Lower Duwamish Waterway Group

Port of Seattle / City of Seattle / King County / The Boeing Company

Appendix H
Coverage Rates for Selected Upper Confidence Limit Methods for Mean of Total PCB in Sediments

Final Feasibility Study

Lower Duwamish Waterway
Seattle, Washington

FOR SUBMITTAL TO:

The U.S. Environmental Protection Agency
Region 10
Seattle, WA

The Washington State Department of Ecology
Northwest Regional Office
Bellevue, WA

October 31, 2012

Prepared by:
Kern Statistical Services, Inc.
5175 NE River Rd.
Sauk Rapids, MN 56379

COVERAGE RATES FOR SELECTED UPPER CONFIDENCE LIMIT METHODS FOR MEAN OF TOTAL PCB IN SEDIMENTS

LOWER DUWAMISH WATERWAY SEATTLE, WASHINGTON

March 30, 2010

Prepared for:
Assessment and Restoration Division
Office of Response and Restoration,
National Oceanic and Atmospheric Administration
7600 Sand Point Way NE
Seattle, WA

Prepared By:
Kern Statistical Services, Inc.
5175 NE River Rd.
Sauk Rapids, MN 56379

Introduction

Over 1300 locations were sampled along the Lower Duwamish Waterway (LDW) in efforts to characterize contaminant concentrations supporting remedial decision making. The data configuration is based on a biased sampling design with higher sample density within areas now identified as Early Action Areas (EAAs) and lower sampling density within the remainder of the river, referred to here as interstitial spaces. Because the sampling design is biased and the sample inclusion probabilities are unknown, ad-hoc methods have been proposed for estimation of upper confidence limits for the mean of contaminant concentrations within the surface sediments. The effect of the sampling bias is apparently large with the un-weighted mean PCB concentration being 1166 ug/kg and the Thiessen polygon weighted average being just 352 ug/kg—nearly a full order of magnitude lower. Understanding the most appropriate approach is of substantive importance. This study uses the sample data to develop a probability model of total PCB concentrations that is then used to test proposed UCL methods in efforts to develop an approach to UCL method selection for the site.

UCL Methods

Statistical methods are not generally available for biased sampling plans. Methods to correct sampling biases in efforts to approximate upper confidence limits for the mean have been proposed for the LDW, although the performance of proposed methods had not been tested prior to this study. In this study, three bias reduction methods were investigated; 1) IDW Interpolation, 2) Site stratification (2 and 11 strata) and 3) Thiessen polygon weighting. The 11 strata represent three interstitial areas, three navigational channel areas and 5 EAAs (Figure 1). The 2 stratum configuration treated all EAAs as one stratum and the remaining areas as the second stratum.

For the interpolated bias correction approach, Hall's bootstrap and the Bootstrap T were applied to the interpolated surfaces to obtain parameter estimates, while for the stratified approach bootstrap re-sampling followed the "naïve" bootstrap for the and the balanced bootstrap with importance sampling (Davison and Hinkley, 1986). The balanced bootstrap with importance sampling method was also used for the Thiessen polygon weighting method. In all seven approaches were tested. UCL methods were tested on the full LDW study area, as well as for data sets restricted to each of the three reaches identified in the feasibility study.

Synthetic Data

Total PCB data and sampling configuration from LDW were used to develop a probability model of the distribution of PCB contamination in sediment consistent with the stratification of the mean among EAAs and interstitial spaces as well as spatial correlation and nugget effect. Generally speaking the PCB concentration varies among EAAs with elevated PCB concentrations and interstitial spaces containing generally much lower PCB concentrations. Data and maps provided by LDWG further subdivide the site into three river segments which were further subdivided into the navigational channel and the remaining areas.

Synthetic populations were developed with simulated stratum means constrained to the observed stratum means using an analysis of variance model of the form

$$\log(PCB_{ij}) = \mu_i + \epsilon_{ij};\ \ i = 1,2,\ldots 11 \text{ and } j = 1,2,\ldots.,1248$$

where μ_i represents the log-mean concentration within the ith stratum and ϵ_{ij} is a mean zero and spatially correlated.

The residuals were subjected to a semi-variogram analysis and the sample semi-variogram and fitted model were plotted.

Equally likely synthetic surfaces were generated by simulating a mean-zero variance 1.0 spatially correlated surface which was then multiplied by the appropriate stratum specific standard deviation and added to the appropriate stratum mean. These re-trended values were exponentiated, arriving at synthetic populations with stratum means and variances; and spatial correlation consistent with the observed data. These synthetic values were used to populate the 10 by 10 foot IDW grid cells defined by the LDWG interpolation grid. These synthetic surfaces differ from interpolated surfaces in that they are not smooth, but rather retain the variance and spatial correlation observed in the sample data. One thousand such surfaces were simulated providing synthetic data with known statistical properties to which sample estimates could be compared.

Findings

1. Residuals were found to be approximately normally distributed (Figure 2).
2. Semivariogram was plotted in Figure 3 showing
    a. The range of influence of the log(pcb) residuals was approximately 70 feet.
    b. Small scale heterogeneity (i.e. nugget effect) constituted approximately 7% (0.17/2.45) of the total variance in log-scale.
3. Example synthetic means are plotted against observed sample means in Figure 4 showing that synthetic data reproduced large scale stratification observed in sample data.
4. One of the 1000 synthetic maps is shown in Figure 5, illustrating the large scale variation of mean concentration among strata as well as the smaller scale fluctuations in concentration characteristic of the interstitial spaces.
5. Confidence limits for the LDW were estimated for total PCBs using 2 methods based on re-sampling the IDW interpolated grids, 2 methods for each of 2 stratified sampling approaches --2 stratum and 11 stratum designs.
    a. All methods resulted in UCLs ranging from approximately 550 ug/kg to 700 ug/kg.
    b. The method that most closely reproduced 95% coverage was the 2 stratum approach with a UCL of 665 ug/kg and coverage rate of 95.3%.
    c. The difference between 550 ug/kg and 700 ug/kg is unlikely to substantively impact remedial decision making.
6. Study results for the LDW are conditional on the biased sample configuration, weighting scheme and analysis method selection. Robustness of methods to changes in sample size and subarea

population statistics was tested by repeating the simulations for each of the three reaches separately. Coverage probabilities are summarized in Table 2.

   a. Coverage rates for methods based on the balanced bootstrap (stratified or Thiessen weights) more closely matched the 95% nominal confidence level than did the interpolation based methods.
      i. The two stratum design with balanced bootstrap was no worse than 3% different from the nominal rate in each reach and was exact for the full LDW.
   b. Methods based on resampling from the interpolated grid, consistently understated the population variance and skewness within reaches as well as at the global level.
   c. Coverage rates for the interpolation based methods were either approximately 100% or 90% primarily due to under or over correction of the sample mean relative to the true population mean.

7. The UCL performance results cannot be generalized to UCL estimation for smaller subareas or other sites.
8. The data spacing within the interstitial spaces is typically on the order of 150 feet indicating that
   a. sample data density is adequate to confirm that large hotspots are unlikely to have been missed, but
   b. interpolated surfaces in the interstitial areas may be poorly constrained
   c. edges of EAAs remain areas of high uncertainty, and
   d. that smaller isolated hotspot areas on the order of 50-150 feet in diameter may remain.

UCL Performance Details

This study included analysis of the coverage rate for each UCL method, including investigation of key parameter estimates mean, variance and skewness to improve understanding of underlying root causes controlling method performance. Following is a summary of the findings for each method and the bias associated with individual parameter estimates.

1. Simulated coverage rates and the biases in parameter estimates are summarized in Table 1.
2. The IDW approach understated the population variance and skewness, but overstated the population mean.
   a. Low bias in the variance and skewness were expected based on mathematical relationship between the population variance and the variance of the smoothed IDW surface, as well as from previous simulations (Kern 2009).
   b. The mean was overstated by the IDW methods which was not expected based on any particular statistical theory, but rather was apparently due to the idiosyncrasies of the particular sampling configuration and the distribution of the underlying population.
   c. The high bias in the mean mitigated understatement of the variance and skewness, but this behavior cannot be expected in general as was shown in previous simulations (Kern 2009) in which the mean estimate was relatively unbiased.
3. Reproduction of coverage probabilities varied among methods, sample weighting assumptions and method of stratification

   a. Coverage rates ranged from 78% for the 11 stratum bootstrap T approach to 99.9% for both of the IDW based approaches.
   b. The 2 stratum approach using importance sampling resulted in the most accurate 95.5% coverage rate with an estimated UCL of 665 ug/kg based on the sample data.
   c. Coverage for the same approach applied to the 11 stratum design was 91%, which is moderately less than the target 95% rate. The estimated UCL based on sample data was 589 ug/kg.
4. The Hall's and Bootstrap T approaches based on IDW interpolation require estimates of the mean and variance and additionally, the Hall's method also requires an estimate of the skewness of the underlying population. For the IDW surface,
   a. the population mean was overstated on average by 14%
   b. the population variance was understated on average by 13%
   c. the skewness was understated on average by 45%
5. The importance sampling approach based on stratified sampling provided a more accurate estimate of the mean which is the only estimate required for the method.
   a. For the 2 stratum case the estimated mean was 3% greater than the population mean on average.
   b. For the 11 stratum case, the estimated mean was 6% less than the population mean on average.
6. For estimation of the mean and UCL for the LDW, the large sample size (N> 1300 locations) is probably the most important factor causing estimated UCLs to be similar.

References

Davison, A.C., Hinkley, D.V. and Schechtman, E. (1986) Efficient bootstrap simulation. Biometrika, 73, 555–566.

Kern Statistical Services, Inc. 2009. Review of Appendix-H: Hall's upper confidence limit for IDW-interpolated data, draft feasibility study. Lower Duwamish Waterway, Seattle, WA.

Table 1. Performance summary for IDW interpolation and stratified sampling based UCL estimates

Method				Average Ratio of Estimated to True Parameters		
		Estimated UCL (mg/kg)	Simulated Coverage Rate	Mean	Variance	Skewness
Interpolated				1.14	0.87	0.55
	Halls Bootstrap	702	99.9%			
	Bootstrap T	545	99.9%			
Stratified Design (2 stratum case)				1.03	NA	NA
	Bootstrap T	629	87.4%			
	Balanced Bootstrap With Importance Sampling	665	95.3%			
Stratified Design (11 stratum case)				0.94	NA	NA
	Bootstrap T	544	72.1%			
	Balanced Bootstrap With Importance Sampling	589	90.8%			
Thiessen Polygon Method	Balanced Bootstrap With Importance Sampling	680	99.2%	1.06	1.15	NA

Table 2. Summary of coverage rates for 5 UCL methods for reaches 1, 2 and 3 and the full LDW study area. Stratified approaches were based on the two stratum configuration.

	Halls Interp	Bootstrap T Interpolated	Bootstrap T Stratified	Balanced Bootstrap Thiessen	Balanced Bootstrap Stratified
Reach 1	90%	89%	92%	93%	94%
Reach 2	100%	100%	96%	97%	98%
Reach 3	91%	91%	88%	97%	92%
Full Site	100%	100%	87%	99%	95%

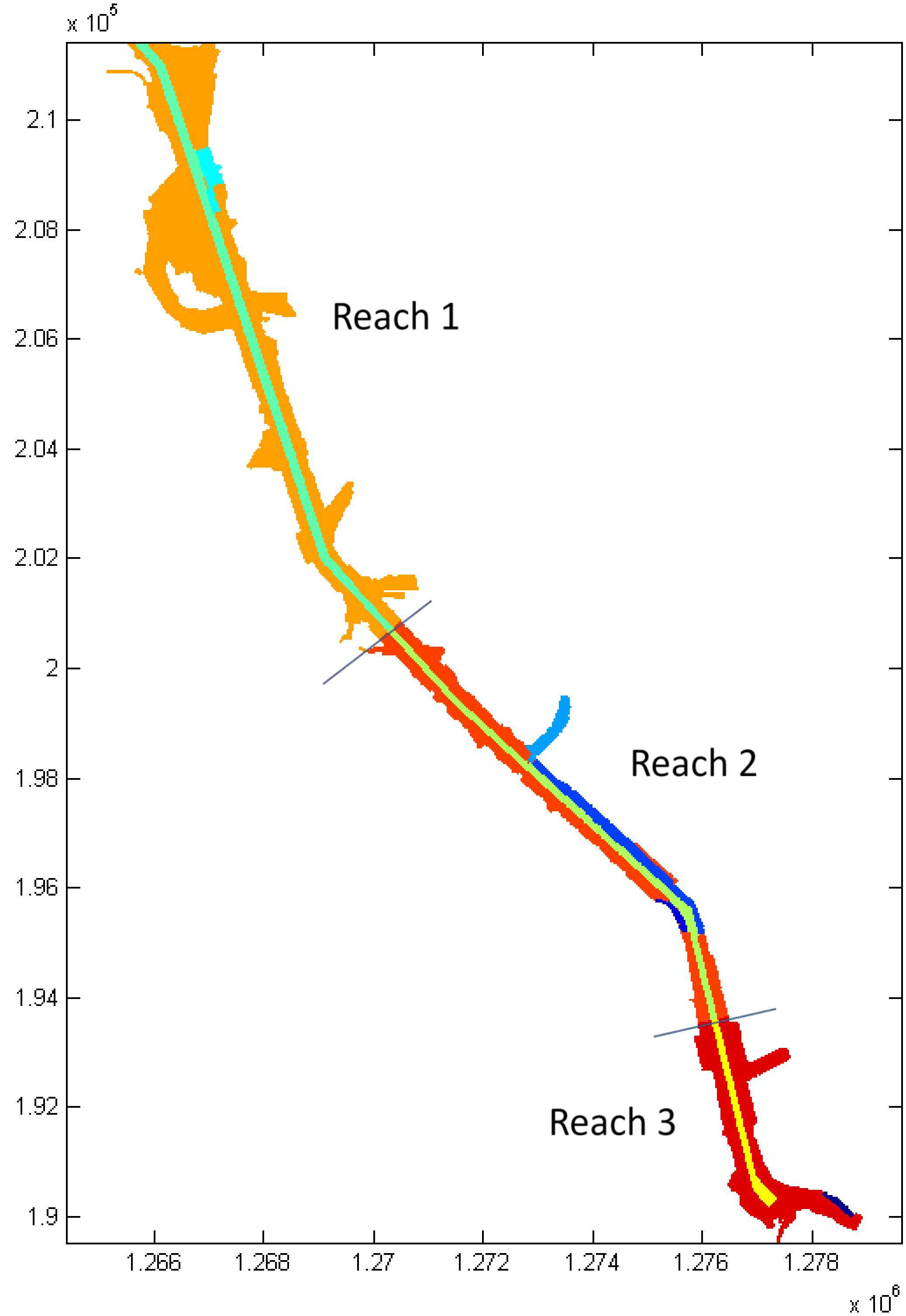

Figure 1. Stratification of study area in the Lower Duwamish Waterway Site.

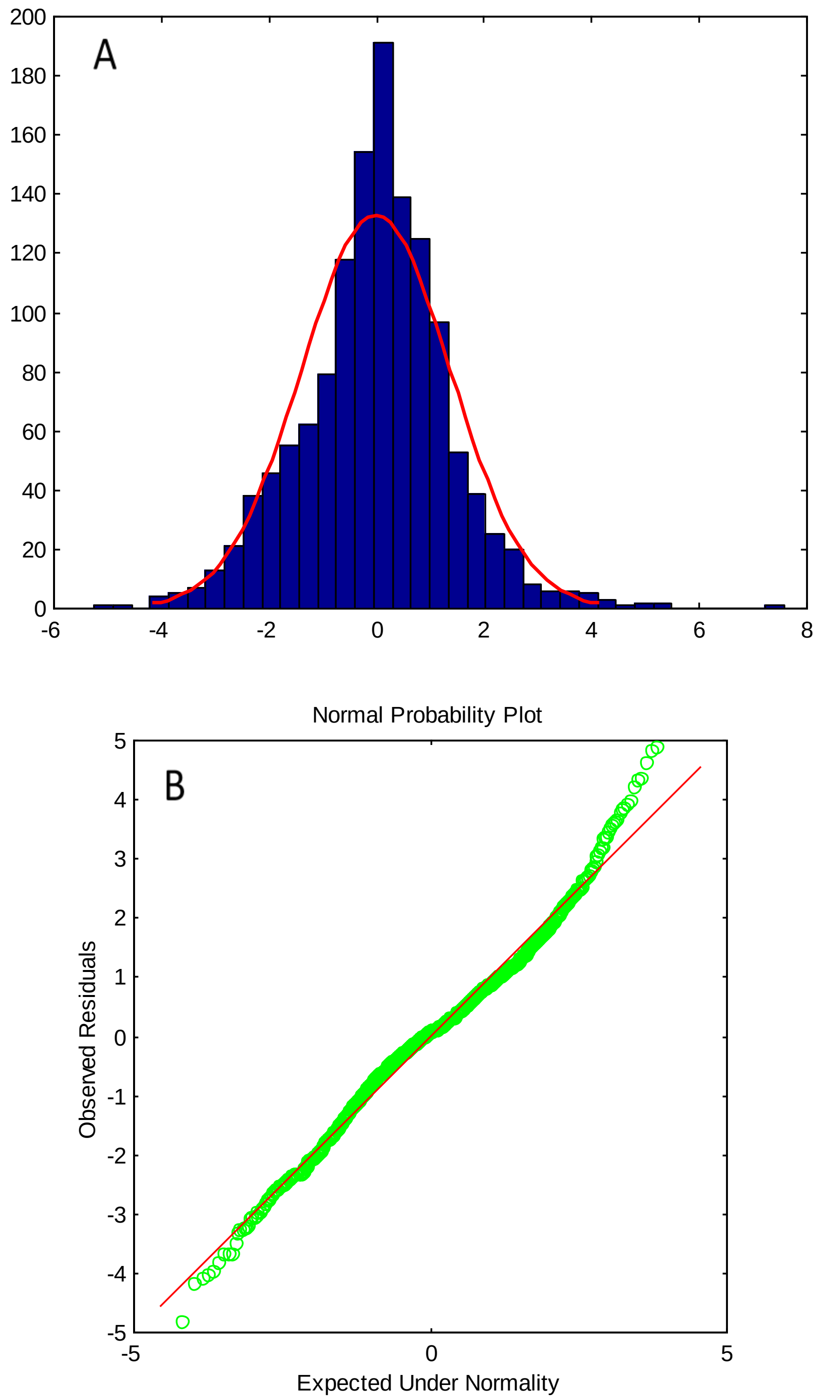

Figure 2. Histogram of residuals of log(PCB) with fitted normal distribution (panel A) and normal probability plot for residuals (panel B). Residuals are similar to a normal distribution (p> 0.10)

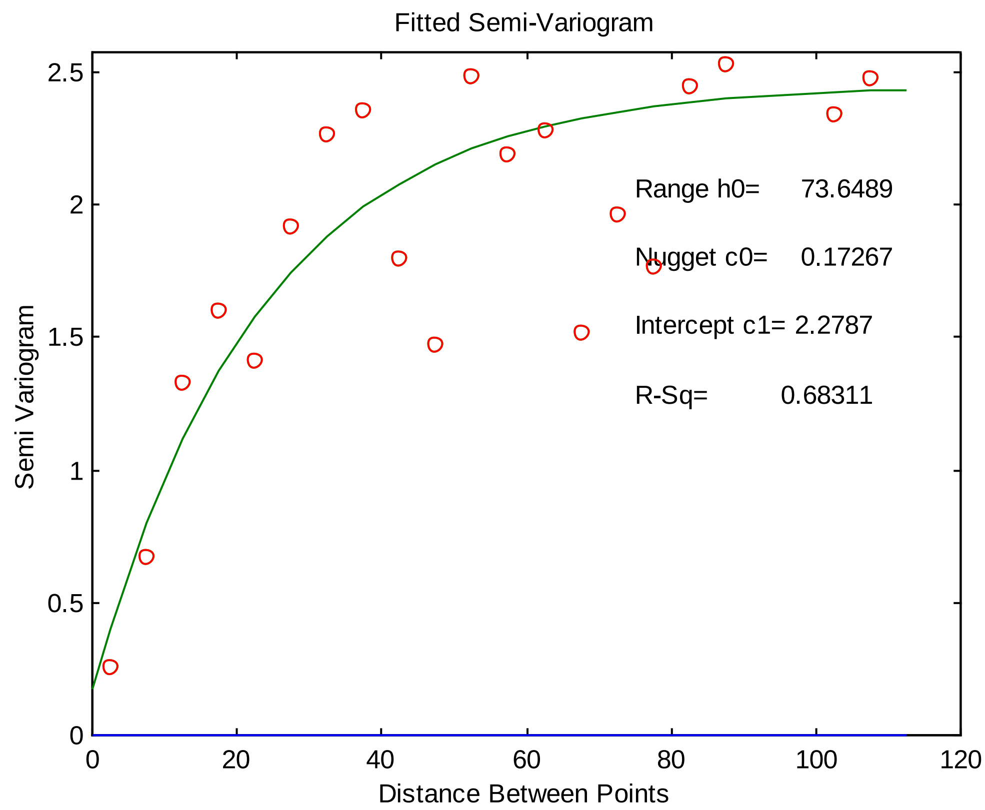

Figure 3. Semivariogram of residual log(PCB) concentration.

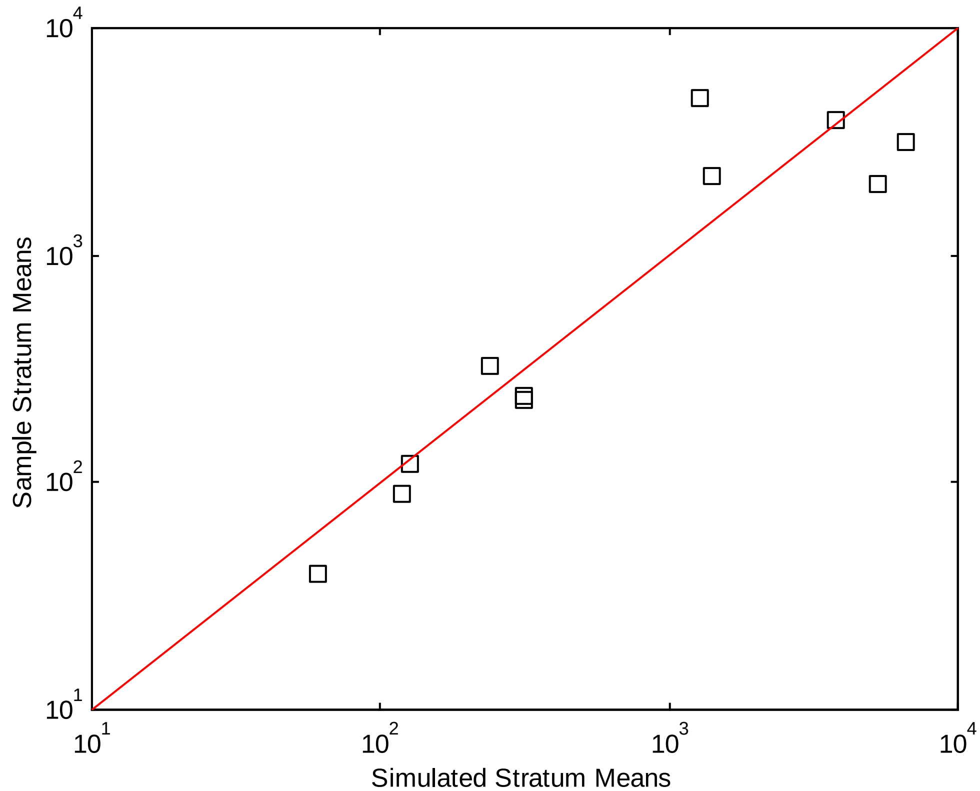

Figure 4. Sample vs. simulated stratum means from one synthetic population.

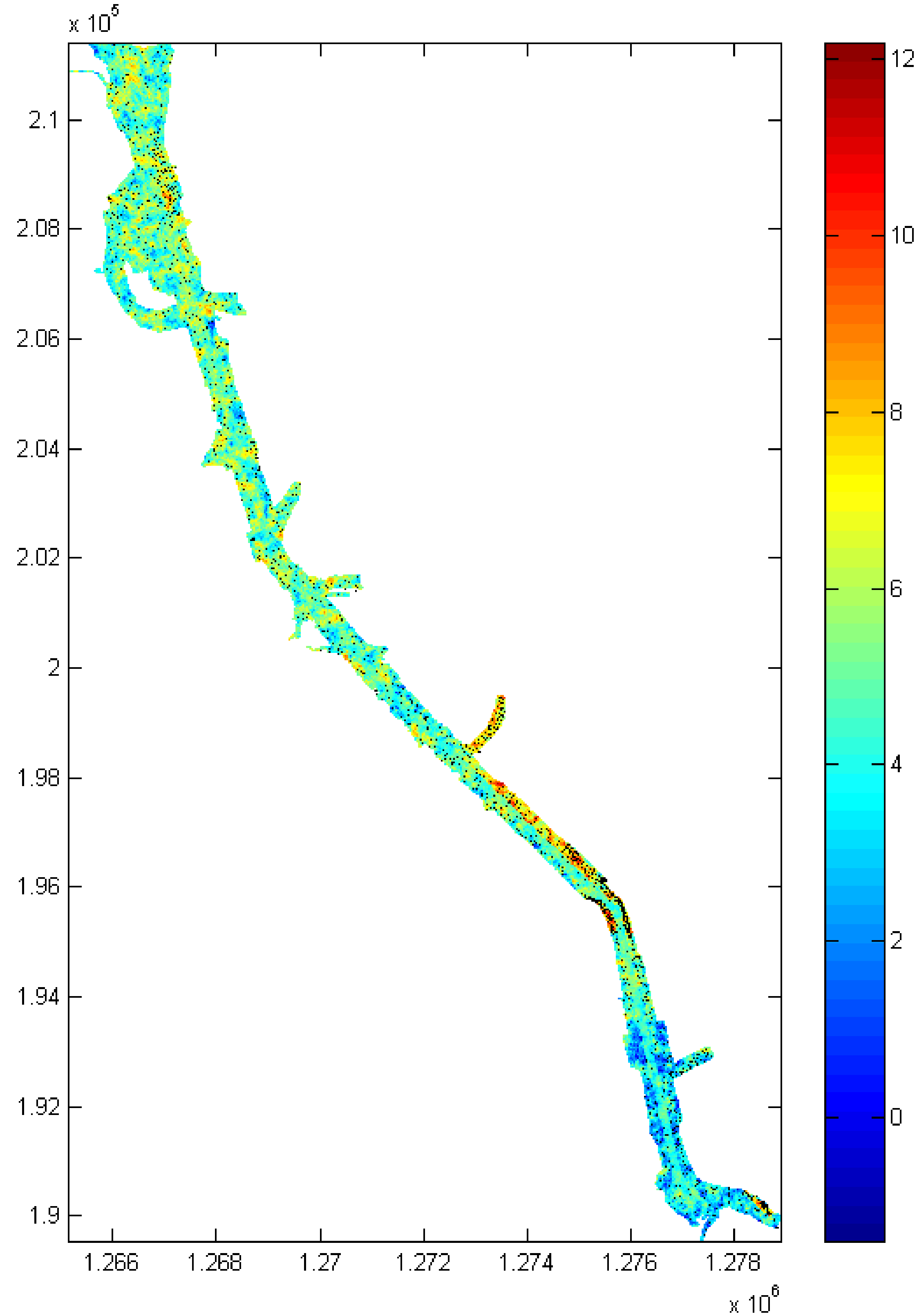

Figure 5. Simulated log(PCB) concentration in surface sediments. One of 1000 realizations generated. Black dots represent sample locations.