specific data and IC predicted using measured pH and organic carbon and projected values of the GI BLM input parameters

When copper concentration correlations were neglected in the FSC calculations, the 10th percentile of the IC distributions did a much better job approximating the FSC. This is shown in Figure C-2. In this case, the 10th percentile of the IC distributions was within 15% of the FSC. This figure also shows the good agreement between IC predicted with data and projected values of the GI BLM input parameters.

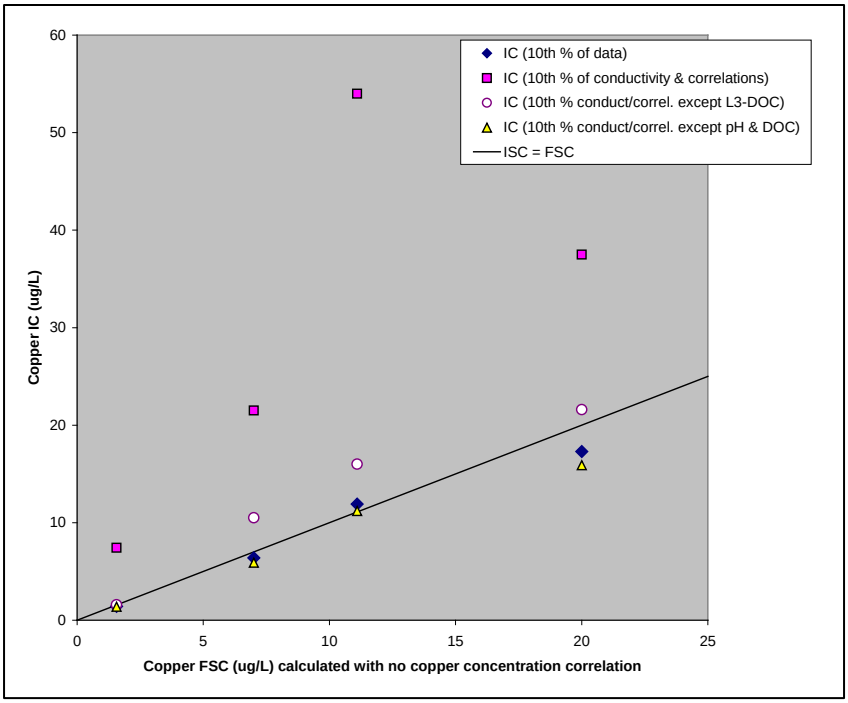

Figure C-2. 10th percentile of the IC distributions using data and projected (predicted) values of the GI BLM parameters

The degree of correlation between filtered copper concentrations and BLM input water quality parameters appears to be an important site-specific factor in determining the relationship between the FSC and the IC. Figure C-3 plots the percentile of the IC corresponding to the FSC for each site as a function of the correlation coefficient between the copper concentrations and the IC, for two cases: (1) FSC calculated by the Monte Carlo method including the observed correlations between concentrations of copper and BLM input water quality parameters, and (2) FSC calculated with no correlation between concentrations of copper and BLM input water quality parameters. In the first case (plotted with dark diamond symbols), the percentile of the IC corresponding to the FSC increases substantially (6th to 46th percentile) as the correlation coefficient between the copper concentrations and the IC increases. If the correlation between concentrations of copper and BLM input water quality parameters is neglected (the second case, plotted in lighter square symbols), the percentile of the IC corresponding to the FSC is considerably lower (4.3rd to 14th percentile). This suggests that correlations between copper concentrations and BLM input parameters should be given careful consideration when calculating FSC.

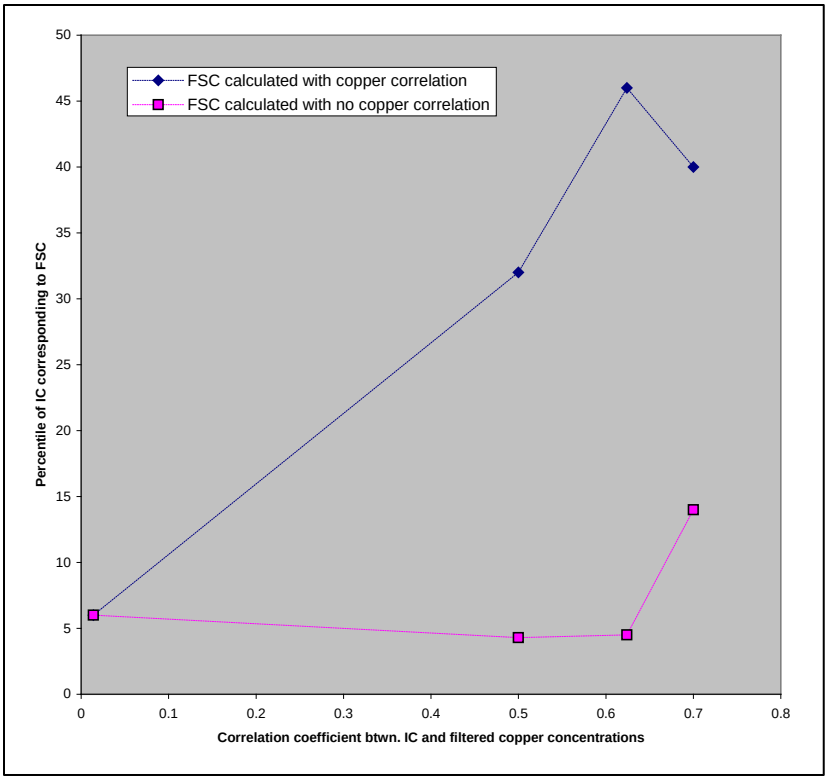

Figure C-3. Percentile of the IC corresponding to the FSC for each site as a function of the correlation coefficient between the copper concentrations and the IC when the FSC is calculated with Copper correlation and when FSC is calculated without Copper correlation

C.4 Combining GI-Conductivity Regressions with Geostatistical Techniques

Geostatistical techniques are attractive because they explain parameter variation arising from spatial correlations, which are otherwise ignored by (and may, in fact, violate the assumptions of) conventional statistics. BLM input water quality parameters (except for pH and DOC) are GIs, the concentrations of which vary in surface water due to dissolution, weathering, ground water-surface water interactions, and other geologic processes in the watershed. Consequently, the concentrations of GI parameters tend to vary according to the regional geology. For example, water hardness has noticeable geographic trends. Areas with limestone geology, such as in the prairie states, tend toward high hardness and alkalinity. Areas of with granite geology, such as parts of the Northeast, tend toward low hardness and alkalinity. The estimation of GI parameter values based on geography thus seems possible. EPA has provided a prototype of a geostatistical approach[8] that demonstrated this potential. That work applied kriging to predict median concentrations of five of the BLM water quality input parameters (pH, DOC, alkalinity, sodium, and calcium) averaged over 8-digit HUCs, using the USGS NWIS as the source of spatial data. Comparison of measured concentrations with kriging predictions were encouraging, especially for DOC and alkalinity. Geostatistical techniques to project BLM GI input

parameters might well be developed from the same nationwide monitoring data used to develop a correlation approach. By the same token, geostatistical techniques based on these data may suffer the same problems experienced when developing the correlation approach. Most significantly, the NWIS data are not randomly distributed in either time or space, and the measurements of the BLM GI parameters are generally uneven (considerable differences in terms of the number of observations for different parameters) and/or inconsistent (i.e., relatively few concurrent measurements of BLM GI parameters).

There may be great value in supplementing the geostatistical approach with classical estimation methods, such as regression and correlation. Examination of the NWIS data suggests that conductivity may be useful for estimating BLM input water quality parameters in conjunction with geostatistics. The literature indicates that conductivity is one of the most widely monitored water quality indicators in the US. Among water quality parameters, the data for conductivity are the most complete and cover the longest time period (Wang and Yin, 1997). In part, this is because conductivity measurements are usually included in automated multiparameter systems for monitoring changes in the quality of surface waters (Allen and Mancy, 1972). A vast amount of conductivity data exists, both in terms of the total number of observations and the number of sites reporting this parameter in comparison to the BLM GI quality parameters. For example, NWIS data for the state of Colorado have almost four times as many observations of conductivity as for calcium, and they are measured at more than twice the number of sites. There are 20 times as many observations of conductivity as for alkalinity, and they are measured at more than seven times the number of sites. Since conductivity data are abundant, and correlate well to the BLM GI parameters (GLEC, 2007), it is reasonable to incorporate conductivity in spatial projections of BLM parameters. This may simplify the geostatistical approach and allow more robust spatial projections of BLM water quality parameters.

Although combining GI-conductivity regressions with geostatistical techniques seems promising for the reasons mentioned above, this approach had never been demonstrated. We conducted a simple test using NWIS conductivity and hardness data from the state of Colorado. We used data from Colorado because many more stations were sampled in comparison to the surrounding states.

The data were processed in a manner similar to the methods used to develop the regressions in Section C.2. For each station, we calculated the 10^{th} percentiles of conductivity and hardness. A regression model was fit to the full dataset (data for all rivers and streams sampled in Colorado between 1984 to 2005). The following regression model was developed to project hardness from conductivity:

$$\ln(\text{hardness}) = 0.984 \cdot \ln(\text{EC}) - 0.870$$

We also kriged the 10^{th} percentiles of conductivity and hardness, using latitude and longitude coordinates reported by USGS for each sampling station. Figure C-4 shows the kriged surface of the 10^{th} percentile of conductivity at all stations in Colorado, Utah, and Wyoming. Data are far more abundant in Colorado, as shown by the density of the dots representing the locations of sampling stations. Figure C-5 shows the kriged surface of the 10^{th} percentile of hardness at all stations in Colorado. Kriging was done using the Vertical Mapper program, version 3.1; no attempts were made to optimize the kriging of conductivity or hardness by parameter adjustment.

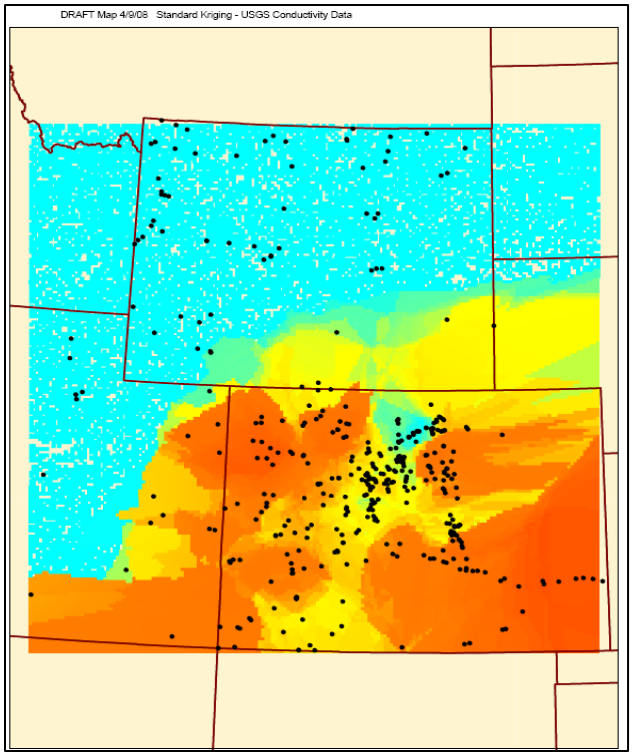

Figure C-4. Kriged surface of the 10th percentile of conductivity at all stations in Colorado, Utah and Wyoming

Dots represent sampling stations; notice that data are far more abundant in Colorado.

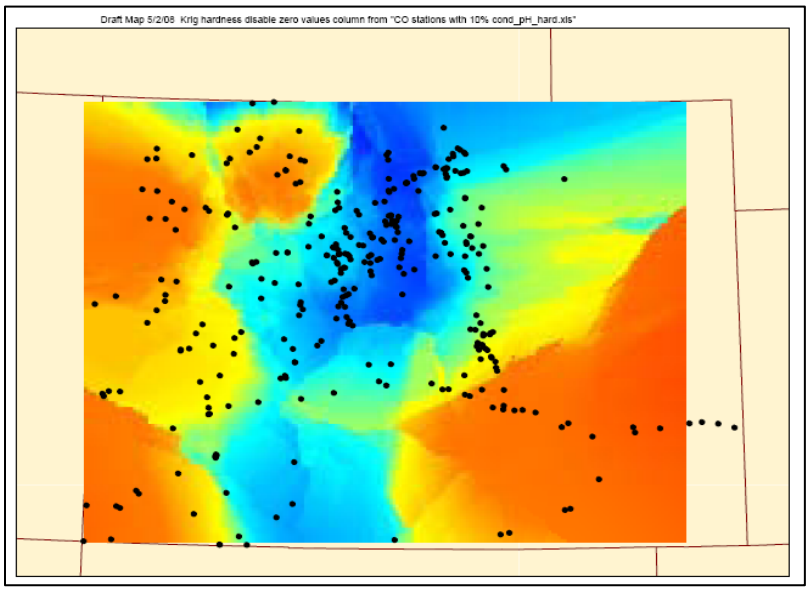

Figure C-5. Kriged surface of the 10th percentile of hardness at all stations in Colorado

Our goal was to see whether combining the kriged conductivities with the conductivity-hardness regression would project the 10th percentiles of hardness better than direct kriging of the hardness

data. For the combined kriging/regression approach, we determined the kriged conductivity values at all of the sampling locations and then projected the 10th percentiles of hardness at these locations using the regression equation. We also determined the directly-kriged 10th percentiles of hardness at all of the sampling locations.

The hardness estimates obtained by each approach were then compared to the 10th percentiles of hardness measured at each station. The results of this comparison are shown graphically in Figure C-6. Both approaches produce estimates of hardness that correlate significantly with the measured data (correlation coefficient r= 0.80 for direct kriging of hardness; r= 0.950 for conductivity kriging + regression projection). However, the kriging+regression approach fits the hardness data substantially better than direct kriging. To quantify this, we calculated the residual sum of squares (RSS), a composite measure of the discrepancy between the data and our alternative hardness estimates. The smaller this discrepancy is, the better the estimation will be. In natural log space, the RSS for the kriging+regression approach is 18.6 (135 degrees of freedom, or df) while the log-space RSS for the direct kriging approach is 73.4 (136 df). Thus, for this test case substantially better estimates of the 10th percentile of hardness were made by the kriging/regression approach compared to direct kriging.

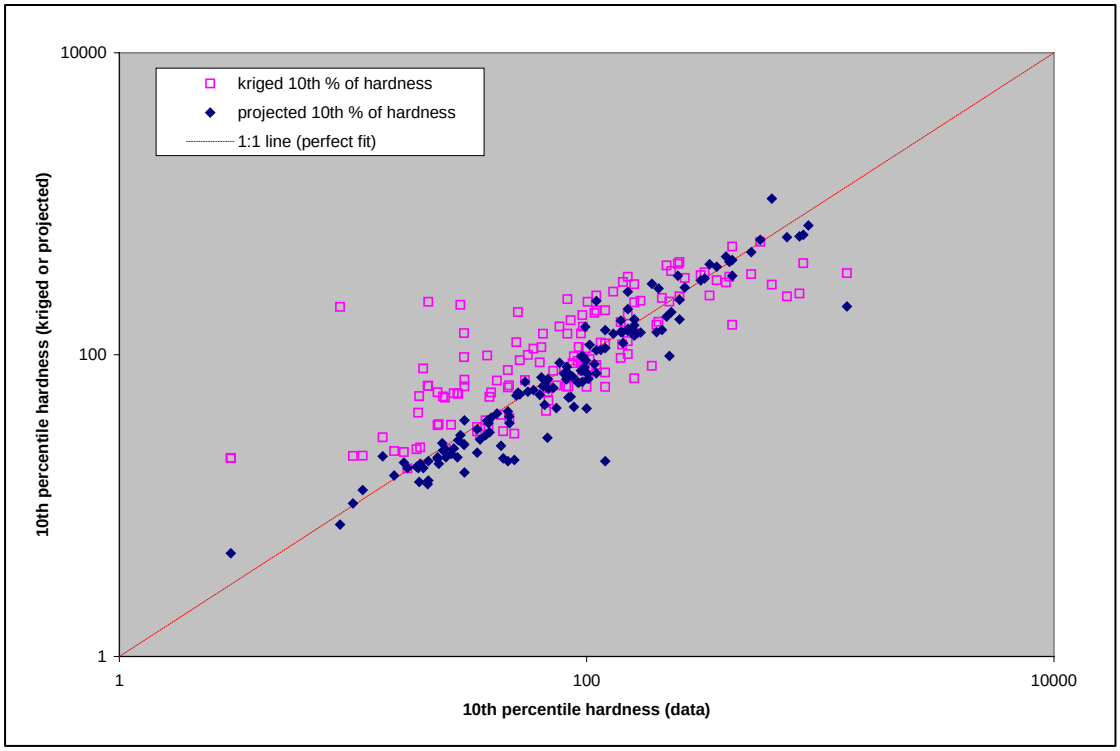

Figure C-6. Comparison of the 10th percentile of hardness at all stations in Colorado with estimates based on (a) direct kriging of hardness data and (b) kriging of conductivity to station locations and projecting conductivity to hardness via regression ("kriging/regression")

As this test demonstrated, combining kriging with regressions to project BLM GI inputs from conductivity appears to improve the accuracy of estimates of parameters used as BLM inputs. Applying the conductivity kriging/regression projection approach on a broader scale should be considered as a

"next step" in developing tools to estimate BLM water quality parameters for sites where there may be few or no data available to characterize water quality. Since direct kriging of most BLM GI parameters has already been done using data from NWIS, it will also be worthwhile to continue comparing the alternative estimates to the observed data in order to obtain the best estimates.

We should also note that although the kriging/regression approach can be used to improve the accuracy of estimates of GI parameters used as BLM inputs, this approach cannot be expected to produce accurate site-specific estimates for the two most important BLM inputs: pH and DOC. As shown in Section C.3, accurate estimates of the GI parameters are less important than pH and DOC in terms of predicting appropriate site-specific IC and FSC. Since our analysis of NWIS data indicates there to be either little no trend between conductivity and pH, and direct kriging produced similarly ambiguous predictions, we must conclude that site-specific data for pH must either be available or be collected for BLM application at a site. This may not be a significant obstacle, since pH data can be cheaply and readily acquired.

Lack of methods to accurately estimate DOC is a bigger problem, since measurements of this parameter are comparatively rare and DOC is a relatively expensive measurement to make. For DOC, analysis of NWIS data again indicates no trend with conductivity, so the kriging/regression approach is not appropriate for this parameter. However, other analyses conducted suggested that DOC could be kriged with some success. And, as was demonstrated for the test sites in Section C.3, the ecoregion and water body-type specific DOC concentration percentiles tabulated by EPA for the National Bioaccumulation Factors Technical Support Document appear to offer reasonable estimates of lower-percentile DOC concentrations. Further development of these approaches for estimating site-specific DOC appears worthwhile, for example by incorporating new data from the Wadeable Stream Assessment and other statistically-based national water quality surveys.

C.5 References

Allen, H.E. and K.H. Mancy. 1972. Design of measurement systems for water analysis. In: Ciaccio, L.L., ed. Water and water pollution handbook. Marcel Dekker, Inc. New York, N.Y.

USEPA. 2002. Development of Methodologies for Incorporating the Copper Biotic Ligand Model into Aquatic Life Criteria: Application of BLM to Calculate Site-Specific Fixed Criteria. GLEC Work Assignment 3-38, Contract No. 68-C-98-134. 63p plus figures.

USEPA. 2007. Approaches for Estimating Missing BLM Input Parameters: Correlation approaches to estimate BLM input parameters using conductivity and discharge as explanatory variables. GLEC Work Assignment 2-34, Task 1, Subtask 1-7. Contract No. 68-C-04-006. 18 p.

Wang, X. and Z.Y. Yin. 1997. Using GIS to assess the relationship between land use and water quality at a watershed level. Environment International. 23(1): 103-114.

Appendix D: Approaches for Estimating Missing Biotic Ligand Model Input Parameters: Projections of Total Organic Carbon as a Function of Biochemical Oxygen Demand

D.1 Introduction

The 2007 Update of the Ambient Water Quality Criteria for Copper (EPA-822-R-07-001) employs the Biotic Ligand Model (BLM) to estimate bioavailability of this metal in toxicity tests used in Criterion Maximum Concentration derivation, which requires data on the 10 input parameters for the BLM, including dissolved organic carbon (DOC). Data for DOC concentrations, in both effluents and receiving waters, are extremely limited. The BLM is very sensitive to DOC concentrations (HydroQual, 2005), which means that to ensure accurate predictions of copper bioavailability and toxicity reliable data on DOC concentrations in the water are needed. Effluent DOC concentrations, which are necessary for application of the BLM to predict copper toxicity associated with a wastewater discharge, are monitored by very few publicly-owned treatment works (POTWs).

Projections of DOC concentrations from biochemical oxygen demand (BOD) values may be a viable solution for surmounting the lack of data on DOC. Effluent BOD (most typically 5-day BOD) is monitored by most POTWs. We expect a positive correlation between BOD and DOC, because the two parameters are conceptually related. While DOC quantifies the concentration of many organic compounds dissolved in water, BOD is a routine surrogate test for estimating the load of organic carbon into the environment. Ideally, one might expect an almost stoichiometric relationship between organic carbon (i.e., DOC) and the oxygen consumed during its metabolization (i.e., BOD). For instance, Fadini et al. (2004) evaluated the possible replacement of BOD for DOC measurements in a number of different wastewater categories. A statistical relationship between effluent BOD and DOC would provide estimates of DOC concentrations, needed for application of the BLM, from routine BOD monitoring data. The effluent contribution to in-stream DOC could then be estimated, for example, by using a dilution model for a site.

Evidence, from analyses of effluent monitoring data from the New York State Department of Environmental Conservation (NYSDEC) Contaminant Assessment and Reduction Project (CARP), suggests that most of the total organic carbon (TOC) in POTW effluent is in the form of DOC. Therefore, a regression between BOD and TOC could be used as a surrogate for the relationship between BOD and DOC. The advantage of using TOC is the greater availability of data. TOC is reported for a significant number of major POTW dischargers.

D.2 Data

In 2006, monitoring data from all major POTWs reporting TOC and 5-day BOD in the United States were downloaded from the U.S. Environmental Protection Agency (EPA) Permit Compliance System (PCS) web site http://www3.epa.gov/enviro/facts/pcs-icis/search.html. Nine POTWs had 30 or more synchronous records of TOC and BOD, while 23 POTWs had at least 10 synchronous records. These numbers include both monthly average and maximum monthly values.

Review of the data indicated several extremely high (>1,000 milligrams per liter [mg/L]) effluent TOC values for discharger CA0079243. We assumed that they presented errors in the reported unit, and divided them by 1,000 to convert from units of microgram per liter (µg/L) to mg/L. TOC and BOD

records were matched by POTW, location (e.g., upstream, downstream, influent or effluent), year and month. Thus, "synchronous" measurements do not necessarily correspond to samples collected on the same day and time. The resulting table had 341 records.

D.3 Results

D.3.1 TOC and BOD at All Monitoring Locations

The first statistical evaluation involved data for all monitoring locations at the eight POTWs reporting 30 or more synchronous records of TOC and BOD. Table D-1 presents the results of least squares regression of the average monthly data: TOC*avg* = *a* + *b* BOD*avg*. A scatter plot of this data is shown in Figure D-1. Table D-2 presents the results of least squares regression of the maximum monthly data: TOC*max* = *a* + *b* BOD*max*. A scatter plot of this data is shown in Figure D-2. Bimodal distributions are observed for TOC and especially BOD in this data set. It should be noted that the BOD concentrations of 200 mg/L or higher were measured in samples of untreated (influent) wastewater; TOC concentrations were also quite high in these samples. Both scatter plots (Figures D-1 and D-2) show a fairly strong correlation between TOC and BOD in the combined data for all POTWs. The linear relationship between TOC and BOD is better defined in the average data (Figure D-1).

Table D-1. Least squares regression of average monthly TOC and BOD data for all monitoring locations

POTW	Location	Intercept (a)	Slope (b)	r^2	df
CA0054372	Effluent Gross Value				
CA0105295	Effluent Gross Value	7.551	-0.379	0.009	41
CA0105295	Raw Sew/Influent	19.935	0.142	0.104	59
CA8000326	Effluent Gross Value*				
CA8000383	Effluent Gross Value	4.952	0.725	0.344	31
CA8000383	Raw Sew/Influent	59.586	0.107	0.038	30
ID0020443	Upstream Monitoring*				
ID0020443	Downstream Monitoring	3.500	-0.400	0.190	2
ID0023981	Effluent Gross Value	6.268	0.281	0.196	26
ID0023981	Upstream Monitoring	3.200	-0.300	0.127	3
LA0073521	Effluent Gross Value*				
TN0023353	Effluent Gross Value	2.628	0.391	0.438	35
All POTWs	All locations	4.828	0.237	0.873	243

*note: POTW/location without regression results indicates less than 2 synchronous data records

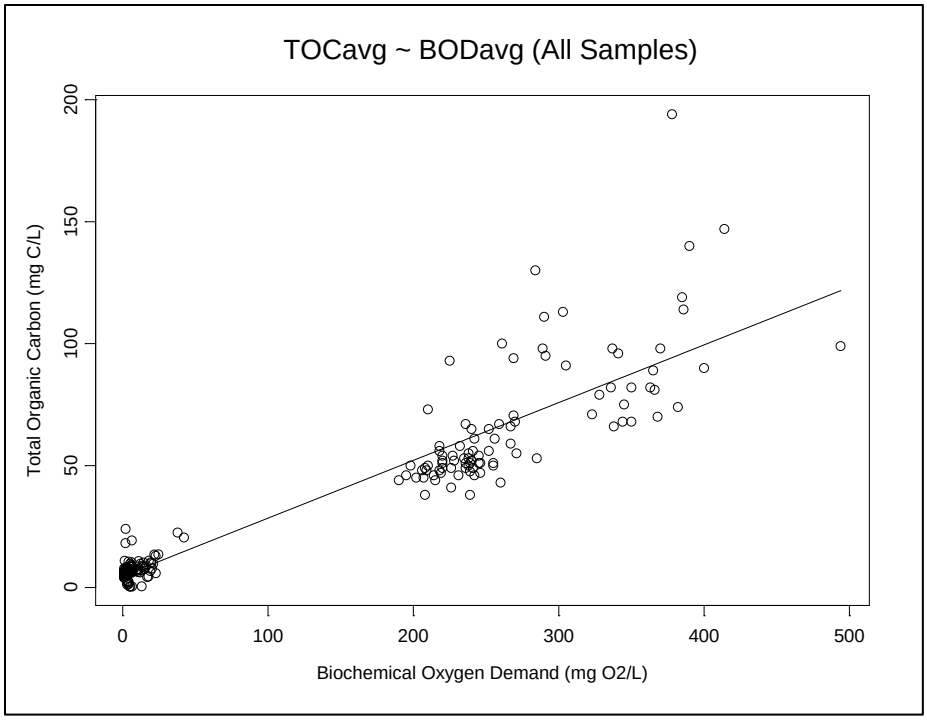

Figure D-1. Scatter Plot of Average Monthly Data (all Monitoring Locations)

Table D-2. Least squares regression of maximum monthly TOC and BOD data for all monitoring locations

POTW	Location	Intercept (a)	Slope (b)	r^2	df
CA0054372	Effluent Gross Value	9.420	0.499	0.154	29
CA0105295	Effluent Gross Value	8.567	-0.114	0.007	50
CA0105295	Raw Sew/Influent	56.439	0.062	0.046	59
CA8000326	Effluent Gross Value	7.307	0.006	0.000	28
CA8000383	Effluent Gross Value	6.674	0.507	0.235	31
CA8000383	Raw Sew/Influent	149.293	0.047	0.052	30
ID0020443	Upstream Monitoring	0.300	1.175	0.039	4
ID0020443	Downstream Monitoring	3.500	-0.400	0.190	2
ID0023981	Effluent Gross Value	6.210	0.208	0.202	28
ID0023981	Upstream Monitoring	3.200	-0.300	0.127	3
LA0073521	Effluent Gross Value	12.989	-0.818	0.110	18
TN0023353	Effluent Gross Value*				
All POTWs	All locations	11.183	0.196	0.700	302

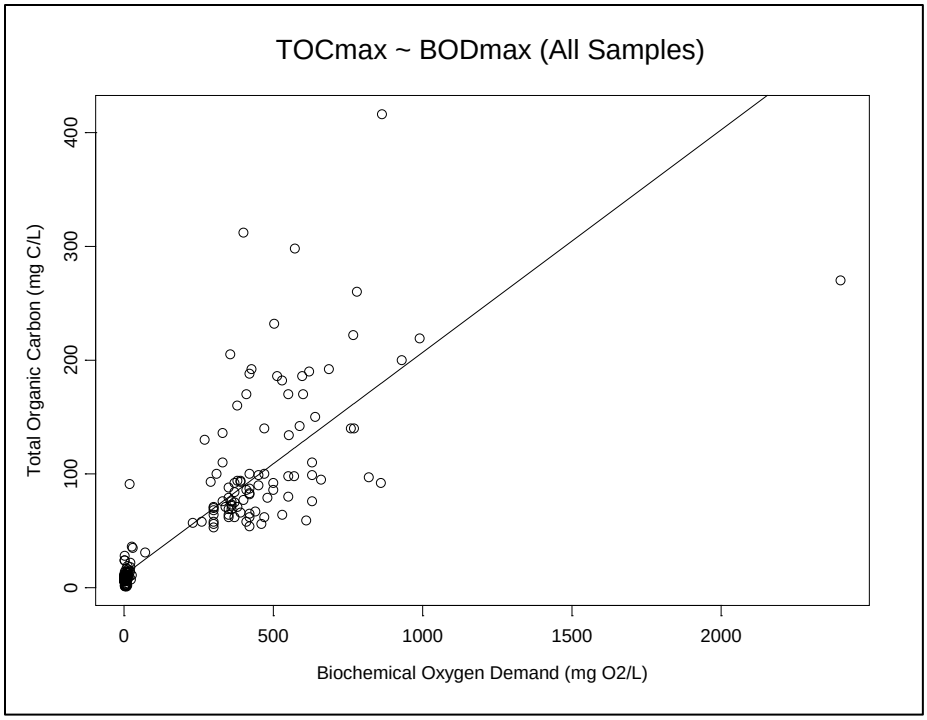

Figure D-2. Scatter Plot of Maximum Monthly Data (all Monitoring Locations)

Results of regression analyses revealed large differences in slopes of the linear model TOC = *a* + *b* BOD among locations and POTWs. Slopes for individual regressions ranged from –0.40 to 0.73 for average, and from -0.82 to 0.50 for maximum BOD and TOC values. Coefficients of determination (r^2) for the regressions were low; for most of them $r^2 < 0.2$. Pooling the data from all POTWs and locations increased the r^2 to 0.87 for average and 0.70 for maximum BOD and TOC values.

Diagnosis of regression analyses revealed that variance in both the average and maximum TOC rose with increasing values of biochemical oxygen demand. Such patterns were also evident from a simple inspection of the plots cited above. Homogeneity of variance, though, is a core assumption of ordinary least squares regression, and its violation compromises the quality of results generated by the analysis. The solution was to perform quantile regression analysis because it does not assume that variance of the response is homogeneous along the range of the independent variable. The fitted model for the 50^{th} quantile (median) was:

$$TOC_{avg} = 5.5647 + 0.2088\ BOD_{avg} \quad (243\ df,\ R^1 = 0.77)$$

D.3.2 TOC and BOD at Effluent Monitoring Locations

Although the quantile regression model above provided a reasonable fit of the data at all monitoring locations, we were specifically interested in the relationship between TOC and BOD measured in POTW effluents. Therefore, we conducted a separate statistical analysis of effluent monitoring data from the 17 POTWs with more than one synchronous record of TOC and BOD retrieved from PCS. TOC and BOD records were again matched by POTW, year, and month. The resulting data tabulation had 373 records.

The results of least squares regression of the average monthly effluent data: TOC*avg* = *a* + *b* BOD*avg* are presented in Table D-3. A scatter plot of this data is shown in Figure D-3. Table D-4 presents the results of least squares regression of the effluent maximum monthly data: TOC*max* = *a* + *b* BOD*max*. A scatter plot of this data is shown in Figure D-4.

Table D-3. Least squares regression of average monthly TOC and BOD data for effluent monitoring locations

POTW	Intercept (a)	Slope (b)	r^2	df
CA0054372				
CA0077691				
CA0079103				
CA0079243				
CA0102822				
CA0105295	7.5512	-0.3789	0.009	41
CA0107492				
CA0109991				
CA8000073				
CA8000326				
CA8000383	4.9522	0.7254	0.344	31
ID0023981	6.2679	0.2808	0.196	26
LA0069868				
LA0073521				
TN0023353	2.6284	0.3909	0.438	35
TN0023531	37.6422	-1.1348	0.076	3
TN0023574	7.4142	0.0882	0.016	25
All POTWs	5.8740	0.2859	0.245	174

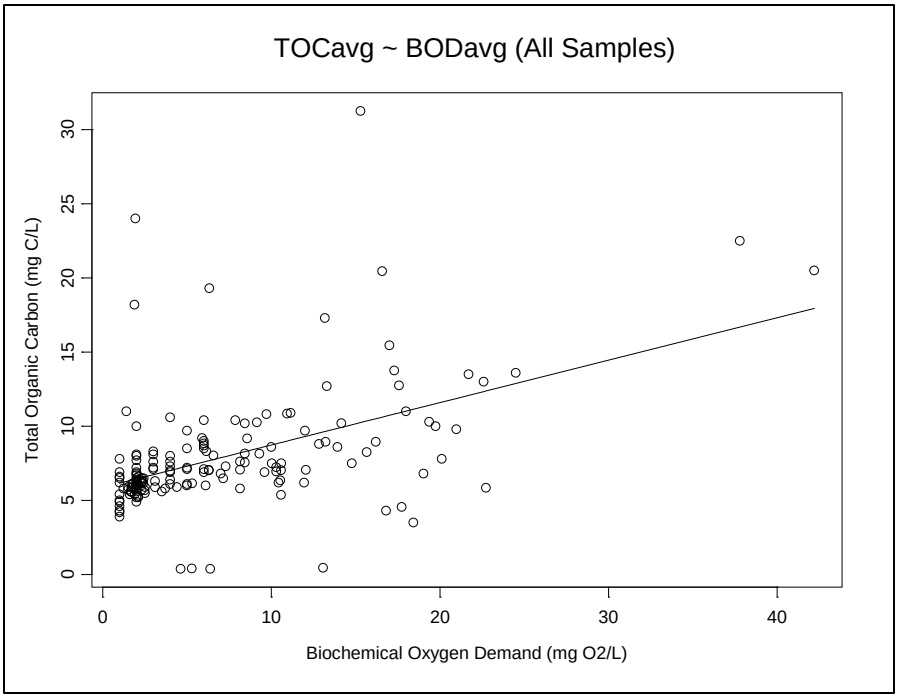

Figure D-3. Scatter Plot of Average Monthly Data (Effluent Monitoring Locations)

Table D-4. Least squares regression of maximum monthly TOC and BOD data for effluent monitoring locations

POTW	Intercept (a)	Slope (b)	r^2	df
CA0054372	9.4197	0.4993	0.154	29
CA0077691	4.2750	0.7008	0.422	6
CA0079103	23.3517	-0.1278	0.035	16
CA0079243	5.7979	-0.0367	0.008	8
CA0102822	6.3526	0.1780	0.195	28
CA0105295	8.5667	-0.1143	0.007	50
CA0107492	5.9043	1.0676	0.027	2
CA0109991	3.0331	0.8458	0.167	11
CA8000073	9.0000	0.0000	0.000	8
CA8000326	7.3073	0.0056	0.000	28
CA8000383	6.6743	0.5070	0.235	31
ID0023981	6.2101	0.2083	0.202	28
LA0069868	7.6475	0.1139	0.535	5
LA0073521	12.9886	-0.8184	0.110	18
TN0023353				
TN0023531				
TN0023574				
All POTWs	6.6930	0.4311	0.276	299

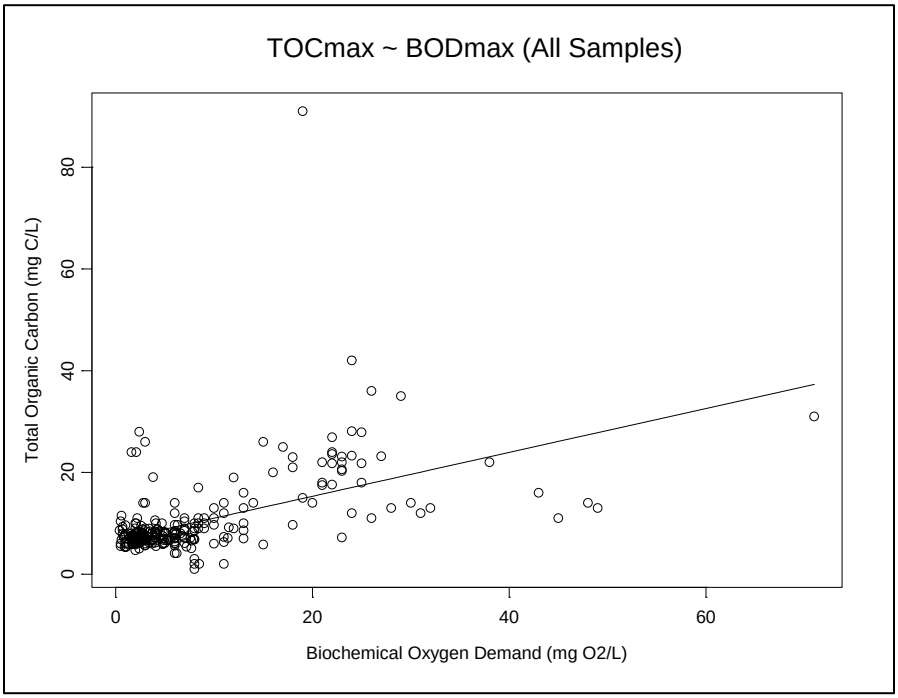

Figure D-4. Scatter Plot of Maximum Monthly Data (Effluent Monitoring Locations)

Results of the effluent regression analyses revealed large differences in slopes of the linear model TOC = *a* + *b* BOD among POTWs. Slopes for individual regressions ranged from –1.13 to 0.73 for average, and from -0.82 to 1.07 for maximum BOD and TOC values. Coefficients of determination (r^2) for the regressions were low; all $r^2 < 0.55$ and for most of them $r^2 < 0.24$. Low coefficients of determination were also recorded for regressions of TOC on BOD values from all POTWs ($r^2 = 0.245$ and 0.276, for average and maximum values, respectively). Further investigations of the effluent regression analyses were performed, because visual inspection of Figures D-3 and D-4 suggested the presence of outliers in the data.

We examined the fit of the linear model, TOC*avg* = *a* + *b* BOD*avg*, by inspecting its residuals (Figure D-5). Studentized residuals were plotted against projected (fitted) TOC values in the left pane, and against quantiles of the standard normal distribution in the right pane. Four suspiciously-low average TOC points in Figure D-4 are labeled '324', '308', plus the two points adjacent to the latter (left pane). This plot reveals that residuals for high-TOC points '489' and '64' are far larger in magnitude than residuals for the four suspiciously-low points. Residuals for these two points greatly deviate from the normal distribution (right pane). Furthermore, points '335' and '336' have much greater leverage than any other (leverages for '335': 0.164, '336': 0.127). Fitting the linear model without points '489', '64', '335', and '336' results in the following parameter values:

TOC*avg* = 6.0388 + 0.2171 BOD*avg* ($r^2 = 0.185$, 170 df) (Equation 1)

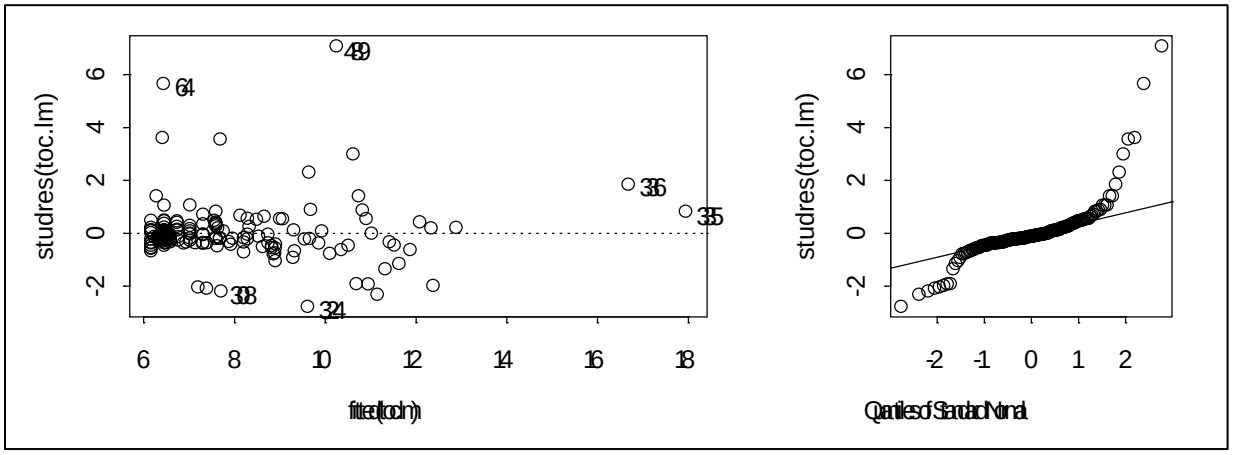

Figure D-5. Residuals of the linear model, TOCavg = a + b BODavg

(Left: plot of Studentized residuals (studres) against projected (fitted) TOC values; Right: plot of studies against quantiles of the standard normal distribution).

It should be noted that this model (Equation 1) projects average TOC concentrations very similar to the regression based upon the uncensored data (i.e., within ± 2 mg/L), for the range of BOD concentrations of interest (less than 30 mg/L).

Diagnosis of the regression analysis, TOC*max* = *a* + *b* BOD*max*, revealed an excessively high residue for the (19, 91) point and very high leverage for the (71, 31) point. The projected regression line without those two points was:

$TOCmax = 6.6242 + 0.4095\ BODmax$ ($r^2 = 0.352$, 297 df) (Equation 2)

This model (Equation 2) fits a single slope for all data. Our results, though, revealed large differences in slopes of regression lines among POTWs (Table D-4). We tested the significance of such differences with an *F*-test, which required fitting two additional models, one with the same slope for all POTWs and the other with a distinct slope for each POTW. The *F*-test compares model fits while taking into account the loss in degrees of freedom associated with the computation of multiple slopes. The estimated *F*-value ($F = 4.92$, 13 df) was highly significant ($P < 0.001$), indicating that distinct slopes are necessary to accurately project maximum TOC from maximum BOD values.

D.3.3 TOC and DOC at CARP Effluent Monitoring Locations

Effluent discharge samples were collected from 11 New Jersey POTWs in 2000 and 2001 for the NYSDEC CARP project (www.dec.state.ny.us/website/dow/bwam/CARP). These samples were analyzed for DOC, particulate organic carbon (POC), and total suspended solids (TSS) by the U.S. Geological Survey. TOC was calculated by adding together DOC and POC concentrations. The results are shown in Table D-5. Effluent DOC concentrations are generally much higher than POC because most of the particulate organic matter is removed from wastewater during secondary treatment. A scatter plot of the TOC and DOC data, Figure D-6, shows the strong linear correlation between TOC and DOC that results from the predominance of DOC in effluent. These data are replotted in Figure D-7 for TOC concentrations less than 50 mg/L.

Table D-5. CARP organic carbon and total suspended solids (TSS) monitoring data for New Jersey discharger

DATE	SITE	DOC (mg/L)	POC (mg/L)	TOC (mg/L)	TSS (mg/L)
Oct. 2-4, 2000	PVSC	43.0	10.5	53.5	51.4
	MCMUA	0.10	6.75	6.85	36.3
	BCMUA	22.2	10.6	32.8	54.1
	JMEU	8.51	8.29	16.8	19.2
	RVMUA	12.2	3.81	16.0	22.1
	LRMUA	8.76	4.71	13.5	9.3
Dec. 11-15, 2000	PVSC	50.3	5.35	55.7	25.9
	MCMUA	260	9.22	269	62.6
	BCMUA	20.0	2.73	22.8	14.4
	JMEU	23.0	5.23	28.2	31.1
	RVMUA	23.4	8.73	32.1	42.0
	LRMUA	10.4	11.4	21.8	55.2
	NHH	14.0	3.07	17.1	22.5
	NBC	28.6	6.67	35.3	23.3
	NBW	21.8	3.38	25.2	7.8
	NHWNY	18.7	5.66	24.3	18.1
	SMUA	15.8	2.89	18.6	6.6
May 21-23, 2001	PVSC	34.5	14.2	48.7	41.1
	BCMUA	15.0	9.17	24.1	11.9

DATE	SITE	DOC (mg/L)	POC (mg/L)	TOC (mg/L)	TSS (mg/L)
	RVMUA	9.26	10.2	19.5	12.0
	LRMUA	14.7	9.34	24.1	10.9
	EMUA	0.25	0.15	0.40	19.9
August 6-9, 2001	PVSC	123	8.74	132	35.6
	MCMUA	20.6	5.34	25.9	22.6
	BCMUA	109	15.4	125	45.6
	JMEU	131	8.58	140	18.1
	RVMUA	8.78	3.39	12.2	6.7
	LRMUA	7.33	5.01	12.3	17.9
	NBC	191	8.39	199	17.4
	NBW	23.4	9.82	33.3	13.5
	EMUA	14.7	5.33	20.1	7.5
	NHWNY	17.7	12.1	29.8	13.8
	SMUA	10.7	2.67	13.4	3.8

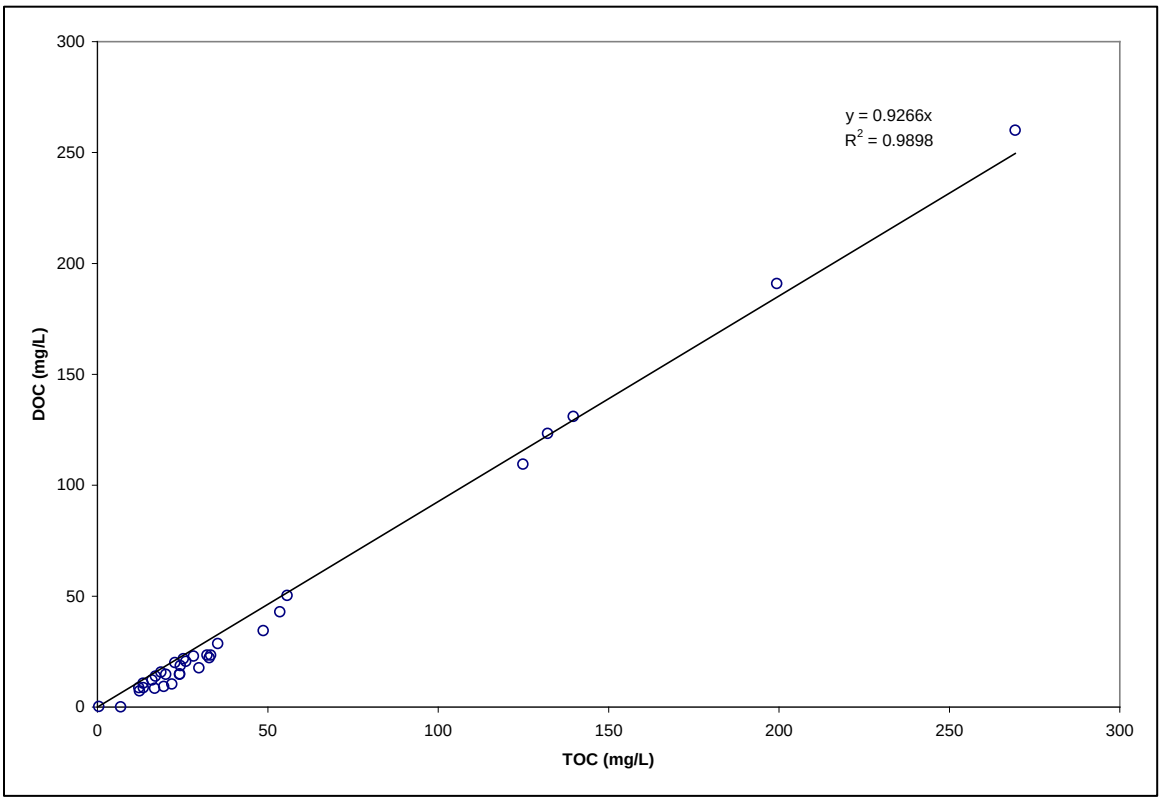

Figure D-6. Scatter plot of TOC versus DOC in CARP effluent monitoring data

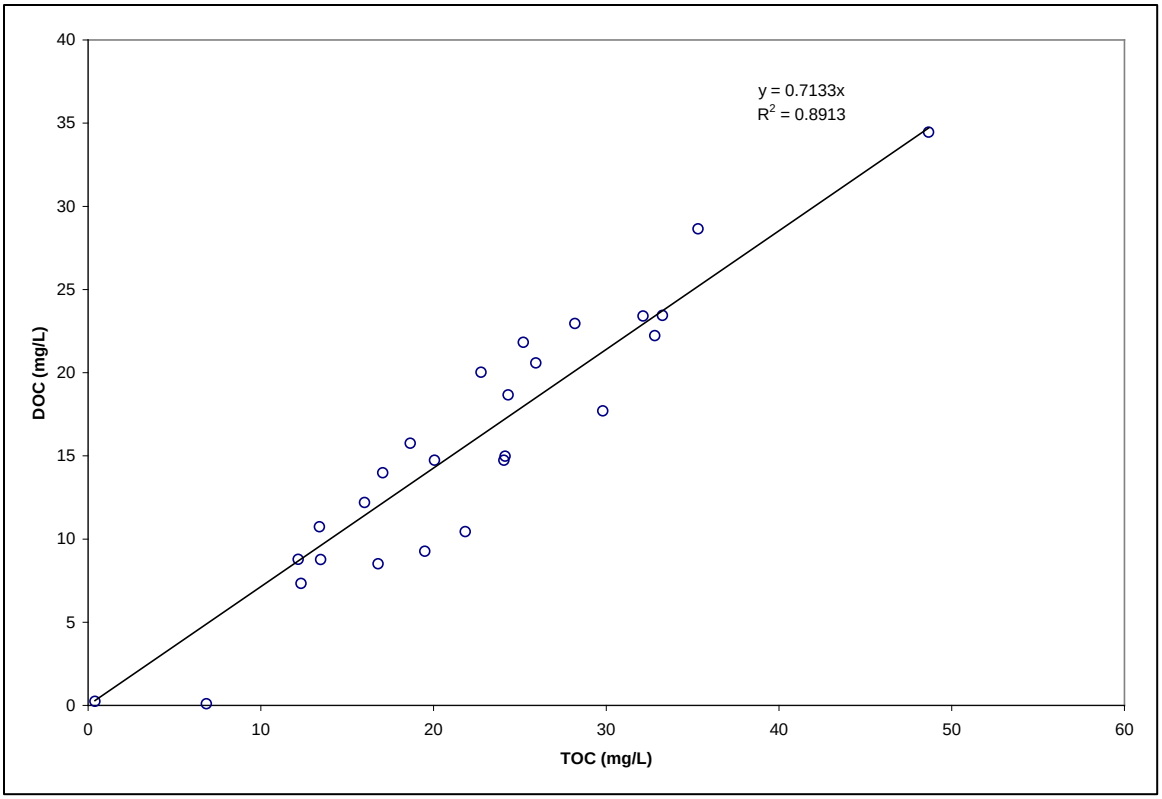

Figure D-7. Scatter plot of TOC versus DOC in CARP effluent monitoring data (TOC <50 mg/L)

Least squares regression of the CARP effluent data (Table D-5) produces the following model:

DOC = 0.9266 TOC (r^2 = 0.9898, 32 df)

This regression was forced through the origin by constraining the intercept to be zero. At the limit of removal efficiency (i.e., as effluent TOC approaches zero), any remaining TOC should be in the form of DOC, as mentioned above. This argument justifies forcing the regression through the origin. If only the data for which TOC falls in the expected range for effluent concentrations (TOC < 50 mg/L) are considered, the regression (again forced though the origin) is:

DOC = 0.7133 TOC (r^2 = 0.8913, 25 df)

For either case, the CARP effluent data show the strong linear relationship between TOC and DOC. Because TOC and DOC are linearly related in POTW effluent, the relationships between BOD and TOC reported above (Sections D.1 and D.2) also apply to DOC.

D.4 Discussion

Initially, we attempted to correlate BOD and TOC concentration measurements using data for all monitoring locations retrieved from PCS for major POTWs. We produced significant linear regression models for both average (Figure D-1) and maximum monthly (Figure D-2) data. Coefficients of determination were 0.87 and 0.70, respectively, for these data when combined for all locations. However, these correlations were substantially influenced by very high (i.e., greater than 50 mg/L) concentrations of BOD and TOC measured in untreated wastewater.

When we repeated the statistical analysis using effluent monitoring data, we found large differences in the slopes of the linear model TOC = *a* + *b* BOD among POTWs. Low coefficients of determination were also recorded for regressions of TOC on BOD values from all POTWs (r^2 = 0.245 and 0.276, for average and maximum values, respectively). In part, this may reflect random errors in the measurements of BOD and TOC, since data quality issues including loss of precision tend to be more frequent and significant at lower concentrations. The greater scatter in the plots of effluent BOD and TOC (Figures D-3 and D-4) may also reflect the limitations of working with the monthly average and maximum data reported by PCS.

Direct inspection of the TOC data in Figures D-3 and D-4 is nevertheless instructive. Aside from some extreme high and low values, the great majority of effluent TOC concentrations are in the range of 5 to 10 mg/L, especially for effluents with BOD concentrations below 10 mg/L. This is true for both average and maximum monthly TOC values. Table D-6 presents summary statistics for average monthly effluent TOC, for all data as well as data categorized according to the following BOD ranges: ≤5 mg/L, 5 to 10 mg/L, 10 to 20 mg/L and > 20 mg/L. As noted in Table D-6, four very low TOC values (≤ 0.5 mg/L) were judged to be anomalies and were therefore censored from the data for these statistics. The same summary statistics are presented for maximum monthly effluent TOC in Table D-7. In this context, the regressions of TOC on BOD values from effluent samples at all POTWs are quite reasonable, despite the low coefficients of determination. For average monthly effluent data, the regression of TOC on BOD is:

$$TOC_{avg} = 5.8740 + 0.2859\ BOD_{avg} \quad (r^2 = 0.245,\ 174\ df)$$

Table D-6. Summary statistics for POTW average monthly effluent TOC concentrations, categorized according to average monthly effluent BOD concentration

TOC_{avg}	BODavg level				
	≤5 mg/L	>5-10 mg/L	>10-20 mg/L	>20 mg/L	All levels
Mean	6.75	8.45	9.83	13.32	7.98
Median	6.17	8.15	8.70	13.25	6.90
Standard Deviation	2.47	2.41	5.26	5.79	3.76
5^{th} quantile	4.90	6.08	4.46	6.53	4.96
95^{th} quantile	10.00	10.58	18.40	21.8	14.52
n	98*	32*	34*	8	172*

*Four suspiciously-low TOC values were censored from the data for these statistics

Table D-7. Summary statistics for POTW maximum monthly effluent TOC concentrations, categorized according to maximum monthly effluent BOD concentration

TOCmax	BODmax level				
	≤5 mg/L	>5-10 mg/L	>10-20 mg/L	>20 mg/L	All levels
Mean	7.95	7.88	14.73	20.57	10.08
Median	7.30	7.85	11.00	20.60	7.90
Standard Deviation	3.28	2.74	15.97	7.75	7.60
5^{th} quantile	5.50	3.00	2.00	11.00	5.4
95^{th} quantile	10.94	12.00	25.55	35.3	23.6
n	164	72	30	35	301

Given the substantial limitations imposed by the data available from PCS, we believe that this regression gives reasonable estimates of TOC in POTW effluents. These are also probably the best available estimates of effluent TOC for dilution calculations to determine DOC concentrations for use in the BLM (for example, the probabilistic dilution framework incorporated in the BLM-Monte software [HydroQual, 2001]). As shown in Section D.3, effluent DOC concentrations can be reliably predicted from TOC values:

$$DOC = 0.7133\ TOC \qquad (r^2 = 0.8913,\ 25\ df)$$

It should be noted that the regressions presented here should not be applied to project water quality in natural receiving waters unimpacted by POTW effluent, because they are based solely on POTW effluent monitoring data. The characteristics of the constituents DOC, TOC, and BOD, as well as the relationships between them, may be quite dissimilar between natural waters and effluents.

D.5 References

Fadini, S.F., Jardim, W.F., J.R. Guimarães. 2004 Evaluation of organic load measurement techniques in a sewage and waste stabilisation pond. Journal of the Brazilian Chemical Society. 15(1): 131-135 (http://jbcs.sbq.org.br/jbcs/2004/vol15_n1/19-152-02.pdf).

HydroQual, Inc. 2001. BLM-Monte User's Guide, Version 2.0. HydroQual, Mahwah, NJ. October, 2001.

HydroQual, Inc. 2005. Biotic Ligand Model Windows Interface, Version 2.1.2, User's Guide and Reference Manual. HydroQual, Mahwah, NJ. June, 2005.