Water Resources Research Institute of The University of North Carolina

Report No. 393

SPACE/TIME ESTIMATION OF FISH TISSUE MERCURY ALONG UNSAMPLED STREAMS IN EASTERN NORTH CAROLINA

By
Eric Money
Marc Serre

Department of Environmental Sciences and Engineering
Gillings School of Public Health
UNC-Chapel Hill
Chapel Hill, North Carolina

July 2010

UNC-WRRI-70242

SPACE/TIME ESTIMATION OF FISH TISSUE MERCURY ALONG UNSAMPLED STREAMS IN EASTERN NORTH CAROLINA

Eric S. Money and Marc L. Serre*

Univ. of North Carolina – Chapel Hill, Gillings School of Global Public Health, Dept. of Environmental Sciences and Engineering, Chapel Hill, NC 27599-7438

*Corresponding Author: marc_serre@unc.edu, 919-966-7014 (phone)

This report fulfills the requirements for a project completion report of the Water Resources Research Institute of The University of North Carolina. The authors are solely responsible for the content and completeness of the report.

WRRI Project No. 70242, July 2010

Acknowledgements

This work was supported by the North Carolina Water Resources Research Institute (grant 70242). This work was done in collaboration with D. Derek Aday and Dana K. Sackett while they received support from WRRI grant 70237. Kyle Messier assembled and wrote parts of this final report.

Abstract

Mercury in fish tissue is a major human health concern. Consumption of mercury-contaminated fish poses risks to the general population, including potentially serious developmental defects and neurological damage in young children. Therefore, it is important to accurately identify areas that have the potential for high levels of bioaccumulated mercury. However, due to time and resource constraints, it is difficult to adequately assess fish tissue mercury on a basin wide scale. We hypothesized that, given the nature of fish movement along streams, an analytical approach that takes into account distance traveled along these streams would improve the estimation accuracy for fish tissue mercury in unsampled streams. Therefore, we used the newly developed river-based Bayesian Maximum Entropy framework (river-BME) for modern space/time geostatistics to estimate fish tissue mercury at unsampled locations in the Cape Fear and Lumber Basins in eastern North Carolina. We also compared the space/time geostatistical estimation using river-BME to the more traditional Euclidean based approach, with and without the inclusion of a secondary variable. Results showed that this river-based approach reduced the estimation error of fish tissue mercury by more than 13%, and that the median estimate of fish tissue mercury exceeded the EPA action level of 0.3 ppm in more than 90% of river miles for the study domain.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Mercury in the Environment

Mercury (Hg) is an extremely reactive element that is present in air, water, and sediments because of both anthropogenic and natural sources (*1*). Mercury is emitted into the atmosphere primarily by volcanic activity (*2*) and the combustion of fossil fuels (*3*). Atmospheric mercury undergoes photochemical oxidation and is released through wet and dry deposition. In an aquatic system, inorganic mercury can be transformed into methylmercury by bacteria in the water and sediment (*4, 5, 6*). Methylmercury is the organic, neurotoxic form that is bioaccumulated by aquatic organisms, makes up the majority of total mercury in fish tissue (95-99%), and poses a significant risk to human health (*7*). Fish consumption is the primary vector for mercury movement from the environment into human populations (*8, 9*). Methylmercury can penetrate mammalian cells and alter cell division, which poses significant risk to developing fetuses and young children. Although methylmercury is primarily a neurotoxin, in high doses it can affect the kidneys and cardiovascular system (*10, 11*).

Many state and local agencies monitor fish tissue mercury (*Fish*Hg) and use this information to issue consumption advisories for particular areas and species of fish. In 2006, in the United States alone, 38% of total lake acreage and 26% of all river miles were subject to fish consumption advisories (*12*). However, assessing the spatiotemporal trends of *Fish*Hg on a larger scale, particularly when based on monitoring data, is a difficult task. In addition, high variability in biotic (e.g., fish size, age, and trophic position) and abiotic (e.g., pH, dissolved organic carbon) drivers of mercury movement in aquatic systems further complicates the development of site- and species-specific consumption advisories (*13-17*). Given these

difficulties, a robust method for estimating fish tissue mercury across broad spatial scales is necessary.

1.2 Geostatistical Estimation Methods

Geostatistical techniques, such as kriging, rely on the fact that many natural phenomenon exhibit spatial autocorrelation. Kriging methods construct a regional model of correlation to estimate variables at unsampled locations based on data from sampled locations (*18-20*). Cokriging uses not only the spatial correlation of a single variable, but also the correlations associated with other environmental variables. There have been numerous examples of cokriging for estimation of environmental variables ranging from soil salinity, suspended sediment, and rainfall, to regional stream quality (*21-24*). It is most beneficial where the primary variable is under-sampled with respect to the secondary variable. Generally, the inclusion of secondary information (i.e., secondary variables) results in more accurate local predictions than when considering a single variable alone (*21, 25*).

While kriging is a linear estimator, a more general approach to estimating at unsampled locations is the Bayesian Maximum Entropy (BME) method of modern space/time geostatistics (*26*). The BME approach is a non-Gaussian, non-linear estimator, which can incorporate non-Gaussian soft data and any non-linear relationship between primary and secondary variables. The BME method provides a rigorous mathematical framework to process a wide variety of knowledge bases. Site-specific knowledge includes both hard data (e.g., monitoring data measured without error) and soft data (e.g., data with associated measurement error). Soft data can also be generated for a primary variable by expressing the relationship between it and one or more secondary variables in terms of a probability distribution (*27*).

With respect to river networks, river-BME has been developed as an extension of the BME framework by incorporating river distances into the geostatistical estimation of water quality parameters on a basin-wide scale (*28, 29*). The traditional approach used in kriging and BME relies on the Euclidean distance for the calculation of the correlation models. Money et al. (28, 29) showed that for some water quality parameters river-BME provided more accurate estimation maps than the Euclidean based approach (Euclidean-BME). These studies, however, only examined ambient environmental parameters that were not inherently restricted by the configuration of the river network. Though estimation accuracy improved in these studies, we hypothesized that the estimation accuracy for variables affected by the configuration of the river network (i.e., mercury in fish whose movements are restricted to the river), could be further improved by using river distance because the distance between estimation points may be better represented by a river distance rather than a straight-line, or Euclidean, distance. This correlation has not previously been examined within a space/time geostatistical framework.

1.3 Secondary Variables Influencing the Bioaccumulation of Mercury

Even though the primary focus of this work was the investigation of estimation accuracy using river distances, it is also beneficial to show how river-BME can be used to incorporate any secondary variables of interest into the space/time estimation process. Previous studies have examined a variety of water quality variables and their relationship to *Fish*Hg. Many found pH to be a reliable predictor of *Fish*Hg (*30-32, 16*). Low pH (i.e., acidic water) has been shown to increase the amount of methylmercury released from sediments, increasing its bioavailability (*33, 34*). Oftentimes, measurements such as pH are more readily available than fish tissue samples; therefore, pH is a good candidate for generating soft data for *Fish*Hg in areas where

fish samples do not exist. Hence, we used pH as our first secondary variable to demonstrate how using a relevant secondary variable in river-BME leads to improved prediction of *Fish*Hg.

The inclusion of a secondary variable can also be used in the spatiotemporal analysis of environmental exposure-effect associations (*35*). For example, the link between water column mercury (*WC*Hg) and *Fish*Hg is not fully understood. The majority of mercury enters from outside the water body as a result of direct atmospheric deposition or indirect runoff and can become associated with sediments; however, some mercury may remain in the inorganic form in the water column, where it is directly transformed into methylmercury by bacteria. According to Southworth et al. (*36*), minute concentrations of aqueous mercury are capable of generating methylmercury at rates significant enough to warrant consumption advisories. It is therefore possible that fish exposure to *WC*Hg results in increased *Fish*Hg. Christakos and Serre (*35*) have proposed the following Physico-Epidemiologic Predictability (PEP) criterion to assess the exposure-effect relationship: An exposure-effect association between *WC*Hg and *Fish*Hg is quantitatively supported if the prediction of *Fish*Hg is improved by integrating soft data generated using *WC*Hg. Hence, we will use *WC*Hg as our other secondary variable to demonstrate how river-BME can be used to investigate a suspected exposure-effect association between an exposure variable and *Fish*Hg.

2. Materials and Methods

2.1 Data and Study Area

The areas under investigation were the Cape Fear and Lumber River Basins in eastern North Carolina (Figure 1). Both basins have active fish consumption advisories, and the entire Lumber Basin was listed as impaired in the state's 303(d) list of impaired waters as required by the Clean Water Act (*37*). The Lumber Basin is approximately 3300 square miles (including portions in South Carolina) and is primarily forested (60%) or agricultural (30%). The Cape Fear Basin, at 9300 square miles, is the largest basin in the state, and contains close to 20% of the total population, or around 2 million people (*38*, *39*).

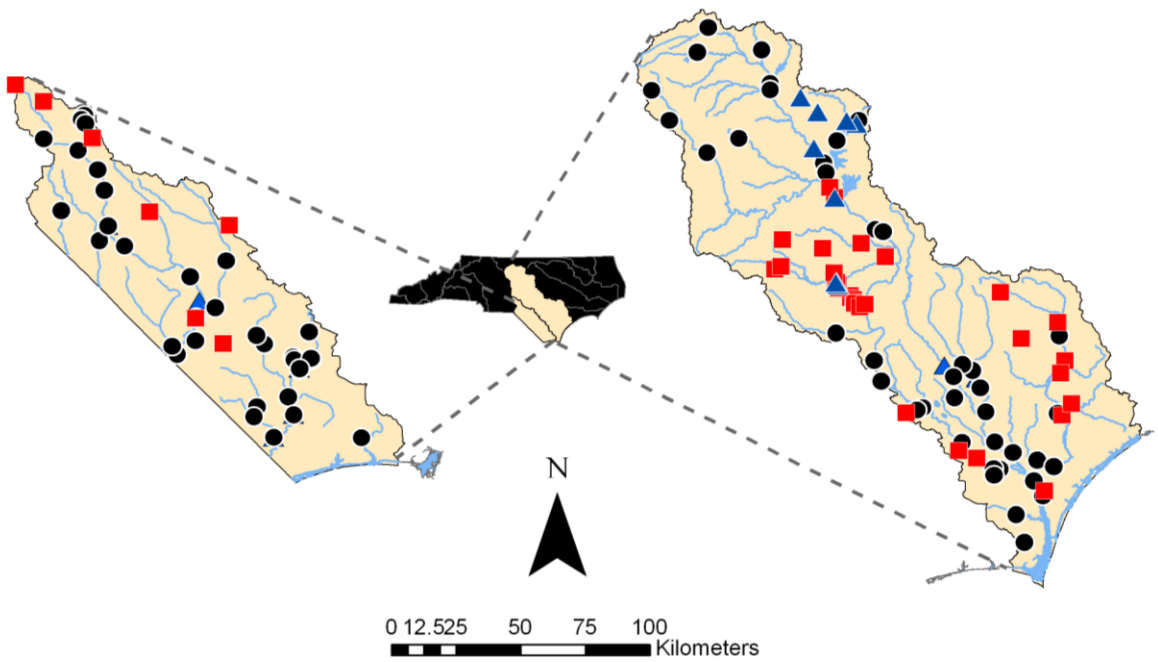

Figure 1: Lumber (Left) and Cape Fear (Right) Basins in North Carolina, with locations for *Fish*Hg (circles), pH (squares), and *WC*Hg (triangles).

Our database was compiled from several sources (Table 1). *Fish*Hg data were obtained from the North Carolina Department of Environment and Natural Resources (NCDENR) and through the North Carolina Division of Water Quality (NCDWQ) Fish Tissue Assessment Program. The database of *Fish*Hg and secondary variables was assembled by researchers at North Carolina State University, and a complete description of this database can be found in Sackett et al. (*16*). Only those data within the Cape Fear and Lumber Basins were used for the period 1990-2004. Collocated pH data and *Fish*Hg were obtained from this database and additional pH measurements were downloaded from the National Water Information System through the United States Geological Survey (NWIS-USGS). Surface water total mercury data were collected by NCDWQ as part of the Eastern Regional Mercury Study (*40, 41*) and combined with data downloaded from the NWIS. The few duplicate measurements (i.e., data collected at the same location on the same day) were averaged to a single value.

Table 1: Data Summary for mercury and pH in the Cape Fear and Lumber Basins, 1990-2004

Data Type	# of Locations	# of Independent Samples	# of Samples collocated with Fish Hg Samples
Fish Hg	75	1663	-
pH	33	356	143
Surface Water Hg*	7	80	35

2.2 Generation of Fish Tissue Mercury Soft Data from Multiple Secondary Variables

The availability of secondary variables provides an opportunity to generate soft data about *Fish*Hg. Money et al. (29) described a linear regression framework for generating soft

data from secondary water quality variables in the context of creating *E. coli* soft data from turbidity measurements. In this study, we use a similar approach to generate soft data for *Fish*Hg based on either pH or *WC*Hg, for reasons described previously.

To fully illustrate the potential effects of river distance on *Fish*Hg estimates, we found the common downstream outlet for the Cape Fear and Lumber basins, essentially treating them as one continuous network. There were a total of 143 points in the combined system where both *Fish*Hg and pH were measured. Using these collocated points, a simple regression analysis was performed using log-transformed data to create a relationship of predicted *Fish*Hg given pH (Eq. 1; R-squared = 0.17):

$$\log\text{-}\mathit{Fish}\mathrm{Hg} = -0.6\ \mathrm{pH} + 3.3 \qquad (1)$$

where log-*Fish*Hg is expressed in log-mg/kg (or ppm) and pH is in standard units. Using this relationship, the log-*Fish*Hg prediction error variance was calculated using the mean of the squared differences between predicted and measured log-*Fish*Hg for a series of given windows of pH values. Finally, for every space/time point where pH (but not necessarily log-*Fish*Hg) was measured, a Gaussian probability distribution function (PDF) was constructed for log-*Fish*Hg with a mean given by Equation (1) and a variance corresponding to the prediction error variance at the measured pH. This resulted in soft log-*Fish*Hg data of Gaussian probabilistic type at 356 space/time points. We should note that it was not our goal to fully model the pH-*Fish*Hg relationship, but rather to devise a simplistic relationship based on the available data and extract what information we could from that relationship to generate soft data. The soft data, in turn, takes into account this inherent uncertainty by way of the variance of the Gaussian PDF.

There were a total of 35 points where both *Fish*Hg and *WC*Hg were measured during the study period. Again, using these collocated points, we constructed a simple relationship to predict *Fish*Hg using log-transformed data (Eq. 2; R-squared = 0.03):

$$\log\text{-}Fish\text{Hg} = 0.25 \log\text{-}WC\text{Hg} - 3.40 \qquad (2)$$

where log-*Fish*Hg is expressed in log-mg/kg (or log-ppm) and log-*WC*Hg is expressed in log-ng/L. For every space/time location where *WC*Hg was measured, a Gaussian probability distribution function was constructed for log-*Fish*Hg, with a mean given by Equation (2) and variance corresponding to the prediction error variance at the measured log-*WC*Hg. This resulted in soft log-*Fish*Hg data of Gaussian type at an additional 80 space/time locations.

The models shown in Equations 1 and 2 are simplified expressions of a complex system that has many potential secondary variables. As discussed earlier, a variety of factors affect the concentration of mercury in fish tissue; however, the intent of this study was to primarily examine how the use of river distance affects the estimation accuracy of *Fish*Hg and how soft data can be incorporated into the analysis; therefore, a simplified model was most appropriate. We should note that we are only using Gaussian soft data in this analysis because it was the only type available; however, as described, BME allows for the incorporation of other types of non-Gaussian data when available.

2.3 Integrating Hard and Soft Data

The BME method of modern space/time geostatistics was used to integrate these hard and soft data and obtain statistical estimates of log-*Fish*Hg at un-monitored locations. Both

Euclidean-BME and river-BME procedures consist of defining the general knowledge (e.g., covariance) and site-specific knowledge (e.g., monitoring data), and integrating these two knowledge bases to obtain a posterior PDF characterizing log-*Fish*Hg at any point on the river.

The BME method (Euclidean-BME) is described fully in Christakos (*27*), while the BME*lib* numerical implementation is described in Serre et al. (*42*), Serre and Christakos (*43*) and Christakos et al. (*44*). The river-BME framework and *riverlib* extensions to BME*lib* are fully described in Money (*45*), along with several water quality applications (*28, 29*).

The *BMElib* package and *riverlib* extensions implement concepts of composite space/time analysis (e.g., composite space/time metrics and neighborhood search, non-separable space/time covariance models, etc.) that result in better geostatistical functions for linear space/time kriging than those provided by classical geostatistics software, where time is included as merely another spatial dimension (*44, 46*).

2.4 Space/Time Covariance Models That Use River Distances

As with previous water quality studies using river-BME (*28, 29*), a covariance model is selected that uses either a Euclidean or river distance. We restrict our model choice to the isotropic exponential covariance model since it has been shown to be permissible when using river distances (45, 47, *48*). Isotropic covariance models are only a function of the distance between points, so that points both upstream and downstream of a given point are equally correlated with that given point. Using this model, the covariance of log-*Fish*Hg between space/time points $\boldsymbol{p}=(\boldsymbol{s},t)$ and $\boldsymbol{p}'=(\boldsymbol{s}',t')$ is expressed as

$$\mathrm{cov}(\boldsymbol{p},\boldsymbol{p}') = c_0 \exp\left(\frac{-3h}{a_r}\right)\exp\left(\frac{-3\tau}{a_t}\right) \tag{3}$$

where t and t' are times, $h=d_\alpha(\boldsymbol{s},\boldsymbol{s}')$ and $\tau=|t-t'|$ are the spatial and temporal lags, respectively , and $d_\alpha(\boldsymbol{s},\boldsymbol{s}')=\alpha d_R(\boldsymbol{s},\boldsymbol{s}')+(\alpha-1)d_E(\boldsymbol{s},\boldsymbol{s}')$ is an α-weighted average of the Euclidean distance $d_E(\boldsymbol{s},\boldsymbol{s}')$ and the river distance $d_R(\boldsymbol{s},\boldsymbol{s}')$. In this study, we used either $\alpha=0$ (Euclidean distance) or $\alpha=1$ (river distance). For each value of α, the parameters (c_0, a_r, a_t) of the covariance model (3) were obtained using a least square fitting between the covariance function and experimental covariance values calculated from the hard log-*Fish*Hg data.

2.5 Comparing Euclidean and River Estimations

A comparison was made between estimations using the river distance described above, and estimation using the typical Euclidean distance, where soft data from measured pH and measured *WC*Hg were included in both cases. A cross-validation analysis was performed to calculate the Mean Square Error (MSE) of four different scenarios to determine the best model for estimating basin-wide log-*Fish*Hg (Table 2). Scenario 1 used the measured log-*Fish*Hg data with Euclidean-BME. Scenario 2 contained the same data as Scenario 1, except river-BME was used. Scenario 3 built upon Scenario 2 by adding in the pH data (incorporated as the soft Gaussian data constructed using Eq. 1). Scenario 4 built upon Scenario 2 by adding in Gaussian soft data from *WC*Hg using Equation 2. The method with the lowest MSE was then used in the assessment and estimation of *Fish*Hg for the Cape Fear and Lumber Basins.

Table 2: Cross-validation scenarios for *Fish*Hg estimations using river-BME and Euclidian-BME

Scenario	Metric Used	Hard Data Used	Soft Data Used
I	Euclidean	Measured log-*Fish*Hg	-
II	River	Measured log-*Fish*Hg	-
III	River	Measured log-*Fish*Hg	Gaussian from log-pH
IV	River	Measured log-*Fish*Hg	Gaussian from log-*WC*Hg

2.6 Estimation of Fish Tissue Hg

Using the selected scenario within the BME framework, we estimated log-*Fish*Hg at equidistant estimation points (i.e., distributed at a fixed interval of 0.1 km) along the combined Cape Fear and Lumber network. For each estimation point, we selected the hard and soft log-*Fish*Hg data situated in its local space/time neighborhood, and calculated the corresponding BME posterior PDF describing log-*Fish*Hg at that estimation point. The variance of the BME posterior PDF provided an assessment of the estimation uncertainty, while the back-log transform of the mean of the BME posterior PDF was used as an approximation of the median estimator for *Fish*Hg concentrations. This estimate was then used to produce chloropleth maps of estimated *Fish*Hg concentration and calculate the fraction of river miles that exceeded the specified action levels.

2.7 Assessment of Impaired River Miles

The fraction of river miles impaired at any given time was calculated by determining the fraction of equidistant estimation points that exceeded a given action level for *Fish*Hg. There are currently three different action levels for *Fish*Hg. The Food and Drug Administration (FDA) has determined a consumer action level of 1.0 ppm (or mg/kg) (*49*). The state of North Carolina has declared a more stringent action level of 0.4 ppm (*50*). In addition, the USEPA has set the most stringent mercury action level at 0.3 ppm (*51*). For the Cape Fear and Lumber Basins, the fraction of total river miles exceeding each of these threshold concentration values was calculated independently.

3. Results and Discussion

3.1 Covariance Analysis

The structure of the covariance model (with parameters c_0, a_r and a_t ; Eq. 3) provided insight into the variability and correlation between *Fish*Hg data points (Figure 2). The variance and temporal range were c_0 = 0.41 log-ppm^2 and a_t = 890 days, respectively, for both models. The variance corresponds to the covariance at a separation distance of zero (i.e. it is the sill of the covariance plot in Figure 2), which characterizes the amount of natural variability in *Fish*Hg (i.e. the inter sample variability of *Fish*Hg). The spatial range a_r, was 58 km for the Euclidean model and increased almost two-fold to 102 km for the river model. Hence the spatial covariance range (a_r) indicated that 95% of the correlation in *Fish*Hg was lost after about 58 km across land, compared to 102 km along the river. This suggested that by accounting for river distance, *Fish*Hg was spatially correlated over longer distances than if the constraints of the river network were not taken into account. This makes sense physically, given that fish are inherently restricted to movement pathways that follow the river network configuration. Conversely, if a Euclidean distance was used, the correlation between measurements of *Fish*Hg can be lost over a short distance because fish do not travel across land. Temporally, *Fish*Hg may remain highly correlated for a period of about 2-3 years. This is understandable given that bioaccumulated mercury may change gradually over time, depending upon the characteristics of the water body and fish community.

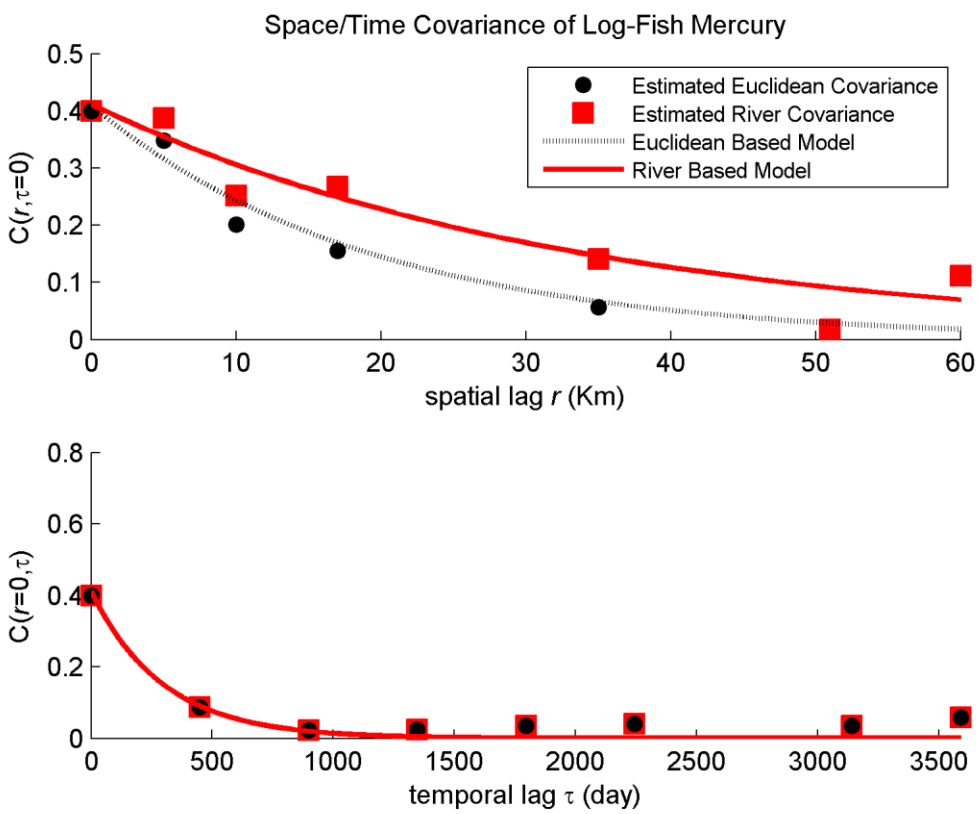

Figure 2: Spatial (top) and temporal (bottom) covariance of log-FishHg in the Cape Fear and Lumber Basins, North Carolina. Experimental covariance values estimated from data are shown with markers, while the covariance models obtained by fitting Eq. 3 to the markers are shown with lines. The covariance was calculated and modeled using both a Euclidean distance (dashed line) and river distance (plain line).

3.2 Cross-validation Analysis

The cross-validation analysis resulted in mean square errors of: MSE_1=0.3040 (log-ppm^2) for Scenario 1, MSE_2 = 0.2637 (log-ppm^2) for Scenario 2, MSE_3 = 0.2556 (log-ppm^2) for Scenario 3, and MSE_4 = 0.2555 (log-ppm^2) for Scenario 4.

Using river-BME over Euclidean-BME, with only hard data, reduced estimation error by $(MSE_1-MSE_2)/MSE_1=13\%$. This was a larger reduction than that obtained in previous studies that examined river-BME for DO (*28*) and for *E. coli* (*29*). Those studies resulted in a 10%-11% decrease in error. This suggests that river-BME is particularly useful for parameters, such as fish data, that are known to be restricted by the river network configuration and affected by the ability of fish to move both upstream and downstream.

The incorporation of pH as a secondary variable in river-BME further reduced the MSE by $(MSE_2-MSE_3)/MSE_2=3\%$. Even though this is a smaller reduction than that seen from soft data in the *E. coli* study by Money et al. (*29*), it is significant because fewer soft data points were available in this study. In the *E. coli*/turbidity study there were over 700 additional soft data points added to the analysis, and covering a much smaller land area. In this study, there were only ~300 additional soft data derived from pH. Hence, the estimation error may be even further reduced if more data points can be included as secondary variables.

The incorporation of *WC*Hg in river-BME also reduced the MSE by $(MSE_2-MSE_4)/MSE_2=3\%$. This result meets the Physico-Epidemiologic Predictability (PEP) criterion, which simply states that a reduction of MSE provides evidence supporting an exposure-effect association between *WC*HG and *Fish*Hg (*35*). This result is significant because the MSE reduction from *WC*Hg is as strong as that obtained from pH, even though there were only ~30 additional soft data points derived from *WC*Hg, while there were ~300 soft data points derived from pH.

Overall, when using river-BME with pH as a secondary variable there was a $(MSE_1-MSE_3)/MSE_1=16\%$ reduction in estimation error as compared to Euclidean-BME without soft data. This suggests that accounting for the hydrogeography of the system as well as variables

that affect the bioaccumulation of mercury will result in more accurate estimations of fish tissue mercury along unsampled streams.

3.3 Assessment of Fish Tissue Mercury

Using river-BME with pH as the secondary variable (Scenario 3), we obtained median estimates of *Fish*Hg concentrations calculated every 180 days over the study period between 1990-2004 (Figure 3 below).

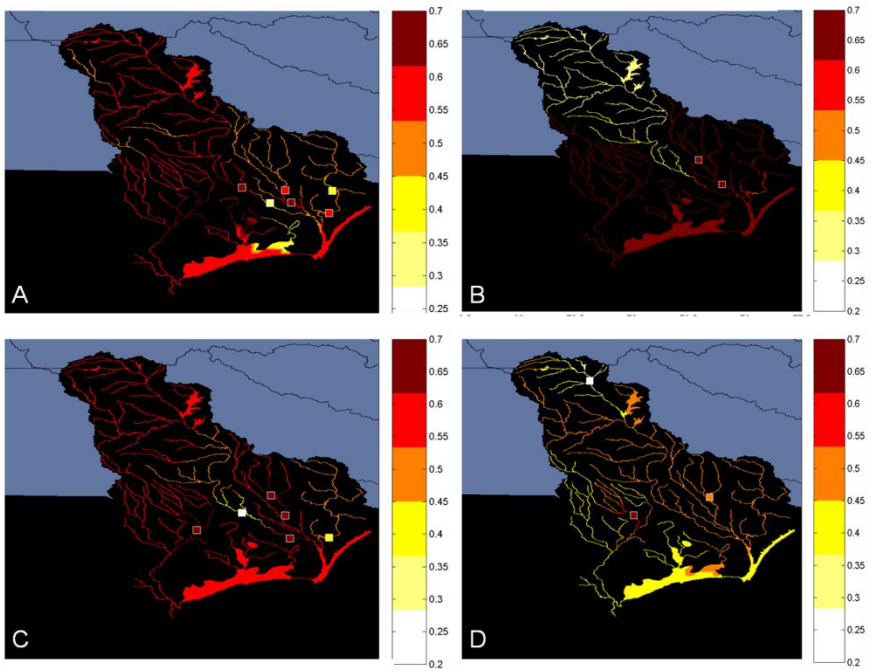

Figure 3: river-BME Fish Tissue Mercury estimations (ppm) in the Cape Fear and Lumber Basins on July 23, 1995 (A); July 2, 1999 (B); June 26, 2000 (C); and May 13, 2004 (D). Squares indicate locations of actual fish tissue measurements.

A movie showing the *Fish*Hg levels estimated across the study area for each of these times can be viewed in supplementary material. In the combined basin, the median estimate of *Fish*Hg exceeded the most stringent action level of 0.3 ppm in more than 90% of river miles for a majority of the study period. There were more fluctuations in the percent of impaired river miles when the action level was increased to the current North Carolina level of 0.4 ppm; however, over 50% of river miles still had median estimates of *Fish*Hg exceeding 0.4 ppm for almost the entire study period. In addition, during the years 1990-1994, between 1-4% of river miles had a median estimate of *Fish*Hg above even the most lenient action level of 1.0 ppm set by the FDA. Another small peak was seen in 1998; however, at least with respect to the FDA, the majority of waters remained below this action level. No river miles had a median estimate of *Fish*Hg exceeding the FDA action level since 1999, according to these results. Generally, the Lumber Basin exhibited higher potential, both spatially and temporally, for contaminated fish. Our modeling results were consistent with current trends; the entire Lumber Basin has been listed as impaired for fish tissue mercury since at least 2002 (*37*).

3.4 Monitoring Assessment for Fish Tissue Hg

One of the benefits of using this type of analysis is the ability to pinpoint areas with insufficient data and provide guidance on future monitoring strategies. One way to determine areas where information is lacking is to examine a combination of the spatial distribution of measured values with a measure of the uncertainty at particular estimation points. In BME, this uncertainty is measured via the variance of the posterior distribution at each estimation point. The larger the variance, the more uncertain the estimation. Higher uncertainty, in this instance, can also be interpreted as a lack of information. Therefore, areas that have a high *Fish*Hg

estimation and a low variance would be good candidates for additional monitoring if no actual samples exist in the vicinity.

A map of the estimation variance can be created similar to the BME estimation map, except the values represent the variance in *Fish*Hg at each point rather than the BME estimate of *Fish*Hg. Figure 4 shows a typical BME variance map for one day during the study period.

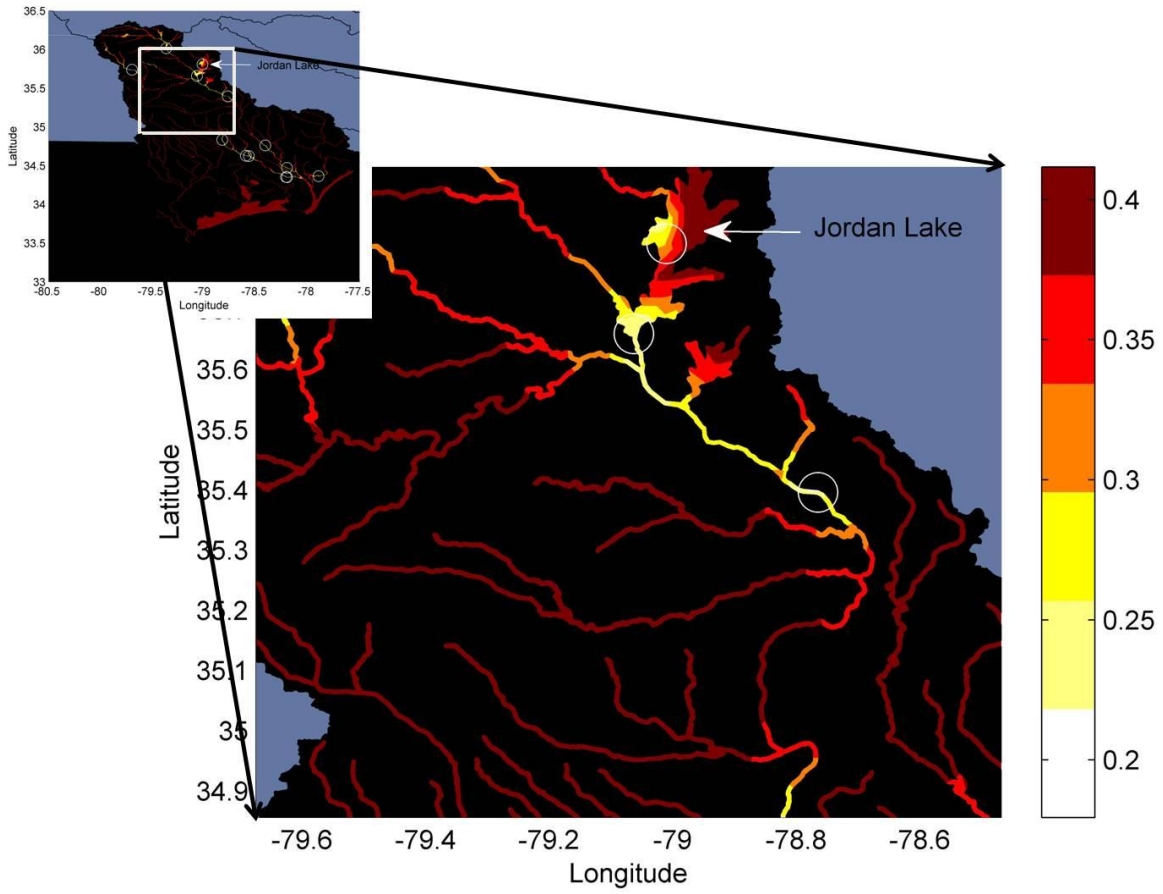

Figure 4: river-BME estimation variance (ppm^2) on 07/07/98. The white box in the upper-left panel represents the zoomed in portion of the map. White circles indicate actual measurements of *Fish*Hg +/- 45 days from the estimation date.

As we can see from Figure 4, areas in yellow-to-white have lower uncertainty associated with the estimation of *Fish*Hg. These areas typically correspond to areas surrounding actual samples. Those areas represented by darker colors, represent higher uncertainty in the estimation of *Fish*Hg, and, therefore, potential areas of insufficient data. Additionally, we can examine this map in relation to the BME river estimation map to determine areas where estimates of *Fish*Hg are high (relative to some threshold) and the variance is low in locations where no samples have been taken.

A preliminary examination suggests that areas in the northwest and southwest quadrant of the study area (primarily in the Lumber Basin) may require additional monitoring to accurately assess the *Fish*Hg in those areas. Further evaluation of the sample data used in the study area also reveals that > 80% of the samples were taken within 30-km of the downstream outlet of the basins (Figure 5), illustrating the need for further monitoring in upstream reaches of these basins.

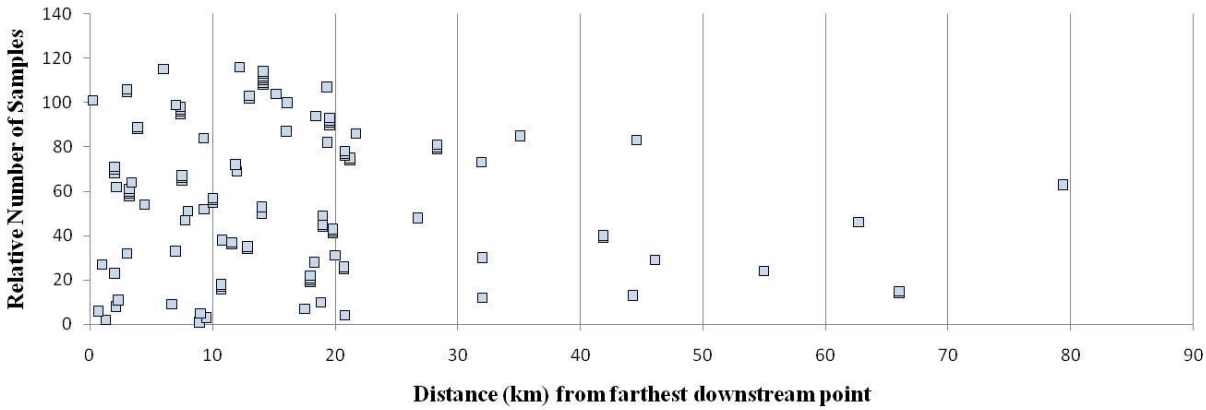

Figure 5: Number of *Fish*Hg samples collected at each monitoring site as a function of upstream river distance from farthest downstream point (i.e. the downstream outlet) of the combined river network. Each square represents a sampling site where one or more samples were collected

Future work should examine each of the maps for the study period to pinpoint exact locations where additional monitoring may provide useful information. Techniques, such as those used by LoBuglio et al. (*52*), could be used to evaluate different monitoring scenarios for optimal knowledge gain and cost-benefit.

4. Summary and Conclusions

In sum, our study adds to the recent literature emphasizing the importance of using river distances to estimate environmental quality parameters along river networks (*48*), and to our knowledge, it is the first study to do so in a space/time river-BME framework applied to a fish environmental parameter. First, we find that using a river geostatistical framework is particularly relevant for a fish environmental parameter because fish have the ability to swim both down and up stream, which is consistent with the way the river distance $d_R(\boldsymbol{s},\boldsymbol{s}')$ is calculated between any points $\boldsymbol{s}$ and $\boldsymbol{s}'$. Second, we demonstrated the ability of river-BME to efficiently integrate a secondary variable such as pH, which resulted in estimation maps that were overall 16% more accurate than maps obtained with the classical estimation approach ignoring river distances and secondary variables. Third, an exposure-effect association between *WC*Hg and *Fish*Hg was quantitatively supported by our data analysis, even though the data available for *WC*Hg were very limited.

River-BME provides a good framework for further decreases in estimation error as more data on these secondary variables become available. Future work will consider the joint integration of secondary data from multiple sources, which should lead to further reductions in estimation error. *Fish*Hg can vary significantly due to a suite of biotic and abiotic factors. Therefore, distinguishing between the natural spatiotemporal trends and species specific trends can be difficult if the monitoring data is heterogeneous, as in this study. Wente (*53*) described an interesting approach to this problem by developing a statistical model for distinguishing trends in *Fish*Hg concentration using a combination of covariance and multiple linear regression. This model serves as the basis for the Environmental Mercury Mapping, Modeling, and Analysis

(EMMMA) program (*54*). Therefore, one way to address this problem would be to integrate EMMMA results into the river-BME analysis. Hence, incorporating additional information about species type and habitat patterns, as well as DOC, sediment data, and other variables could result in even more accurate maps. Overall, the framework developed in this study is a good starting point and can aid environmental managers in the identification of important bioaccumulation factors and areas where sampling resources and advisories can be efficiently targeted.

This study examines a combination of knowledge sources that may improve the estimation of fish tissue mercury concentrations in two large river basins in North Carolina. Estimation maps were produced that were on average 16% more accurate when accounting for river distance and the secondary variables pH and water column total mercury. Both secondary variables contributed to an overall decrease in estimation error, albeit small due to the limited amount of data points available for these secondary variables. The use of river-BME in this study contributed 13% of the reduction in estimation error and provides a good framework for further decreases in estimation error with the addition of other secondary factors in future work. Soft data from pH contributed as much error reduction as soft data from surface water mercury. Generally pH data is much more reliable and easier to measure, therefore state and local agencies may consider using pH measurements to aid in the assessment of fish tissue mercury in areas where samples may be scarce. However, this work demonstrates that water column total mercury data, when available, can provide a valuable alternate source of information to estimate fish tissue mercury levels. Furthermore, incorporating additional information about species type and habitat patterns could result in even more accurate maps. Overall, the framework developed

in this work is a good starting point and can aid environmental managers in identifying important bioaccumulation factors and areas where sampling and advisory resources can be targeted.

5. Recommendations

These results will benefit state planners interested in resource allocation and in the identification of impaired stream segments for federal classification purposes. Several recommendations can be made to highlight the utility of this analysis and make improvements for future guidance. First, we recommend that a stakeholder workshop be developed to inform the state environmental community on the methodology and application of the project outcomes. This workshop could involve the researchers of this study, along with the North Carolina Dept. of the Environment and Natural Resources, the North Carolina Division of Public Health, and others who manage or are affected by fish consumption advisories. The workshop would examine data and analysis needs as well as ways to disseminate this information to other stakeholders, including the public.

Second, we recommend future analysis to integrate EMMMA (Environmental Mercury, Mapping, and Modeling Analysis) results into the river-BME estimations, thereby incorporating additional information about species type and habitat patterns, as well as DOC, sediment data, and other variables that could result in even more accurate maps.

Finally, we recommend that future monitoring strategies should take into account the cost of monitoring vs. the benefit of knowledge gained from a particular location. One use of the analysis presented here is to identify target areas where sampling will be most beneficial and identify 'hot spots' of potentially high fish tissue mercury levels. This information can be used to direct the sampling of not only fish tissue mercury, but also other, more routine, water quality parameters that can act as sufficient indicators of bioaccumulation potential.

REFERENCES

(1) Morel, F.M.M.; Kraepiel, A.M.L.; Amyot, M. The chemical cycle and bioaccumulation of mercury. *Annual Review of Ecology and Systematics*. **1998**, *29*, 543-566.

(2) Nriagu, J.; Becker, C. Volcanic emissions of mercury to the atmosphere: global and regional inventories. *Sci Total Environ*. **2003**, *304*, 3-12.

(3) Pacyna, E.G.; Pacyna, J.M.; Steenhuisen, F.; Wilson, S. Global anthropogenic mercury emission inventory for 2000. *Atmospheric Environment.* **2006**, *40*, 4048-4063.

(4) Gilmour, C.C.; Henry, E.A.; Mitchell, R. Sulfate stimulation of mercury methylation in fresh-water sediments. *Environ Sci Technol.* **1992**, *26*, 2281-2287.

(5) Lin, C.C.; Jay, J.A. Mercury methylation by planktonic and biofilm cultures of Desulfovibrio desulfuricans. *Environ Sci Technol.* **2007**, *41*, 6691-6697.

(6) Ullrich, S.M.; Tanton, T.W.; Abdrashitova, S.A. Mercury in the aquatic environment: a review of factors affecting methylation. *Critical Reviews in Environmental Science and Technology*. **2001**, 31:241.

(7) Bloom, N.S. On the chemical form of mercury in edible fish and marine invertebrate tissue. *Can J Fish Aquat Sci*. **1992**, *49*, 1010-1017.

(8) Clarkson, T.W. The three modern faces of mercury. *Environ Health Perspect*. **2002**, *110*, 11-23.

(9) Scheuhammer, A.M.; Meyer, M.W.; Sandheinrich, M.B.; Murray, M.W. Effects of environmental methylmercury on the health of wild birds, mammals, and fish. *Ambio*. **2007**, *36*, 12-18.

(10) Chan, H.M.; Egeland, G.M. Fish consumption, mercury exposure, and heart diseases. *Nutrition Reviews*. **2004**, *62*, 68-72.

(11) Mergler, D.; Anderson, H.A.; Hing Man Chan, L.; Mahaffey, K.R.; Murray, M.; Sakamoto, M.; et al. Methylmercury exposure and health effects in humans: a worldwide concern. *Ambio*. **2007**, *36*, 3-11.

(12) U.S. Environmental Protection Agency. *2005/2006 National listing of fish advisories*. USEPA: Washington, DC, 2007, EPA-823-F-07-003.

(13) Huckabee, J.W.; Elwood, J.W.; Hildebrand, S.G. Accumulation of mercury in freshwater biota. In *Biogeochemistry of Mercury in the Environment*; Nriagu, J.O. ed. Elsevier/North-Holland Biomedical Press: New York, NY, 1979, 277-302.

(14) MacCrimmon, H.R.; Wren, C.D.; Gots, B.L. Mercury uptake by lake trout, *Salvelinus namaycush*, relative to age growth, and diet in Tadenac Lake with comparative data from other Precambrian Shield Lakes. *Canadian Journal of Fisheries and Aquatic Sciences*. **1983,** *40*, 114-120.

(15) Cope, W.G.; Wiener, J.G., Rada, R.G. Mercury accumulation in yellow perch in Wisconsin seepage lakes – relation to lake characteristics. *Environ. Tox. Chem*. **1990**, *9*, 931-940.

(16) Weiner, J.G.; Spry, D.J. Toxicological significance of mercury in freshwater fish. In *Environmental Contaminants in Wildlife – Interpreting Tissue Concentrations*; Beyer, W.N.; Heinz, G.H.; Redmon-Norwood, A.W. eds. Lewis Publishing: Boca Raton, FL, 1996, 297-339.

(17) Sackett, D.K.; Aday, D.D.; Rice, J.A.; Cope, W.G. A state-wide assessment of mercury dynamics in North Carolina waterbodies and fish. *Transactions of the American Fisheries Society*. 2009, In Press.

(18) Delhomme, J.P. Kriging in the Hydrosciences. *Advances in Water Res*. **1978**, *1* (5), 251-266.

(19) Cressie, N. The origins of kriging. *Math. Geol*. **1990**, *22* (3), 239-252.

(20) Stein, M.L. *Interpolation of spatial data: Some theory for kriging*. Springer: New York, NY, 1999, 247p.

(21) Darwish, Kh.M.; Kotb, M.M.; Ali, R. Mapping soil salinity using collocated cokriging in Bayariya, Oasis, Egyp. In *Proceedings of the 5th International Symposium on Spatial Data Quality*. ITC Enschede: The Netherlands, 2007.

(22) Li, Z.; Zhang, Y-K; Schilling, K.; Skopec, M. Cokriging estimation of daily suspended sediment loads. *J Hydrol*. **2006**, *327*, 389-398.

(23) Seo, D.J.; Krajewski, W.F.; Azimi-Zonooz, A.; Bowles, D.S. Stochastic interpolation of rainfall data from rain gauges and radar using cokriging. *Water Resour. Res*. **1990**, *26* (5), 915-924.

(24) Jager, H.I.; Sale, M.J.; Schmoyer, R.L. Cokriging to assess regional stream quality in the southern blue ridge province. *Water Resour. Res*. **1990**, *26* (7), 1401-1412.

(25) Goovaerts, P. *Geostatistics for natural resources evaluation*. Oxford University Press: London, 1997.

(26) Christakos, G.; Li, X.Y. Bayesian maximum entropy analysis and mapping: A farewell to kriging estimators? *Math. Geol*. **1998**, *30*, 435-462.

(27) Christakos, G. A Bayesian/maximum-entropy view on the spatial estimation problem. *Math. Geol*. **1990**, *22* (7), 763-776.

(28) Money, E.; Carter, G.; Serre, M.L. Using river distances in the space/time estimation of dissolved oxygen along two impaired river networks in New Jersey. *Water Research*. **2009a**, *43* (7), 1948-1958.

(29) Money, E.S.; Carter, G.; Serre, M.L. Modern space/time geostatistics using river distances: Data integration of turbidity and *E. coli* measurements to assess fecal contamination along the Raritan river in New Jersey. *Environ. Sci. Technol*. **2009b**, *43* (10), 3736-3742.

(30) Grieb, T.M.; Driscoll, C.T.; Gloss, S.P.; Shofield, C.S.; Bowie, G.L.; Porcella, D.B. Factors affecting mercury accumulation in fish in the upper Michigan peninsula. *Environ. Tox. Chem*. **1990**, *9*, 919-930.

(31) Rose, J.; Hutcheson, M.S.; West, C.W.; Pancorbo, O.; Hulme, K.; Cooperman, A.; DeCesare, G.; Isaac, R.; Screpetis, A. Fish mercury distribution in Massachusetts, USA lakes. *Environ. Tox. Chem*. **1999**, *18* (7), 1370-1379.

(32) Song, S.S.; Warren-Hick, J.; Keating, D.; Moore, R.J.; Teed, R.S. A predictive model of mercury in stream fish. *Environ. Sci. Technol*. **2001**, *35* (5), 941-947.

(33) Ullrich, S.M.; Tanton, T.W.; Abdrashitova, S.A. Mercury in the aquatic environment: A review of factors affecting methylation. *Critical Reviews in Environmental Science and Technology*. **2001,** *31* (3), 241-293.

(34) Miller, D.R.; Akagi, H. pH affects mercury distribution, not methylation. *Ecotoxicol Environ. Saf.* **1979**, *3* (1), 36-38.

(35) Christakos, G.; Serre, M.L. Spatiotemporal study of environmental exposure-health effect associations. *Journal of Exposure Analysis and Environmental Epidemiology*. **2000**, *10* (2), 168-187.

(36) Southworth, G.R.; Peterson, M.J.; Bogle, M.A. Bioaccumulation factors for mercury in stream fish. *Environmental Practice*. **2004,** *6*, 135-143.

(37) N.C. Dept. of Environment and Natural Resources. *North Carolina Water Quality Assessment and Impaired Waters List (2006 Integrated 303(d) and 303(b) reports).* NCDENR: Raleigh, NC, 2006.

(38) NCDENR. *Cape Fear basinwide assessment report*. NCDENR: Raleigh, NC, 2004.

(39) NCDENR. *Lumber river basinwide assessment report*. NCDENR: Raleigh, NC, 2007.

(40) N.C. Division of Water Quality. *2002/2003 Eastern Regional Mercury Study summary*. NCDWQ: Raleigh, NC, 2004.

(41) NCDWQ. *Eastern Regional Mercury Study extension report*. NCDWQ: Raleigh, NC, 2007.

(42) Serre, M.L.; Bogaert, P.; Christakos, G. Computational investigations of Bayesian maximum entropy spatiotemporal mapping. In *Proceedings of the 4th Annual Conference of the International Association for Mathematical Geology*; Buccianti, A.; Nardi, G.; Potenza, R.; De Frede Editore: Naples, Italy, 1998, 1, 117-122.

(43) Serre, M.L.; Christakos, G. Modern geostatistics: Computational BME in light of uncertain physical knowledge – the equus beds study. *Stoch. Envrion. Res. Risk Assess*. **1999**, *13* (1), 1-26.

(44) Christakos, G.; Bogaert, P.; Serre, M.L. *Temporal GIS: Advanced Functions for Field Based Applications*; Springer: New York, 2002, 217 p.

(45) Money, E.S. Modern Space/Time Geostatistics Using River Distances: Theory and Applications for Water Quality Estimation and Mapping. Ph.D. Dissertation, University of North Carolina at Chapel Hill, Chapel Hill, NC, 2009.

(46) *BMElib*, version 2.0b; BME Laboratory: Chapel Hill, NC, 2008 (accessed January 5, 2008).

(47) Ver Hoef, J.M.; Peterson, E.E.; Theobold, D.; Spatial statistical models that use flow and stream distance. *Environ. Ecol. Stat*. **2006**, *13*, 449-464.

(48) Ver Hoef, J.; Peterson, E. A moving average approach for spatial statistical models of stream networks. *Journal of the American Statistical Association*. In press.

(49) Food and Drug Administration. *FDA Consumer Advisory*; FDA: Washington, DC, 2001.

(50) Williams, L.K. *Health effects of methylmercury and North Carolina's advice on eating fish.* N.C. Dept. of Health and Human Services: Raleigh, NC, 2006.

(51) USEPA. *Notice of Availability of Water Quality Criterion for the Protection of Human Health: Methylmercury*; USEPA: Washington, DC, 2001, Federal Register 66 (5), 1334-1359.

(52) LoBuglio, J.N.; Characklis, G.W.; Serre, M.L. Cost-effective water quality assessment through the integration of monitoring data and modeling results. Water Resources Research. 2007, 43, W03435, doi:10.1029/2006WR005020.

(53) Wente, S.P.; A statistical model and national data set for partitioning fish-tissue mercury concentration variation between spatiotemporal and sample characteristic effects. *U.S. Geological Survey Scientific Investigation Reports*, *USGS2004-5199*. 2004, 15 p.

(54) Hearn, P.P.; Wente, S.P.; Donato, D.I.; Aguinaldo, J.J. EMMMA: A web-based system for environmental mercury mapping, modeling, and analysis. *U.S. Geological Survey Open File Reports, USGS2006-1086*. 2006, 13 p.

APPENDIX A: Supplementary Material

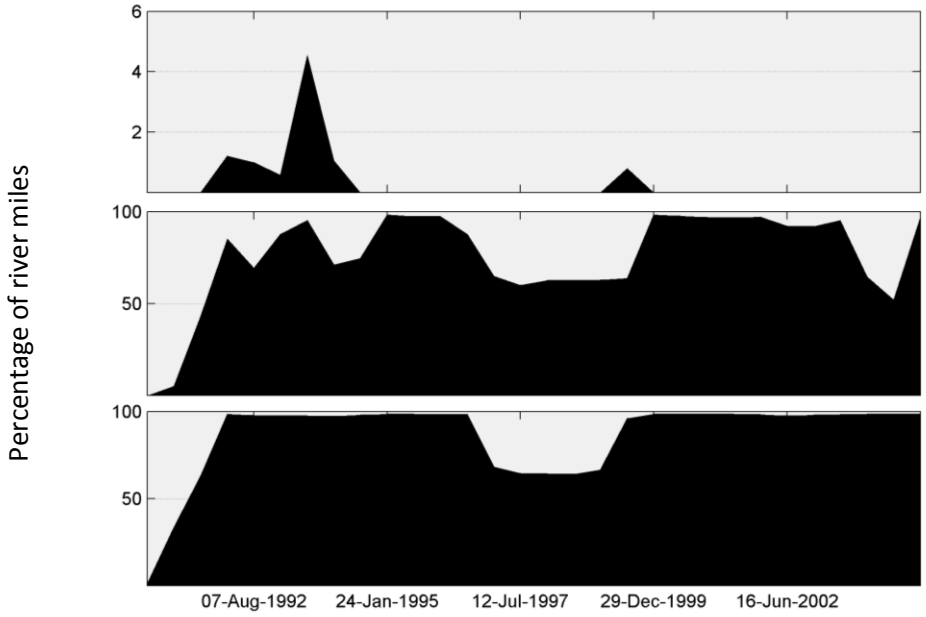

Figure S1: Percentage of river miles with fish tissue mercury median estimate exceeding Mercury Action Levels set by the FDA (top; 1.0ppm), North Carolina (middle; 0.4ppm), and the EPA (bottom; 0.3ppm).

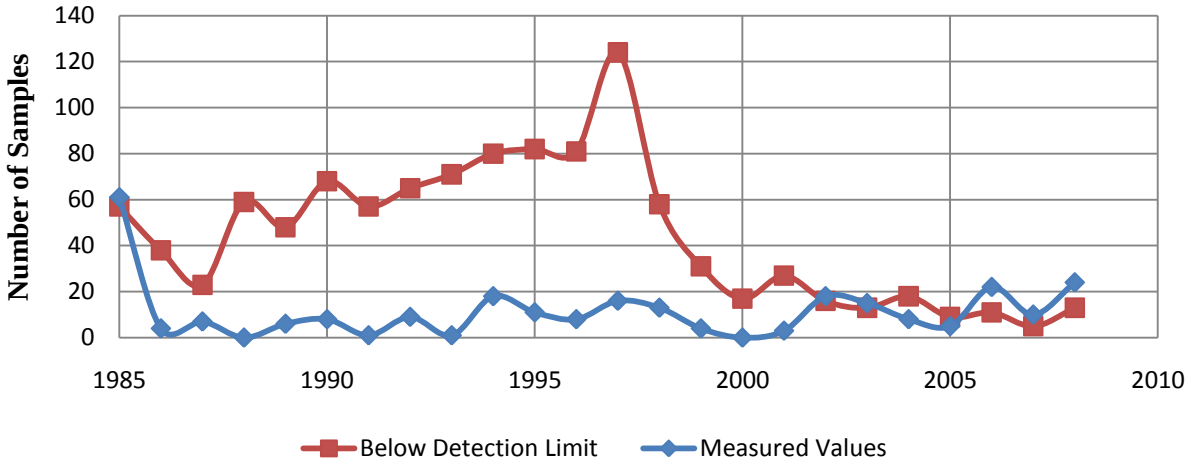

Figure S2: Temporal distribution of surface water Hg measurements for North Carolina over the study period. Squares indicate the number of samples recorded as 'BDL'or below the detection limit of the method used; Diamonds indicate the number of actual measured values.

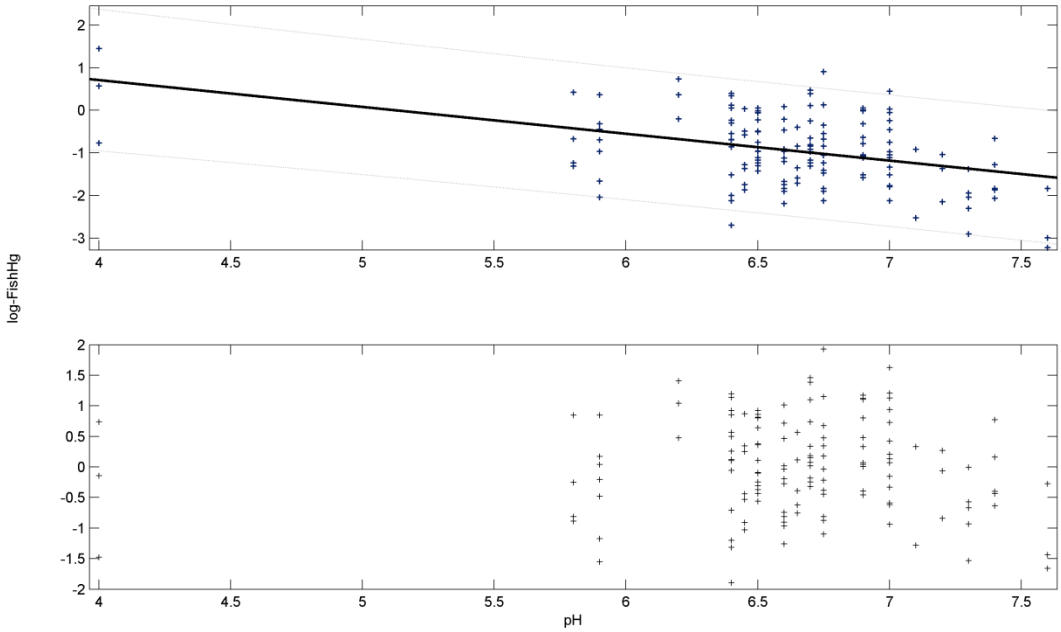

Figure S3: (top) Regression scatter plot of log-*Fish*Hg vs. pH used to derive *Fish*Hg soft data.

Dashed lines represent the 95% prediction bounds for new observations; (bottom) scatter plot of the residuals; p-values for the model coefficients were < 0.001.

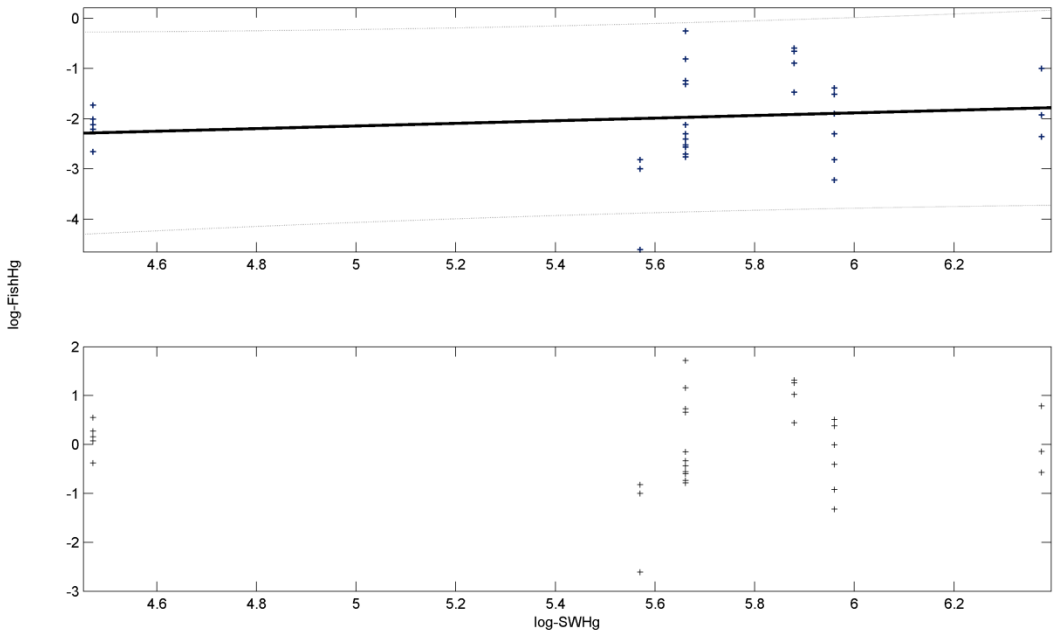

Figure S4: (top) Regression scatter plot of log-*Fish*Hg vs. log-*SW*Hg used to generate FishHg soft data; Dashed lines represent the 95% prediction bounds for new observations; (bottom) scatter plot of the residuals; intercept p-value: 0.035; variable coefficient p-value: 0.37.

Movie of Spatiotemporal Trend

Movie S1 can be viewed as an animated GIF at the following online location:

http://www.unc.edu/depts/case/BMElab/studies/HgFish_NC/CapefearLumber_HgFish_1991_2004.GIF

Movie S1: Space/time distribution of *Fish*Hg in the Cape Fear and Lumber Basins, every 180 days, between 1991-2004.