Lower Duwamish Waterway Group

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

Appendix H – Computing Hall's Upper Confidence Limit (UCL) for IDW-Interpolated Data Draft 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 Field Office
Bellevue, WA

April 24, 2009

Prepared by: AECOM

710 Second Avenue, Suite 1000 ♦ Seattle, Washington ♦ 98104

Table of Contents

List of Tables

H.1 UCL95 Introduction

In the feasibility study (FS), the Lower Duwamish Waterway Group (LDWG) is using a 95th percentile upper confidence limit (UCL95) of the estimated spatially-weighted average concentration (SWAC).

The UCL95 is used to estimate the upper bound on the average exposure concentration that a particular receptor might be exposed to in a specified area over time (EPA 1989; EPA 1992). Use of interpolated, spatially-weighted chemical concentrations accounts for the fact that receptors can move around the exposure area, and partially corrects for non-random sampling within the site. As one line of evidence for evaluating the effectiveness of remedial alternatives, the UCL95 serves as both a relative gauge of remedial alternative effectiveness (i.e., the lower the UCL95, the more effective the alternative) and a benchmark to evaluate how well the alternatives are achieving the preliminary remediation goals. The UCL95 has come into widespread use for monitoring and compliance efforts. The existence of high-quality statistical software makes UCL95 estimates easy to calculate for virtually any size dataset.

For this FS, the UCL95 was determined for three risk-driver chemicals based on spatially interpolated data; these chemicals are total polychlorinated biphenyls (PCBs), carcinogenic polycyclic aromatic hydrocarbons (cPAHs), and arsenic.[1] A UCL95 estimate was calculated for each of these three chemicals, assuming each of the five remedial alternatives under review are implemented as currently proposed.

H.2 Datasets

LDWG developed SWACs from interpolated concentrations of PCBs, cPAHs, and arsenic using the inverse distance weighting (IDW) methodology (ENSR 2008; see Appendix A) applied to the baseline surface sediment data from the Remedial Investigation (RI) (refer to Section 2 for a detailed description of the RI dataset).

For the UCL95 evaluation, estimates were based on interpolated concentrations of the three chemicals at three points in time following remediation:

- Year 0 (immediately following remediation),
- Year 10 (10 years post-remediation)
- Year 30 (30 years post-remediation).

[1] Dioxins/furans and SMS chemicals are also risk-driver chemicals for the LDW. However, insufficient dioxin/furan data exist site-wide to generate spatially-interpolated datasets, and SMS chemical exceedances are expressed as Thiessen polygons (SWACs are not derived). Consequently, these data are not included in this analysis.

The surface sediment concentration estimates for each chemical at each time point were calculated using the bed composition model (BCM), described in Section 5 of the FS. The BCM couples the physical sediment transport model with the baseline concentrations of chemicals in surface sediments to predict future surface sediment concentrations under different remedial action scenarios.

For each of the five remedial alternatives under review, the concentration within the remedial footprint was "replaced" with a post-remedy bed sediment replacement value. These input values account for residual surface contamination from the resettling of contaminated sediments suspended during remedial activities and the redistribution of sediment from adjacent areas. As described in Section 5 of the FS, input parameters for the BCM were developed using various lines of evidence (e.g., data from upstream Green/Duwamish River inflows).

The method for calculating the UCL95 on the SWAC for use in this FS is further described in Section H.3.

H.3 Hall's UCL Statistical Program: Methods

The UCL95 for the LDWG FS was calculated using the Hall's method, which is described in Attachment 4 of the *Human Health Risk Assessment for the GE/Housatonic River – Rest of River* project (Weston Solutions, Inc. 2005). Hall's method is a bootstrap method well-suited for skewed datasets of unknown distribution. The Hall's method makes no assumptions about the shape of the data distribution (i.e., normal or log normal), and is not adversely affected by the skewness of the data. The method does assume that the concentrations analyzed are representative of the underlying distribution of concentrations within the area. Thus, Hall's method does not account for non-random or targeted sampling of an area (for example, delineation of an area of elevated chemical concentrations).

Hall's method was approved by the U.S. Environmental Protection Agency (EPA) for use on the Housatonic River project (Weston Solutions, Inc. 2005). The Housatonic project team used a software program ("hallbig2") to perform their Hall's UCL95 calculations. This program was written using Pascal code, and runs as an executable file (*.exe). This software was used to calculate the UCL95 estimates for the LDW project, with minor modifications as described below. ProUCL, which is software distributed by EPA, also performs UCL95 calculations. However, ProUCL only accommodates input arrays with fewer than 66,000 values; the IDW dataset has over 182,000 data values based on 10-ft by 10-ft grid cells. Although the original "hallbig2" program was limited to use with far fewer samples, it was possible to modify the program to accommodate larger datasets (see Section H.3.2). Another reason for selecting the "hallbig2" program over ProUCL was that it has been tested and vetted by its use on the Housatonic Superfund project (for which ProUCL developer Anita Singh was also a consultant).

Methods, inputs, and modifications to the "hallbig2" program used for the LDW are discussed below.

H.3.1 Degrees of Freedom

The spatial weighting created in the IDW-interpolated datasets introduced a complication in calculating the UCL95. Specifically, in calculating a UCL on sample data, the degrees of freedom are determined by subtracting 1 from the number of samples used in the calculation. An IDW interpolation, however, yields a concentration estimate at every grid node, thus yielding many more values than in the baseline surface sediment dataset. For example, the degrees of freedom based on the actual sample size for PCBs would be 1,326 for the LDW; by contrast, the number of grid nodes at which PCB concentrations were interpolated is over 182,000.

Consequently, the degree of freedom value used in the Hall's calculation was based either on the original sample size or the assumed post-remedy sample size (see H.4.3), not the number of grid nodes. Thus, if "*n*" original samples were used to interpolate concentrations at "*m*" grid nodes, the degrees of freedom used in the UCL calculation was *n*-1.

H.3.2 Program Modifications

The original "hallbig2" program, which was used to calculate the UCL95 for the Housatonic project was limited in that it could accommodate only datasets of 800 or fewer original sample values. However, the LDW RI datasets have original sample sizes of 1,327 for PCBs, 828 for cPAHs, and 852 for arsenic. Additionally, the "hallbig2" program requires that input data be grouped into concentration ranges distributed between the minimum and maximum values, akin to bins in a histogram. The "hallbig2" program could not accept the number of values per concentration range (or bin) generated by the LDW datasets.

To account for these limitations, LDWG enlisted Scott Ferson[2] (Applied Biomathematics) to help modify the "hallbig2" program to accommodate the LDW interpolated datasets ("hallbig3"). Mr. Ferson increased the capability of the program to accommodate sample sizes of 1,500 and adjusted the program to accept datasets with as many as 100,000 interpolated values per concentration range. The program was also enhanced to generate streamlined summary output from multiple scenarios that were sequentially calculated.

[2] Scott Ferson was the lead contractor on the Housatonic Project assigned to run the UCL95 statistical program. He has been retained by EPA Region 10 as a statistical consultant for the LDW RI/FS project.

H.4 Sensitivity Analysis

Three aspects of the UCL95 estimate were evaluated, specifically checking for bias, coverage percentage, and effect of the degrees of freedom. Bias potentially associated with the software program modifications was evaluated by comparing results from the modified program to results obtained from the original program. Coverage measures the frequency with which the UCL95 estimate is above the true average value. This was evaluated by comparing the UCL95 estimate against the estimated SWAC from the original sample data. The effect of the degrees of freedom on the UCL95 estimates was assessed to understand how such estimates may be affected in the future when applied to the smaller post-remediation sample sets. UCL95 estimates for three different sample sizes (thus different degrees of freedom) were compared.

H.4.1 Bias Evaluation

A quality assurance/quality control check was performed to evaluate the effect of the program modifications on results by comparing the output of the final LDW modified program with the output of the original "hallbig2" program used on the Housatonic project (Table H-1). To conduct this evaluation, the interpolated dataset was limited to those values in the tribal clamming area, because this data subset was small enough to meet the input requirements of the original program. The resultant UCL95 estimates from both versions of the software program were comparable; the relative percent differences in the UCL95 for PCBs, cPAHs, and arsenic was 0.13, 0.12, and 0.00, respectively.

H.4.2 Coverage Evaluation

As discussed in Attachment 4 of the *Housatonic Human Health Risk Assessment* (Weston Solutions, Inc. 2005) and reiterated here, all UCL calculation methods provide only an estimate of the upper bound of the true average concentration. Because the true average concentration can never be truly known, the UCL95 estimate is designed to be above this unknown value for 95% of all input samples, assuming that all possible input samples (datasets) were collected in the same manner.

This is a statistical concept frequently difficult for non-statisticians to understand. Consider if the entire sampling program and interpolation had been repeated 100 times for the LDW. Each dataset would yield another UCL95 estimate, resulting in potentially 100 different estimates. The 95% confidence level indicates that 95 of the 100 estimates will be greater than the true average concentration. The frequency with which the calculated UCL95 estimate exceeds the true mean is referred to as the method's coverage rate; the coverage rate measures how the method performs its function as an upper estimate of the mean. A method's coverage rate depends on the characteristics of the input data to which it is applied. Thus, it is important to evaluate a method's actual coverage rate when applied to data with characteristics expected to be encountered.

The "hallbig2" program was determined to have acceptable coverage for the Housatonic project (Weston Solutions, Inc. 2005). For purposes of the LDW FS, a qualitative assessment of coverage was conducted by comparing the UCL95 estimates with the SWAC estimates for post-remediation surface sediment conditions immediately following construction (Year 0), as shown in Table H-2. In all but one case, the UCL95 estimate was larger the SWAC estimate, even for PCBs, which have the highest degrees of freedom of all three chemicals (assuming the original dataset sample size). The lone instance wherein the SWAC was higher than the UCL95 was under Alternative 5 for PCBs; the relative percent difference in the two values was only 1.9. This suggests the "hallbig3" program is appropriately conservative (i.e., functions well to calculate an upper estimate of the mean).

H.4.3 Degrees of Freedom Evaluation

The RI surface sediment dataset formed the foundation for the IDW interpolation analysis. This dataset has 800+ values for each of the three chemicals. However, post-remediation datasets will likely be substantially smaller than the RI dataset.

Thus, to evaluate the influence of smaller datasets on the UCL95, the Hall's UCL95 estimates were compared using three sample sizes (Table H-3):

1. Original RI dataset
2. Reduced dataset of 400
3. Reduced dataset of 100.

The two smaller sample sizes were used to bracket the likely post-remediation sample sizes. These sizes were based on an assumed post-construction monitoring program.

The degrees of freedom made little difference in the case of arsenic, except for the baseline and Alternative 1 at Year 0 (Table H-3). For PCBs and cPAHs, the lower degrees of freedom (i.e., smaller sample sizes) resulted in generally higher UCL95 estimates, as would be expected to retain the same coverage rate. This was particularly true for the small remedial footprints (Alternatives 1, 2, 3a, and 4a).

The relative percent differences between the highest (RI) and lowest (100) degrees of freedom are shown on Table H-3. Generally, for any given chemical at a particular time, the lowest relative percent differences between the highest and lowest degrees of freedom are seen for Alternatives 3 and 4. The degree of freedom has very little influence on the UCL95 between sample sizes of 400 and 100. Using the smaller sample sizes results in slightly higher UCL95s for all times than those generated using the original data sample size, but the UCL95s are within the uncertainty of sampling accuracy. The FS is conservative because it assumes the smaller sample size of 100; the true UCL95 could be lower.

H.5 References

ENSR 2008. *Revised Memorandum: Inverse Distance Weighting Methodology for Interpolating Surface Sediment Chemistry in the Lower Duwamish Waterway Feasibility Study*. August 15.

EPA (U.S. Environmental Protection Agency) 1989. *Risk Assessment Guidance for Superfund, Volume I - Human Health Evaluation Manual (Part A)*. Interim Final. EPA/540/1-89/002. Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, DC.

EPA (U.S. Environmental Protection Agency) 1992. *A Supplemental Guidance to RAGS: Calculating the Concentration Term*. Publication 9285.7-081. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, DC.

Weston Solutions, Inc. 2005. *Human Health Risk Assessment GE/Housatonic River Site Rest of River*. Prepared for Environmental Remediation Contract General Electric/Housatonic River Project, Pittsfield, Massachusetts for submittal to U.S. Army Corps of Engineers, Concord Massachusetts and U.S. Environmental Protection Agency, Boston, Massachusetts. Submitted to EPA February 11, 2005.

Table H-1 Bias Evaluation of Original Hall's Program to Modified Version

Program	Total PCBs			cPAHs			Arsenic		
	No. of Samples	Mean	UCL95	No. of Samples	Mean	UCL95	No. of Samples	Mean	UCL95
hallbig2 (Housatonic)	475	510.88	755.54	298	359.49	463.76	309	12.69	20.6959
hallbig3 (LDW-modified)	475	510.88	754.54	298	359.49	463.22	309	12.69	20.6959
Relative Percent Difference			0.132			0.116			0.00

Notes:
The purpose of this exercise was to check that slight modifications to the Hall's Bootstrap program yield results consistent with the unmodified program that was vetted and accepted by EPA on the Housatonic project.

hallbig2 is the program version used on the Housatonic HHRA.

hallbig3 is a modified version of hallbig2 that allows a higher maximum sample count and higher maximum IDW grid cell count. The program was modified with the assistance of Scott Ferson (Applied Biomathematics).

The test datasets used for the comparison correspond to the tribal clamming exposure area. The IDW interpolations were performed using parameters established previously (ENSR 2008). The clamming area was selected for the test dataset because the number of grid cells in the LDW-wide dataset is >180,000, which exceeds the limits of hallbig2. The number of grid cells in the clamming area dataset is 44,255, small enough to be processed by hallbig3. The datasets used for comparison correspond to Alternative 1 at Year 0.

The "number of samples" corresponds to samples from the RI dataset that are located within the tribal clamming area and also within a region extending 150 ft from the edge of the clamming area. The search radius for the IDW interpolation is 150 feet. This captures all samples that influence the interpolated values within the exposure area.

Table H-2 Degree of Coverage Evaluation: Hall's UCL95 Values vs. SWAC for Post-Remediation IDW Datasets

Post-Remediation (Year 0)						
	Total PCBs (µg/kg dw) RV = 75		**Arsenic (mg/kg dw) RV = 13**		**cPAHs (µg/kg dw) RV = 200**	
Alternative	**UCL95**	**SWAC**	**UCL95**	**SWAC**	**UCL95**	**SWAC**
0	508	342	19.9	15.9	421	393
1	253	186	19.9	16.0	377	357
2	166	143	13.0	12.3	334	315
3a/4a	155	128	12.9	12.2	278	267
3b/4b	136	114	12.8	12.1	270	259
3c/4c	128	108	12.8	12.2	259	249
3d/4d	128	108	12.8	12.2	259	249
5	68	70	12.2	12.1	199	195

Notes:

UCL95 values calculated for site-wide post-remediation IDW interpolated datasets

The post-remedy bed sediment replacement value (RV) for total PCBs in the Alternative 5 dataset is 50 µg/kg dw.

The DOF used for the UCL95 calculations correspond to the sample sizes in the original RI dataset.

	PCBs	Arsenic	cPAHs
DOF =	1,326	851	827

DOF = degrees of freedom; IDW = inverse distance weighting; RV = post-remedy bed sediment replacement value (concentration) assigned to grid cells within actively remediated areas.

Table H-3 Degrees of Freedom Evaluation: Hall's UCL95 Values for Post-Remediation IDW Datasets

Post-Remediation (Year 0)												
	PCBs (µg/kg dw) RV = 75				Arsenic (mg/kg dw) RV = 13				cPAHs (µg/kg dw) RV = 200			
Alternative	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3
0	508	508	508	0.04	19.9	23.9	23.9	18.17	421	444	629	39.57
1	253	253	253	0.05	19.9	23.9	23.9	18.19	377	391	464	20.63
2	166	174	174	4.56	13.0	13.0	13.0	0.00	334	345	409	20.15
3a/4a	155	155	155	0.07	12.9	12.9	12.9	0.00	278	284	311	11.13
3b/4b	136	136	136	0.36	12.8	12.8	12.8	0.00	270	276	303	11.67
3c/4c	128	128	128	0.23	12.8	12.8	12.8	0.00	259	265	293	12.24
3d/4d	128	128	128	0.23	12.8	12.8	12.8	0.00	259	265	293	12.16
5	68	70	74	8.30	12.2	12.3	12.5	1.77	199	202	224	11.93

Post-Remediation (Year 10)												
	PCBs (µg/kg dw) RV = 75				Arsenic (mg/kg dw) RV = 13				cPAHs (µg/kg dw) RV = 200			
Alternative	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3
0	185	185	185	0.15	12.6	12.6	12.6	0.00	292	312	312	6.51
1	96	120	120	22.43	12.6	12.6	12.6	0.00	272	279	346	23.99
2	87	100	100	13.45	10.9	10.9	10.9	0.00	254	258	275	7.67
3a/4a	84	94	94	11.73	10.9	10.9	10.9	0.00	242	245	260	7.38
3b/4b	77	85	85	9.55	10.8	10.8	10.8	0.00	239	242	258	7.71
3c/4c	76	82	82	7.34	10.8	10.8	10.8	0.00	236	239	256	8.23
3d/4d	76	82	82	7.39	10.8	10.8	10.8	0.00	236	239	256	8.18
5	66	69	73	10.07	10.6	10.6	10.7	0.76	222	226	258	14.90

Post-Remediation (Year 30)												
	PCBs (µg/kg dw) RV = 75				Arsenic (mg/kg dw) RV = 13				cPAHs (µg/kg dw) RV = 200			
Alternative	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3	DOF_1	DOF_2	DOF_3	RPD: 1 vs 3
0	107	107	107	0.28	11.0	11.0	11.0	0.00	259	259	259	0.00
1	71	74	86	18.79	11.1	11.1	11.1	0.00	245	258	282	14.11
2	70	72	83	17.38	10.2	10.3	10.3	0.63	240	248	279	15.16
3a/4a	68	71	80	16.22	10.2	10.3	10.3	0.77	238	246	276	14.74
3b/4b	68	70	80	15.75	10.2	10.3	10.3	0.76	238	246	276	14.76
3c/4c	68	70	79	15.80	10.2	10.3	10.3	0.77	238	246	275	14.69
3d/4d	68	70	79	15.80	10.2	10.3	10.3	0.77	238	246	275	14.69
5	66	68	75	12.78	10.2	10.3	10.3	0.94	231	235	276	18.00

Notes:

UCL95 values calculated for LDW-wide post-remediation IDW interpolated datasets.

The post-remedy bed sediment replacement value (RV) for total PCBs in the Alternative 5 dataset is 50 µg/kg dw.

DOF_1 is based on the number of samples used in the IDW interpolation. DOF_2 and DOF_3 are based on assumed lower sample numbers of 400 and 100, respectively.

	PCBs	Arsenic	cPAHs
DOF_1 =	1,326	851	827
DOF_2 =	400	400	400
DOF_3 =	100	100	100

DOF = degrees of freedom; IDW = inverse distance weighting; RPD = relative percent difference; RV = post-remedy bed sediment replacement value (concentration) assigned to grid cells within actively remediated areas.