

## SPATIAL TRENDS IN WATER CHEMISTRY AND THE BIOTIC LIGAND MODEL

James N. Carleton\*

**ABSTRACT:** In a proof-of-concept exercise, a large database of spatially-referenced surface water monitoring data was examined to look for geographic trends in five chemical constituents that are key inputs to a model for predicting metal toxicity to aquatic organisms. Continuous prediction maps of concentration were generated using kriging to interpolate between site-median values measured at several thousand separate locations throughout the continental United States. Concentration surfaces generated this way were averaged over 8-digit USGS cataloging unit (HUC) polygons to produce block-averaged mean estimates of site-median concentrations. Pairwise comparisons indicated distinct trends between various HUC-averaged concentrations of different chemical constituents. Analyses of data from several hundred sampling locations where all five constituents had been measured revealed similar relationships between constituents. Principal components analyses 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 quantile maps for block-averaging is suggested as a possible approach for addressing spatial variability, in developing concentration values to use as model inputs, when site-specific monitoring data are lacking.

### BACKGROUND

The U.S. Environmental Protection Agency recently released proposed water quality criteria for copper (Cu). 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 Cu and render it biologically unavailable, or compete with Cu 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 ( $\text{Ca}^{2+}$ ), magnesium ( $\text{Mg}^{2+}$ ), sodium ( $\text{Na}^+$ ), sulfate ( $\text{SO}_4^{2-}$ ), potassium ( $\text{K}^+$ ), chloride ( $\text{Cl}^-$ ), dissolved organic carbon (DOC), and pH), model-predicted toxicity is most sensitive to five:  $\text{Ca}^{2+}$ , alk, pH,  $\text{Na}^+$ , 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 conservative (realistic but 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 body of water chemistry monitoring data collected in recent years by the US Geological Survey, and recorded in their NWIS (National Water Information System) database. The data includes monitoring information from several thousand separate surface water sampling locations throughout the United States. The primary objective of this analysis is to look for any obvious spatial trends in typical concentrations of the five most sensitive BLM inputs, 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 United States.

### Description of Data

Although NWIS contained data from 207153 sampling events at 13824 individual sampling locations in the continental U.S., all ten 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, 5900;  $\text{Ca}^{2+}$ , 10940; DOC, 3726;  $\text{Na}^+$ , 10424; pH, 11780. Numbers of sampling events at individual locations ranged from 1 to 2605, with a mean of 15, and a mode of one (*i.e.* most sites were sampled only once). Examination of the spatial distribution of numbers of sampling events per site (not shown)

\* Modeling Team Leader, EPA, Office of Water, Office of Science & Technology, ph: 202 566-0445, fax: 202 566-0409, carleton.jim@epa.gov

revealed that the most intensive sampling tended to occur in midwestern and western states. 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.

**DATA ANALYSIS**

Because, as mentioned, distributions of environmental sampling 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 of interest, the first step involved simply mapping the sampling locations as points, color-graded by median concentration. Figure 1, 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.

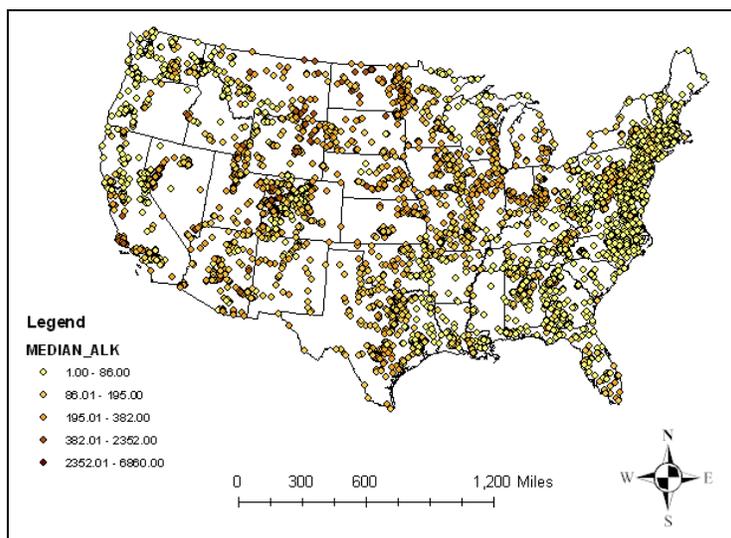


Figure 1. Median measured alkalinity (mg/L as CaCO<sub>3</sub>) at NWIS locations.

The next step in data visualization involved the calculation of median concentrations averaged over each 8-digit HUC containing sampled sites. These display essentially the same information as the point displays, but with a degree of smoothing and summarization provided by the spatial averaging process, to make visual interpretation of general trends easier (Figure 2).

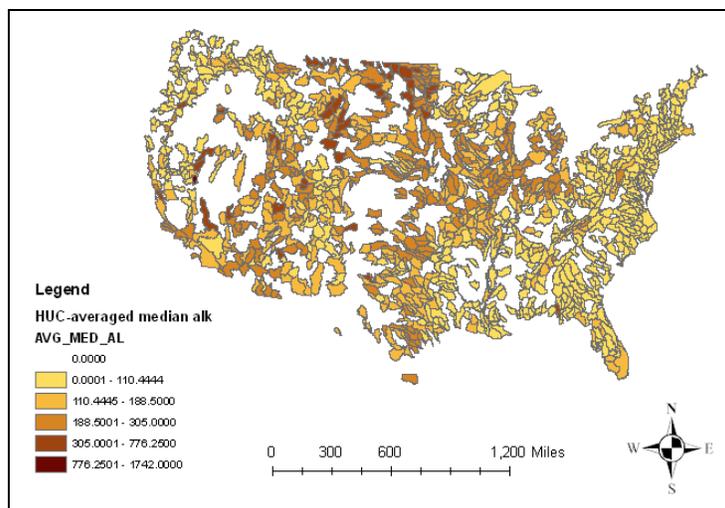


Figure 2. HUC-averaged mean median observed alkalinity in the continental U.S.

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 apparently roughly consistent with gross variations in median concentration (Figure 2). One problem with using HUCs for spatial aggregation is that not all HUCs contain NWIS sampling stations, as the blank areas in Figure 2 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 that lacked NWIS sites.

For each of the five key constituents, the Geostatistical Analyst extension in ESRI’s ArcGIS® software 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 professional 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.

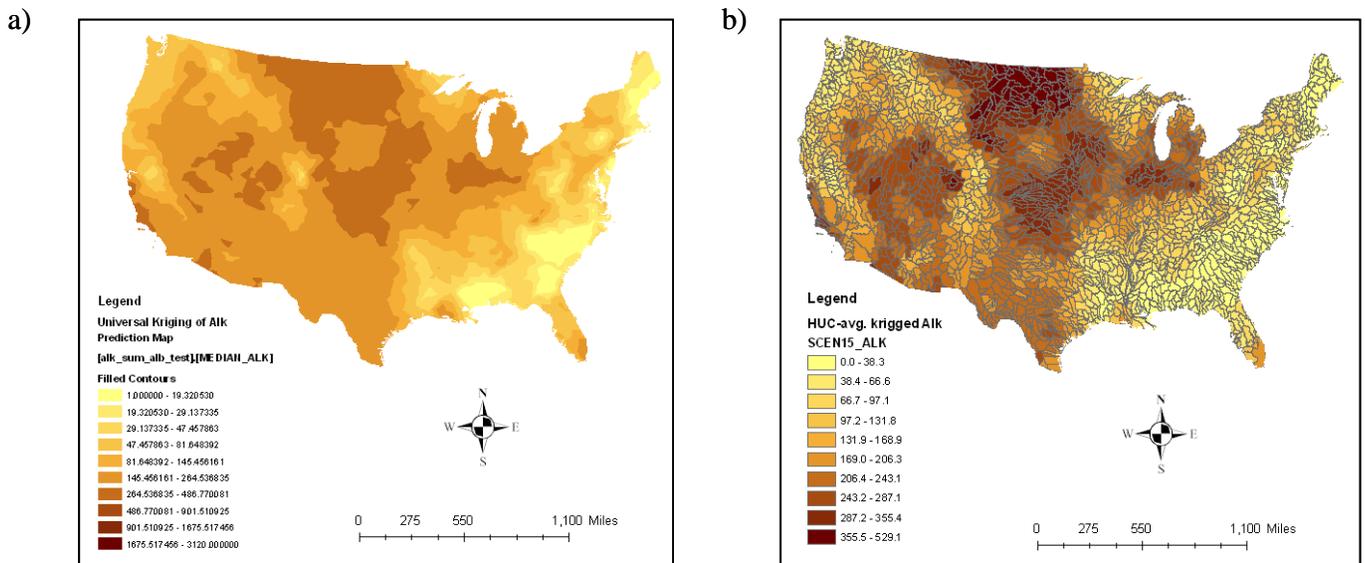


Figure 3. Kriging prediction maps a) median alkalinity, and b) same, averaged over 8-digit HUC polygons.

Prediction surface maps were generated for each constituent using the above sets of kriging options. Figure 3a, which displays the results for alkalinity, shows patterns that are generally consistent with those in the data (Figure 1). 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 3b shows the predicted values averaged over HUC polygons by using the Zonal Statistics function of the ArcGIS Spatial Analyst extension.

For HUCs containing NWIS sampling locations, linear regression plots of predicted vs. measured concentrations provided a check on the accuracy of the kriging predictions. R-squared values for the five constituents were: 0.537 (alk), 0.238 (Ca<sup>2+</sup>), 0.686 (DOC), 0.351 (Na<sup>+</sup>), and 0.139 (pH). In most cases, a handful of outliers appeared to be responsible for smaller-than-expected correlation coefficients.

A scatterplot matrix of cross-constituent comparisons revealed some interesting, non-random relationships between HUC-averaged concentrations (Figure 4a). 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 scatterplot matrix of this data (Figure 4b) to also be constructed. Similarities between the kinds of relationships in Figure 4a and 4b suggest that the predicted HUC-mean median values are reasonable.

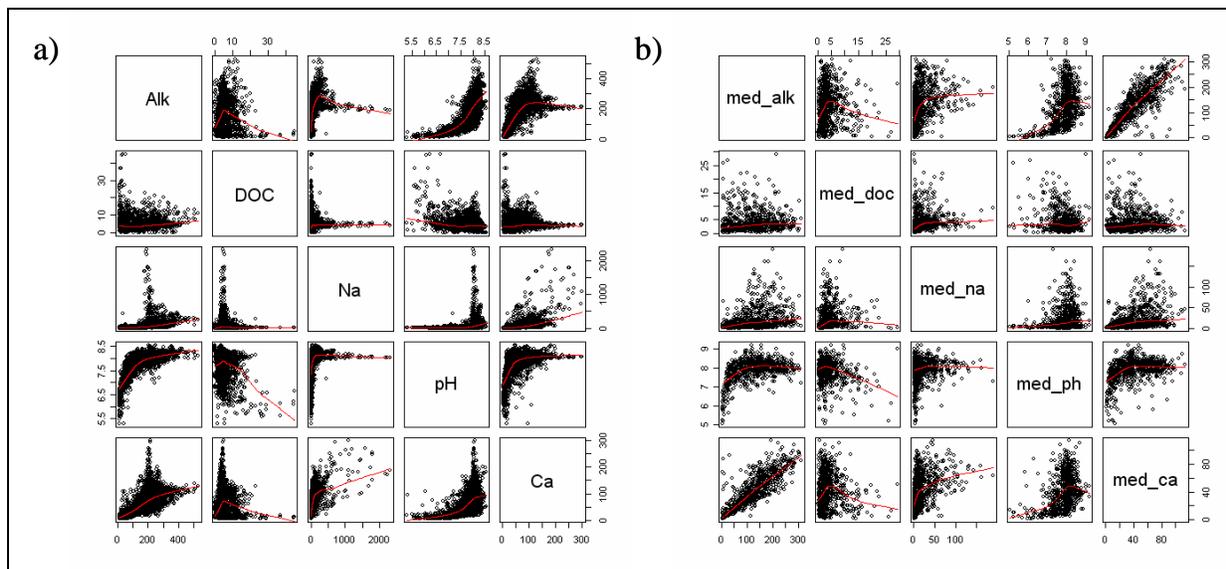


Figure 4. Scatterplot matrices of median concentrations from a) kriging predictions averaged over 8-digit HUC polygons; b) 772 separate monitored sites.

In addition to scatterplots, 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 1). Although not identical, the coefficients were generally similar between the two datasets, again suggesting that the kriging predictions are reasonable.

Table 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. The first component comprised 80 and 88% of the variance in the HUC-based and site-based analyses, respectively. This component loaded entirely onto alk, Na<sup>+</sup>, and Ca<sup>2+</sup> in both cases. For the HUCs, component 1 was primarily loaded on Na<sup>+</sup>, while for the sites, it primarily loaded on alk.

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 more conservative (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 5 displays the 25<sup>th</sup> 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 6) measured inside the same areas.

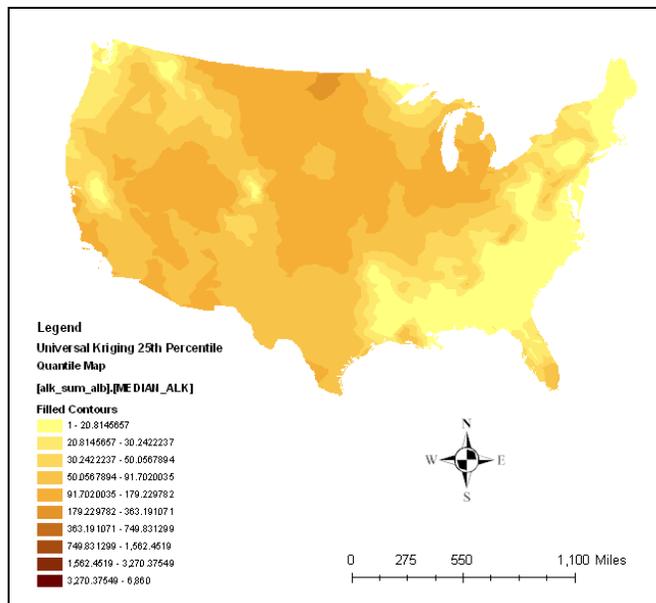


Figure 5. Kriging 25<sup>th</sup> percentile map of median alkalinity.

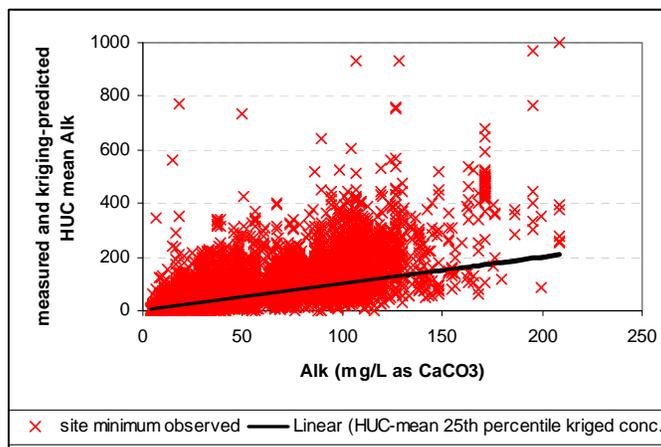


Figure 6. Comparison of observed site-minimum alkalinities with HUC-mean 25<sup>th</sup> percentile kriging-predicted values

DISCUSSION

Although geostatistical techniques such as kriging can be used to infer spatial patterns in water concentration over large areas, results should be interpreted cautiously, especially when considering locations where data is lacking, for example in mountainous regions of the western U.S. (Figure 2). The use of HUCs for spatial averaging of surface water concentrations is also not without conceptual difficulties. First, only about 45% of HUC8s 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 equally 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.

### **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: alkalinity, Na<sup>+</sup>, and Ca<sup>2+</sup>. Results suggest that block averaging of kriging predictions over irregularly spaced sampling points can provide estimates that preserve much of the character of the interrelationships between different measured entities. The use of suitable low-quantile kriging predictions is suggested as a way to estimate conservative (protective) water chemistry concentrations that could serve as regional default values for input to the Biotic Ligand Model for quantifying metal ecotoxicity.

### **Acknowledgements**

I am grateful to the following people for the assistance they provided on various aspects of this work: Luis Cruz, Charlie Delos, Christina Jarvis, and Lauren Wisniewski of EPA, and Professor Shunlin Liang of the Geography Department at the University of Maryland, College Park.

### **DISCLAIMER**

The content of this paper reflects the personal views of the author, and does not necessarily reflect official policy of the U.S. Environmental Protection Agency or any other agency.

### **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* 33(12):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.