Appendix D

Characterizing Background Populations

1.0 Introduction

Washington Department of Ecology (Ecology) contracted with TerraStat Consulting Group to help identify statistical tools that could be used to distinguish between populations of Natural Background and Regional Background, as well as how to best summarize available data for background for the different objectives required under Sediment Cleanup programs. Statistical tools that can be used to separate a dataset into distinct populations with overlapping concentrations generally rely on either *a priori* description of the underlying characteristics for one of the populations so that separation of a mixture distribution can be accomplished; or data sets with sufficient spatial coverage such that the concentration surfaces can be modeled and locations where changes in the concentrations occur can be identified. Neither of these scenarios is met by the data that are generally available for Puget Sound background sites, so this work product ultimately became the description of a process for describing and generating appropriate summary statistics from existing Regional Background and/or Local Natural Background data sets.

TerraStat reviewed existing approaches to the description of Background that have been used by various agencies; these definitions and approaches are briefly touched on in **Section 2.** Knowledge from Ecology staff, from other agency approaches, and about the statistical tools appropriate for the type of data available for Puget Sound Background areas was integrated to describe a process for identifying and characterizing Background. A description of the process and important considerations to this process are included in **Section 3**; applications to three Puget Sound case studies are included in **Section 4**. Final recommendations regarding the description of Background are provided in **Section 5**.

2.0 Approaches to Background

TerraStat reviewed approaches to Background as described in USEPA guidance (US EPA 1995, US EPA 2002), and as used by WDOE for several local sites (summarized by NewFields 2011), by OR DEQ for the lower Willamette River, and by the ACOE for DMMP sites (DMMP 2011). These approaches all started with the *a priori* selection of the geographic boundaries that constitute background and then proceeded to summarize the background data set in different ways for different purposes.

Some working definitions of background in use by EPA in the CERCLA (Superfund) program (US EPA 2002) are "substances or locations that are not influenced by the releases from a site, and are usually described as <u>naturally occurring</u> or <u>anthropogenic</u>:

1) <u>Naturally occurring</u> - substances present in the environment in forms that have not been influenced by human activity *[matches Ecology's definition for 'Natural Background']*; and,

2) Anthropogenic - natural and human-made substances present in the environment as a result of human activities (but not specifically related to the CERCLA site in question)" *[matches Ecology's definition of 'Regional Background']*.

In the EPA documents reviewed, the background data set is not extracted and statistically separated from a larger dataset, but rather it is acquired via appropriate sampling within the boundaries of the area expected to meet the background definition. EPA emphasizes that background sampling is a strategic, biased sampling event with sampling to occur "in areas expected to be outside the area influenced by the site." So information about the fate and transport of contamination from the site, as well as best professional judgment (BPJ) must be used to identify the background areas. It is also recommended that possible concentration outliers should be identified and decisions regarding these outliers be made by the project team (US EPA 2010).

The reviewed approaches used by Washington DOE, Oregon DEQ, and Puget Sound DMMP all define the geographic boundaries for the background population using BPJ and knowledge about the site, which is consistent with EPA's recommendations. Once the background population has been identified, the summary statistics suited to the intended application may be calculated (e.g., central tendency value such as 95UCL on the mean for exposure related questions, or an upper threshold of the distribution such as 90/90 UTL for a not-to-exceed value for station-by-station comparisons).

3.0 Background Considerations and Recommendations

As with any sampling or analysis plan, it is important to clearly state the objective up front.

Sampling Objective: To characterize a "Background" population in order to set clean-up goals for a site, and to delineate site clean-up boundaries.

Caveats:

- Current DOE guidance utilizes several definitions of "Background", including Puget Sound-wide Natural Background, Local Natural Background, and Regional Background. The method(s) proposed herein will not separate the available data into these separate populations; but they will allow the proper estimation of summary statistics from whatever boundaries are considered to constitute the relevant background population.
- We provide no statistical recommendations for how to separate Local Natural from Regional Background –pattern matching requires knowledge of the source pattern; and existing sampling locations may be insufficient to adequately describe the spatial patterns needed to statistically identify boundaries of the different types of background.
- Identifying the site-influenced areas needs to be done first with site knowledge combined with spatial contouring. Areas near known point sources, or areas adjacent to those point sources with elevated concentrations, are excluded *a priori* from the possible background population. If there are obvious trends away from the site, then a boundary can be determined with the help of contour mapping.
- Methods to identify univariate and multivariate outliers should be used, and samples that are clearly different from the others should probably be excluded from the background population.

3.1 Designing a Sampling Plan to Characterize Background

The ideal situation where a new survey is designed specifically to characterize background involves first determining an appropriate boundary for the Local Natural or Regional Background area, and then taking a spatially-balanced random sample that uniformly covers the area using a systematic random sample (e.g., random samples within a grid). Other methods for achieving a spatially-balanced random sample can also be used (e.g., generalized random tessellation stratified [GRTS] sampling used by US EPA [2011], and implemented via the RRQRR algorithm in GIS), but are not generally necessary for this situation. If there are areas that are found to have elevated concentrations and are suspected to be due to a point source, the boundary of the regional background area can be adjusted after sampling (i.e., the areas with high concentrations and their corresponding data can be excluded). If the gridded samples are found to be spatially autocorrelated, the gridded data points can be subsampled or methods that account for autocorrelation can be used to generate summary statistics. Otherwise, the set of samples taken on the grid can be assumed to be an independent random sample, and the appropriate summary statistics can be easily generated in ProUCL (for example).

3.2 Working With Existing Data to Characterize Background

The following recommendations address situations when compilations of existing data sets are being used to characterize background. In these cases, the "ideal" design may not have been used to generate the background data set, and the background area may have been sampled non-uniformly and non-randomly.

For the case studies examined in this report, sampling points were located unevenly through space and time, and collected for different purposes. In these case studies, the boundary of the characterized area is delineated by the locations of the existing samples and these existing sampling locations may be insufficient to fully characterize the background population. Within the compilation of data available to use for background, there may be non-randomness, non-independence, and more than one population represented. The simplest statistics (e.g., 95th UCL on the mean) assume an independently and identically distributed (i.i.d.) random sample. If the dataset violates these basic assumptions, then the simple statistics may be biased, and the variance poorly estimated. These basic assumptions are relevant to a number of available methods for summarizing data (i.e., bootstrapping, generating a trend surface via interpolation, kriging, etc.).

Given a data set consisting of multiple studies, there are three main steps to be followed to define a background concentration distribution and produce unbiased estimates of summary statistics. These are:

- Step 1. Delineate Background by excluding areas near known point sources, and areas suspected to be of a different population based on proximity to local influences (e.g., developed shorelines). Initial evaluation of the compiled data should include identification of possible outliers.
- Step 2. Determine the extent of autocorrelation and/or trend in samples from the background area.
- Step 3. Generate upper bound estimates for the regional background concentration distribution using an independent subset of the data based on the results from Steps 1 and 2.

3.2.1 Step 1. Drawing Background Boundaries

Unless new data are collected, the background population will be defined by the area that has been sampled. However, the sample locations should be examined to ensure that areas near known point sources are excluded. In this stage it is also important to identify possible outliers indicative of an unsuspected source signal, potential mixture distributions due to the presence and blending of two or more strong and different local signals (e.g., Bellingham Bay), and to determine spatial gaps in the background data set. For example are there internal areas which have not been sampled, but which should be included in the background population? Is the existing boundary too limited, and more sampling is required? Or, is the existing boundary too broad, and should some areas with unusual contamination patterns be excluded until more information is available?

If the samples were not all collected using a random or systematic random design within the total area (e.g., some samples were targeted to address questions regarding local sources), then adjustments to the data set are needed to reduce the risk of bias in the summary statistics from over-sampling sub-regions of the population. We recommend two possibilities for this adjustment: 1) adjust the boundary of the background population to be a union of circles surrounding each sampled point. If the circles do not overlap, the samples then comprise a systematic sample of the population thus defined. Or, 2) use a spatial interpolation method (e.g., kriging, or area weighted averaging) to estimate a concentration surface for a larger background boundary. There may not be enough data to accomplish the latter alternative, and the boundary for the former alternative may not be acceptable. In that case, more data must be collected. These two alternatives are discussed in more detail in Section 3.2.3.

3.2.2 Step 2. Examine Trend Characteristics and Autocorrelation of Samples within the Background Data Set

The presence of trends, differences in mean concentrations, and spatial autocorrelation within a data set require special attention. Ignoring trends or autocorrelation can result in biased estimates of population parameters and summary statistics that are not representative of the entire background population.

Spatial autocorrelation is important to identify so that only the independent samples may be used. Clusters of samples that targeted a particular sub-region of the background area should not be allowed to overly emphasize conditions of that sub-region in the description of the entire background area.

If there are concentration trends, or areas with clearly different concentrations (i.e., separate strata), these should be removed prior to estimating autocorrelation (i.e., autocorrelation is estimated from data that exhibit no trend and have a zero mean). In this report, a relatively simplistic approach to evaluating trends is used; an in-depth evaluation and description of a trend surface is beyond the scope of this report. For each case study in Section 4.0, several surface concentration models were used to evaluate potential trends in concentrations. Least squares polynomial surface models of orders 0 to 4 (i.e., from no trend up to a 4th order polynomial) were considered. (The total number of samples for the case studies is 26 to 27, which is probably too few for adequate fitting of the 4th order polynomial model. However, it is considered for illustration purposes.) The five polynomial regression models were compared using Aikake Information Criterion corrected for sample size (AIC_c, Burnham and Anderson, 2002). Note that the AIC_c is based on the maximum likelihood, which is a function of the residuals, and this metric may be

somewhat compromised if autocorrelation is present in the residuals. For final trend models, the process is iterative – trend should be re-evaluated after autocorrelation is removed. Autocorrelation is not expected to have a large effect on the trend evaluation in this context of these case studies, mainly due to limitations of the spatial distribution of the small data sets.

There are many methods for evaluating spatial autocorrelation (e.g., using GIS or other spatial statistics packages). The simple method used here could be done manually in MS Excel, or in R (R Development Core Team, 2011). The case studies do not have a regularly spaced grid of samples, so the boundary of autocorrelation is estimated by evaluating correlation among pairs of points within a certain distance of each other. An autocorrelation boundary can be estimated if there are a reasonable number of points that are close enough together to be autocorrelated. Pairs of sample points are grouped into bins of similar distances. For example, if there are at least six pairs of points within 200m of each other, the distance bins could be 0-200m, then 200-400m, etc. Theoretically, any existing positive autocorrelation would be highest in the first bin. The autocorrelation is estimated by Pearson's linear correlation coefficient between concentrations for all possible station pairs within a distance bin.

The presence of autocorrelation should be tested on the residuals from the best-fit trend model (i.e., the detrended data). The closest distance that could tested for autocorrelation was the smallest distance yielding at least six pairs. In spatial statistics literature, six is a small number of pairs on which to test the autocorrelation (e.g., Journel and Huijbregts, 1978), and is considered to be a bare minimum for a correlation test. For this small sample size, a significance test of the autocorrelation within each distance bin used $\alpha = 0.20$ in order to limit Type II errors (i.e., failing to reject the null hypothesis when autocorrelation is present). This binned hypothesis testing approach is useful given the data limitations (i.e., insufficient pairs of samples at sequentially increasing distances) and the objective of estimating the minimum distance between independent samples. If the data were to be used to estimate a kriged trend surface, then a smoothed autocorrelation function is required, where autocorrelation is described as a continuous function of distance.

3.2.3 Step 3. Calculate Summary Statistics

In the preceding two steps, the valid background samples and the background population boundary have been identified, along with potential trends and autocorrelations within the data set. At this point, we consider two methods for generating appropriate summary statistics in the presence of autocorrelation and/or trends:

Method 1. Adjust the background boundary to be simply the union of the set of independent circles (radius > autocorrelation range) surrounding the existing sampling locations, and treat samples from these independent circles as an independent data set;

or

Method 2. Generate a concentration surface for the defined background boundary and use the surface for generating upper bound estimates (e.g., 95 UCL on the mean and 90/90 UTL).

For *Method 1*, the boundary of the background population is redefined to be exactly the area that has been independently and systematically sampled. This is simply a union of circles around each sampled point, with the radius of the circles greater than or equal to the autocorrelation range. Only one observation within each circle can be used, so that the data set is a uniform independent sample from the defined population. The radius can be larger than the autocorrelation range in order to make the background boundary larger, but this may cause overlapping circles and require subsampling to maintain a uniform and independent systematic sample of the defined population. It also increases the uncertainty, because now a single observation is representative of a larger (unsampled) area, where the concentrations are unknown. If there are unsampled gaps among the sampled locations, these areas are not actually part of the background population as defined. Without information about whether the concentrations in the unsampled gap areas are homogeneous or trending, it is perhaps an overstatement of the available information to assert that the concentrations remain constant within these unsampled areas. Consequently, the background population as defined may not be contiguous.

Method 1 is a fairly simple and reasonable choice if the area that has been sampled reflects an adequate boundary for background. This method requires minimal assumptions, and does not extrapolate beyond, or interpolate between points.

Method 2 may be desirable if the area described by the union of circles around each sampled point is not acceptable, and more samples cannot be collected. For this method, a concentration surface is generated by kriging or another surface contouring method to estimate concentrations for the areas between the sampled points. *Method 2* can be complex and requires more assumptions about the behavior of the data, but it is the only way to estimate concentrations across a broader boundary area. For example, if a large unsampled area is located between two areas of high concentrations, then the surface model would predict concentrations in that area to be similar to neighboring concentrations, rather than simply excluding that area from the described population. Uncertainty in the estimate of the concentrations in interpolated areas reflects prediction error from the model, so the upper bound on the mean for the total background area will quickly increase as you spatially interpolate or extrapolate beyond your data. This approach is not advisable when the data are sparse. Note that the simplest trend surface model would use Thiessen polygons to divide the area into polygons represented by one sampled point. This is a model that assumes constant concentrations within each polygon, and estimates of uncertainty are not readily available (although bootstrapping could be used).

3.2.3.1 Process for Estimating Summary Statistics from Existing Data Using Method 1

The autocorrelation distance d is defined to be the smallest distance between data points for which independence can be assumed. This distance may be derived from site-specific data, results from similar data sets in other areas, or best professional judgment (BPJ). To estimate background summary statistics using *Method 1*:

a. Establish a sample boundary radius $r \geq d$. The larger the radius, the larger the boundary of the background area, but the smaller the overall sample size may be (only one sample per circle is

permitted). There is a tradeoff between a large background area with high uncertainty and a smaller, perhaps non-contiguous background area with lower uncertainty.

b. Identify all samples greater than *r* away from all other samples. These samples are assumed to be independent, random observations given our estimated autocorrelation range. If this captures all of the samples (i.e., only one sample per circle), then the existing data are i.i.d. random samples from the background population and summary statistics can be generated on this dataset, with no need to subset the data. Otherwise:
c. Identify clusters that have two or more samples within *r* of each other. Permute all possible ways of selecting one or more independent samples (>*r* apart) from each cluster.
d. Combine the samples from b) and c) to form all the possible combinations of samples from the data set that are > *r* apart from every other sample. For each of these permutations, generate summary statistics (mean, variance, 95 UCL on the mean, 90th percentile, and 90/90 UTL).
e. Use the distribution of each statistic (e.g., 95 UCL on the mean) to find the best summary statistic for the population (e.g., the maximum 95UCL on the mean will be most likely to capture the true background population mean). Each of the individual UCLs represents a slightly different background boundary, but each is a possible representation of the regional background. The permutation distributions of these summary statistics reflect part of the uncertainty in the background boundary. A highly variable distribution of background 95 UCLs indicates that different background boundaries can produce widely different results, an indicator of small-scale changes in concentration and the need, perhaps, for a re-evaluation of the first steps of the background boundary definitions (i.e., Steps a and b).

3.3 Sampling to Augment Existing Background Data

For existing background data sets, there are likely to be four features that could be improved by additional sampling: 1) temporal consistency; 2) spatial extent; 3) sampling density; and 4) sample size adequacy.

Temporal consistency: Placing new samples in the areas with outdated sample results may be a priority.

Spatial extent: For the spatial extent (boundaries) of the background population, BPJ is required to decide whether the outer boundaries are sufficient to fully capture the target background conditions. Addressing the boundary question uses BPJ and possibly geophysical modeling information about the mechanisms distributing the sediments throughout the area from anthropogenic contributions (e.g., what are the boundaries of influence from non-point source runoff) or natural sources (e.g., what are the boundaries of influence from river sediments).

Sampling density: If the sampling locations are sparsely distributed, then the union of overlapping circles for the point locations sampled may not result in a contiguous background area. Large unsampled or under-sampled areas within the background area should be sampled to reduce uncertainty. Existing data may indicate that trends are present, so sampling at fairly regular intervals along that trend is recommended. Bathymetry and hydrologic flow patterns provide information about sedimentation or disturbance patterns. This information can be used to identify areas within the background population that have potentially different contamination levels, and where additional sampling is recommended.

Overall Sample Size: If more data are needed to satisfy one or more of the three features described above, grid sampling is recommended. The minimum grid size should be *d* (the autocorrelation range). The actual grid size will be determined by budget in most cases, but sample sizes needed for a desired estimate of precision can be estimated using existing data. These calculations, however, assume that additional samples will have the same mean and variance as the existing data, which may not be a valid assumption when sampling from a patchy distribution or a trending surface. At best it provides an informed guess regarding the change in precision of the mean estimate (i.e., width of the 95UCL on the mean) with additional samples.

4.0 Case Study Examples

Ecology provided dioxin/furan TEQ concentrations (ng/kg, dry weight) for three case study sites: Fidalgo Bay, Port Gardner, and Bellingham Bay. For each of these case studies, we illustrate the process of defining the regional background boundaries, evaluating concentration trends in the data, estimating the autocorrelation distance, and using Method 1 (Section 3.2.3.1) to estimate upper bound summary statistics.

4.1 Fidalgo Bay Case Study

Ecology staff used BPJ to identify samples that were too near to point sources, or were from a different area-of-influence and therefore considered to be inappropriate for comparison to Fidalgo Bay Site concentrations. The sampling locations are shown in Figure 1; the grid overlaid on the map is a 0.5 km square grid used simply to illustrate the scale of the distance between samples. The latitude and longitude for the sampling locations and their associated TEQ concentrations (ng/kg, dry weight) are provided in Appendix Table A-1.

4.1.1 Trends and Autocorrelation in Fidalgo Bay

The first step in estimating the autocorrelation range is to remove any existing trends that may be present in the samples. For the Fidalgo Bay data set the first-order polynomial (i.e., a linear trend) fits the trend surface best (indicated by the lowest AICc for all trend surface models considered). There appears to be a linear increasing trend to the northwest (Figure 2).

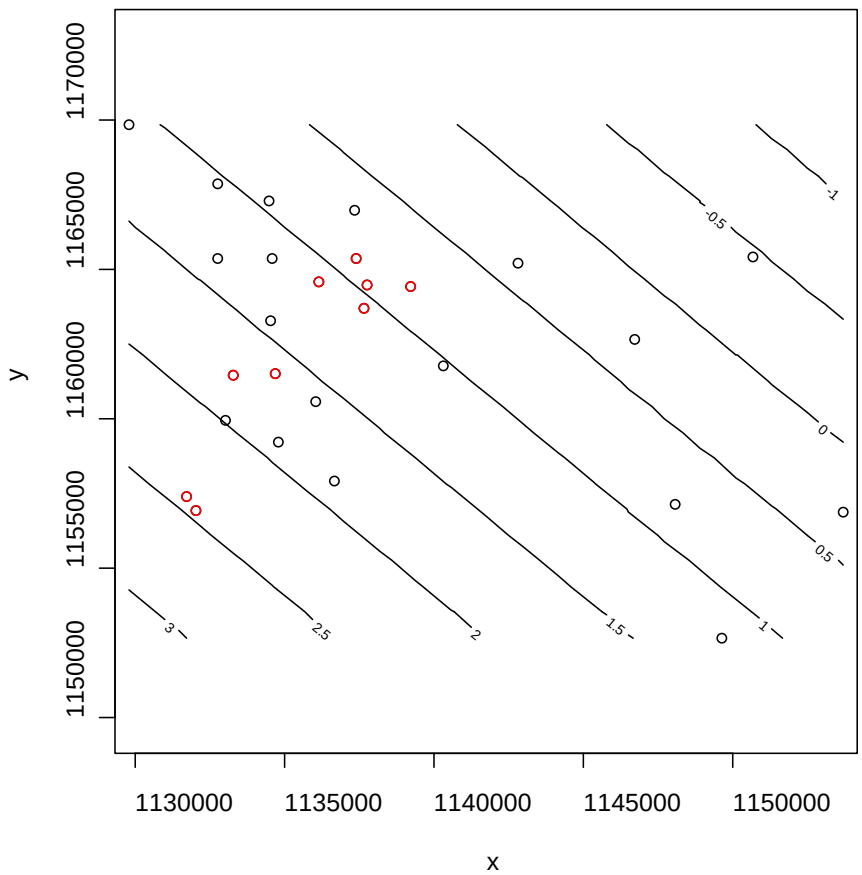

Figure 2. Fidalgo Bay sample locations with contours based on first-order polynomial fit. Samples that are <460m from other samples are highlighted in red.

The maximum distance between the six closest pairs of points in Fidalgo Bay is approximately 460m apart, so bin sizes are set to multiples of 460m. The data points within the first distance bin are highlighted with a different color in Figures 1 and 2. A distance of 460 m is the smallest autocorrelation range that we can test with these data. There may be autocorrelation present at smaller distances, but we cannot test whether the correlation is significant because of insufficient numbers of data pairs available for that distance. The correlation results after removing the linear trend from the data are shown in Table 1.

Table 1. Autocorrelation Results for Fidalgo Bay.

Bin Endpoints (m)	N	Pearson's Correlation Coefficient	one-tailed p-value for parametric test
0-460	6	0.338	0.512
460 – 920	33	0.246	0.168
920-1380	42	0.0894	0.573
1380 – 1840	59	-0.00570	n/a

The autocorrelation coefficient decreases with distance (Table 1). The autocorrelation in the first bin is not significantly different from zero (α=0.20). Sample sizes after the first bin increase dramatically, so it is appropriate to use a lower α-level to assess significance in the second bin (e.g., α=0.10 or 0.05), from which we would conclude that the autocorrelation is not significantly different from zero in this distance range (460 – 920 m). We varied the size of the first bin, and the autocorrelation was evaluated for pairwise distances from zero to increasingly larger maximum distances, up to 914m, but no strong correlations were seen at any of these distances. There are insufficient samples close to each other to estimate autocorrelation less than 460m. There may be autocorrelation at smaller ranges, but it is not a testable hypothesis on this dataset, so we could assume that the data set is roughly uncorrelated at the distances that were sampled. However, we don't want to underestimate autocorrelation, so we also compare the estimation results by subsampling the data (*Method 1,* Section 3.2.3.1) based on a minimum separation distance of approximately 460m (for convenience we round up to 500m).

4.1.2 Method 1 Applied to Fidalgo Bay

Subsets of independent samples were generated from the full data set, as described above. The autocorrelation range used in this approximation was 500m.

There were 27 samples from acceptable non-site affected stations in Fidalgo Bay. Of these 27 samples, 15 were more than 500m away from any other samples. The remaining 12 samples were grouped into three clusters of two or more samples each. All permutations were constructed of independent samples within each of these clusters and combined with the other 15 samples (24 possible permutations). For each of the permutations, summary statistics for TEQ values were generated (i.e., mean, variance, bootstrapped 95 UCL on the mean, and 90/90 UTL for the best fit gamma distribution). The distributions of these TEQ summary statistics are shown in Figure 3. The red lines indicate the values calculated by assuming that the data are uncorrelated at the distances that were sampled (i.e., we have an independent data set). We can see that the full data set (n=27) produced a 95 UCL on the mean (1.55 ng/kg TEQ) that was lower than some of the permutations. For the permutations, 95UCL values range from 1.4 to 1.8 ng/kg TEQ. For the 90/90 UTL, permutation values ranged from 2.5 to 3.6 ng/kg, and the observed data had a TEQ value of 3.1 ng/kg.

4.1.3 Fidalgo Bay Conclusions and Recommendations

The trend surface regression models indicate a linear trend in Fidalgo Bay background concentrations. For this case study, the observed data can be used to estimate background population characteristics as is, or, to ensure independence, after the data have been subsampled to generate a set of independent observations using the autocorrelation range (at 500 m). From the permutation distribution of UCLs, we could choose the maximum value (1.9 ng/kg) as this is the value most likely to capture the true mean, even though the coverage will likely exceed 95%.

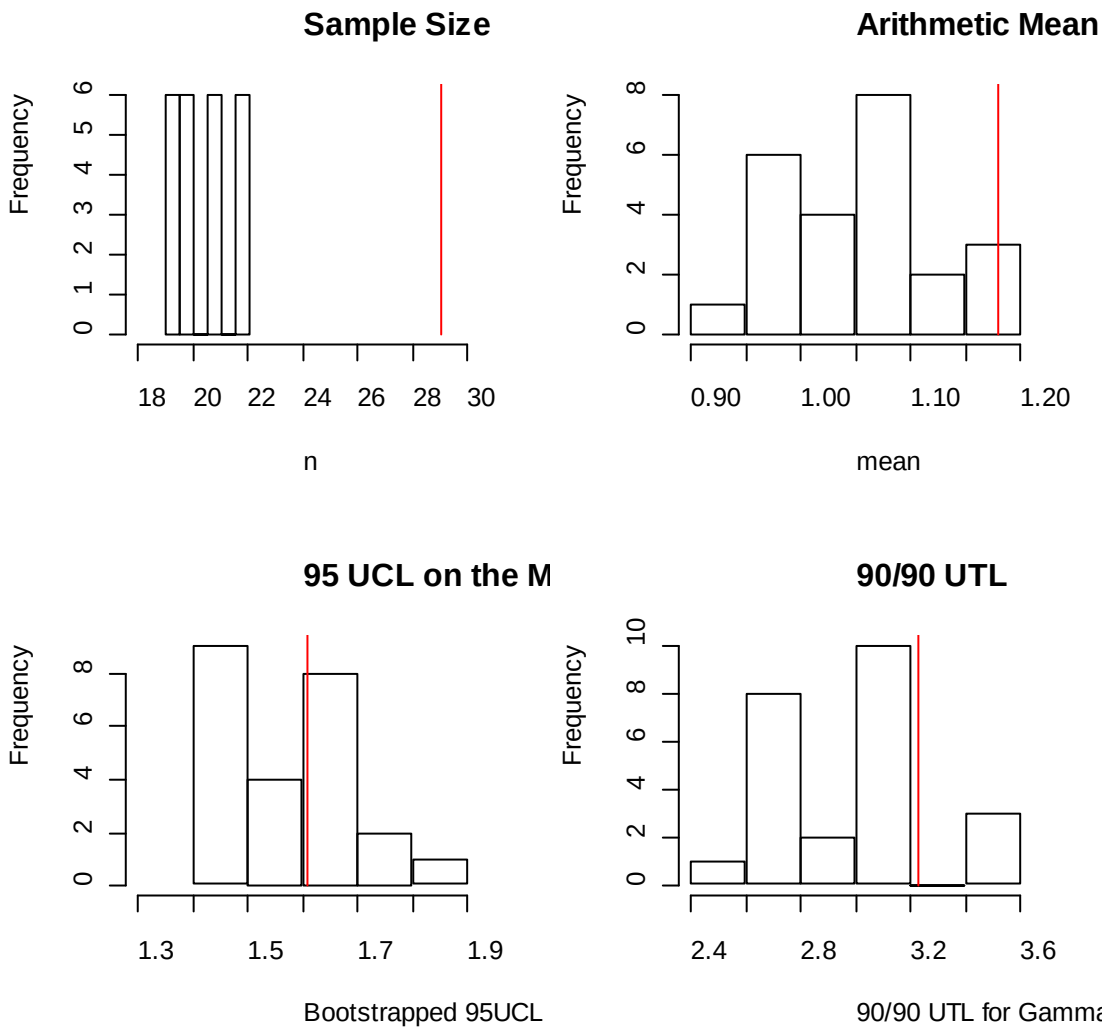

Figure 3. Distribution of summary statistics on TEQ values (ng/kg, dry wt) for the permutations of independent samples (>500m) at Fidalgo Bay. The red lines indicate the values for the observed data set.

For sample adequacy, we consider the following:

Temporal consistency: The data within background areas are collected from 2007 and 2010, so time period is probably not an issue and all these data are useable.

Number of samples: The existing data are not significantly different from a gamma distribution (ProUCL, alpha=0.05). We plot the number of samples vs. the width of the gamma confidence interval on the mean (Figure 4). The figure shows that our sample size of 27 provides a UCL width that is 37% of the mean; we're not on the steepest part of the curve but are on a part of the curve where it's starting to flatten out for our sampled population. Doubling the sample size is expected to decrease the UCL half-width to 24%, assuming that the mean and the variance stay the same. This assumption may not be realistic given that there is a trend in these data – samples collected from a different area will affect

both the mean and the variance, so this graph provides simply a ballpark estimate of expected sample size adequacy.

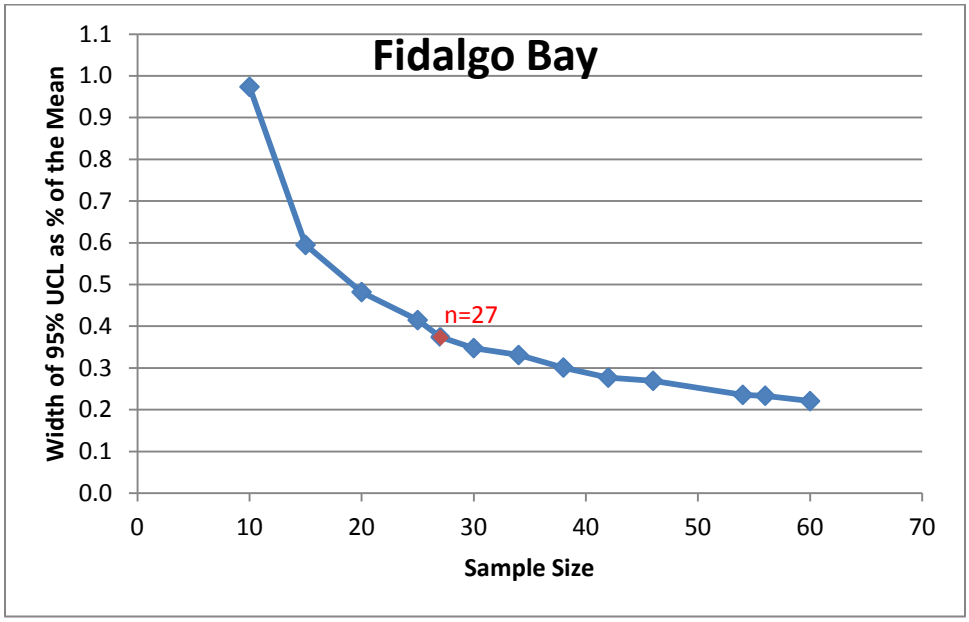

Figure 4. Sample size vs. precision of the mean using Fidalgo Bay data, fit using a gamma distribution.

Spatial Extent: The best trend surface was a 1st order polynomial, decreasing to the northwest away from the Anacortes shoreline. Currently, most of the samples are closer to Anacortes in Fidalgo Bay, so these might describe Regional Background, whereas samples in Padilla Bay describe Local Natural Background, for example. There are rather few samples in Padilla Bay, so if this area is included as part of Regional or Local Natural Background, it would be wise to place more samples on that side of the Bay.

Sampling Density: Based on the autocorrelation range test, we recommend samples no closer than 500m apart (rounding up from 460m). This minimum spacing is expected to achieve independent samples. Any new samples would be placed as evenly as possible within the desired boundaries for the background population, and at least 500m away from any other new or existing samples.

- Option #1: Spatial extent for regional background is a line drawn from Anacortes west to Hat Island and south to March Point (?). Take 5-10 more samples within any of the available grid squares, trying to achieve uniform distribution of samples throughout the area and minimum separation between samples of 500m. Pros: this provides a good spatial coverage within the delineated regional background of Fidalgo Bay. Cons: none, assuming that the boundary for background is sufficient.

- Option #2: Include Padilla bay in the Background characterization. Pursue Option #1, plus additional 5-10 samples from Padilla Bay. Use a grid approach to try to achieve a uniform distribution throughout the area with minimum separation of 500m. Pros: larger background area; provides data to test if Padilla Bay is a separate population. Cons: The higher cost over

Option #1 may not be necessary; the shallower depths and the large Intertidal areas in Padilla Bay may make this an inappropriate background data set for the subtidal sites in Fidalgo Bay.

- Option #3: Include Padilla Bay in the Background characterization, and fill as many squares as possible in both Fidalgo and Padilla Bay, increasing the grid size to 1km (we exchange small scale accuracy for broad scale information). Pros: More information about both areas and a dataset that may be sufficient for drawing a surface contour map. Cons: Cost and potentially sampling overkill.

4.2 Port Gardner Case Study

Ecology staff used BPJ to identify samples that were too near to point sources, or were from a different area-of-influence and therefore considered to be inappropriate for comparison to Port Gardner site concentrations. The sampling locations are shown in Figure 5; the grid overlaid on the map is a 0.5 km square grid used simply to illustrate the scale of the distance between samples. The latitude and longitude for the sampling locations and their associated TEQ concentrations (ng/kg, dry weight) are provided in Appendix Table A-2.

4.2.1 Trend and Autocorrelation in Port Gardner

For the Port Gardner data set, the trend surface is best fit by the third order polynomial (the model with the lowest AICc; Figure 6).

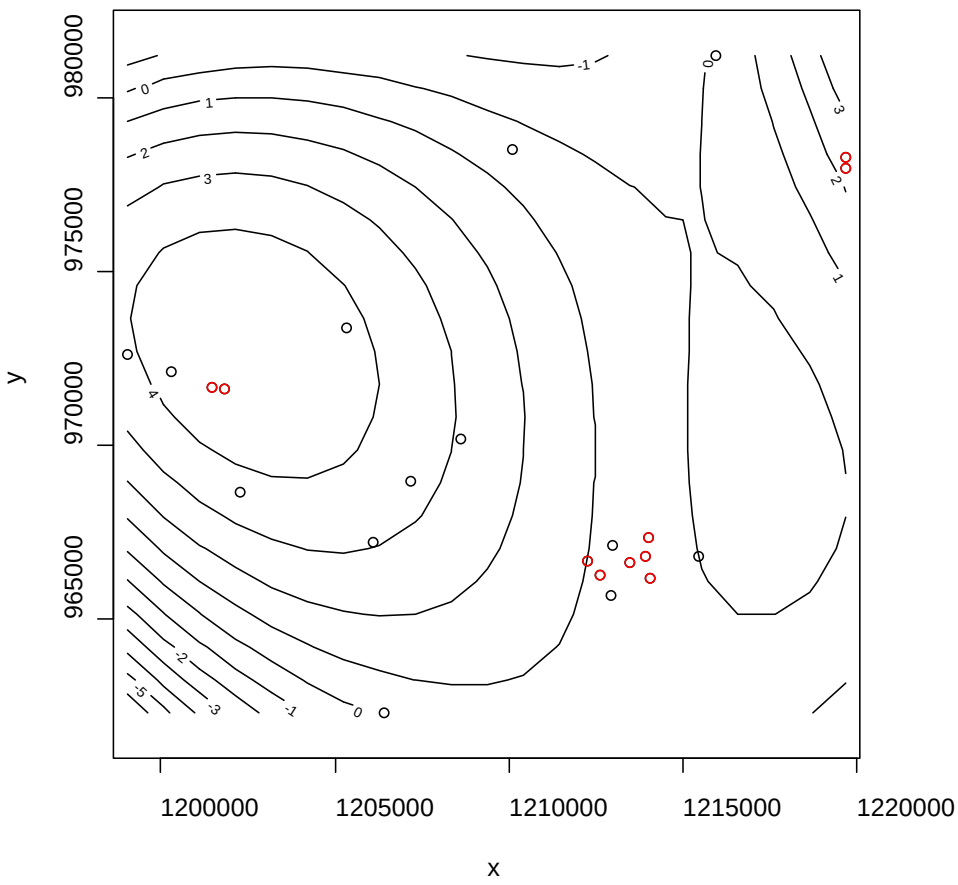

Figure 6. Port Gardner sample locations with contours based on third-order polynomial fit. Samples that are <200 m from other samples are highlighted in red.

The maximum distance between the six closest pairs of points is 200m, so the bin sizes were set to multiples of 200 m. The autocorrelation results are shown in Table 2.

Table 2. Autocorrelation Results for Port Gardner data.

Bin Endpoints (m)	N	Pearson's Correlation Coefficient	one-tailed p-value for parametric test
0-200	6	0.735	0.048
200-400	15	-0.261	n/a
400-600	15	-0.0598	n/a
600-800	5	-0.300	n/a

With the polynomial trend removed, there is evidence that samples within 200m of each other are still correlated. We are limited by the number of samples and the distances among them in our estimation process of the autocorrelation range. The data indicate that samples within 200m of one another should not be treated as independent samples. The minimum distance for independence is at least 200m. Samples between 200m and 400m apart were not correlated, though this appears to be strongly influenced by a single elevated sample in the cluster of stations near Weyerhaeuser (the cluster of stations close to shore in Figure 5).

4.2.2 Method 1 applied to Port Gardner

Subsets of independent samples were generated from the full data set, as described above. The autocorrelation range used in this approximation was 200m. There were 26 samples from acceptable non-site affected stations in Port Gardner. Of these 26 samples, 17 were more than 200m away from any other samples. The remaining nine samples were grouped into four clusters of two or more samples each. All permutations were constructed of independent samples within each of these clusters and combined with the other 17 samples (24 possible permutations).

There were a few samples that were just beyond 200m apart, so we also calculated results for autocorrelation range of 305m. For this range, there were 14 samples more than 305m away from any other samples. The remaining 12 samples were grouped into four clusters of two or more samples each and all permutations were constructed of independent samples within each of these clusters and combined with the other 14 samples (28 possible permutations).

Sample sizes ranged from 18 to 20 for the permutations at 305m distance; and were always 21 for the permutations at 200m distance. Arithmetic mean values were higher for the 305m distance: ranging from 1.86 to 2.01ng/kg TEQ compared to a range of 1.77 to 1.84 ng/kg TEQ for the 200m distance. The distributions of the 95UCL and the 90/90 UTL values for the TEQ of these two sets of permutations are shown in Figure 7. The larger values for the 305m distance partially reflects the smaller sample size (even for an identical distribution, a smaller n generates a larger UCL and UTL because of the greater uncertainty). But the larger values in the 305m distance permutations also indicate the presence of some small scale spatial variability, as observed in the tight cluster of samples near Weyerhaeuser (the cluster of stations close to shore in Figure 5).

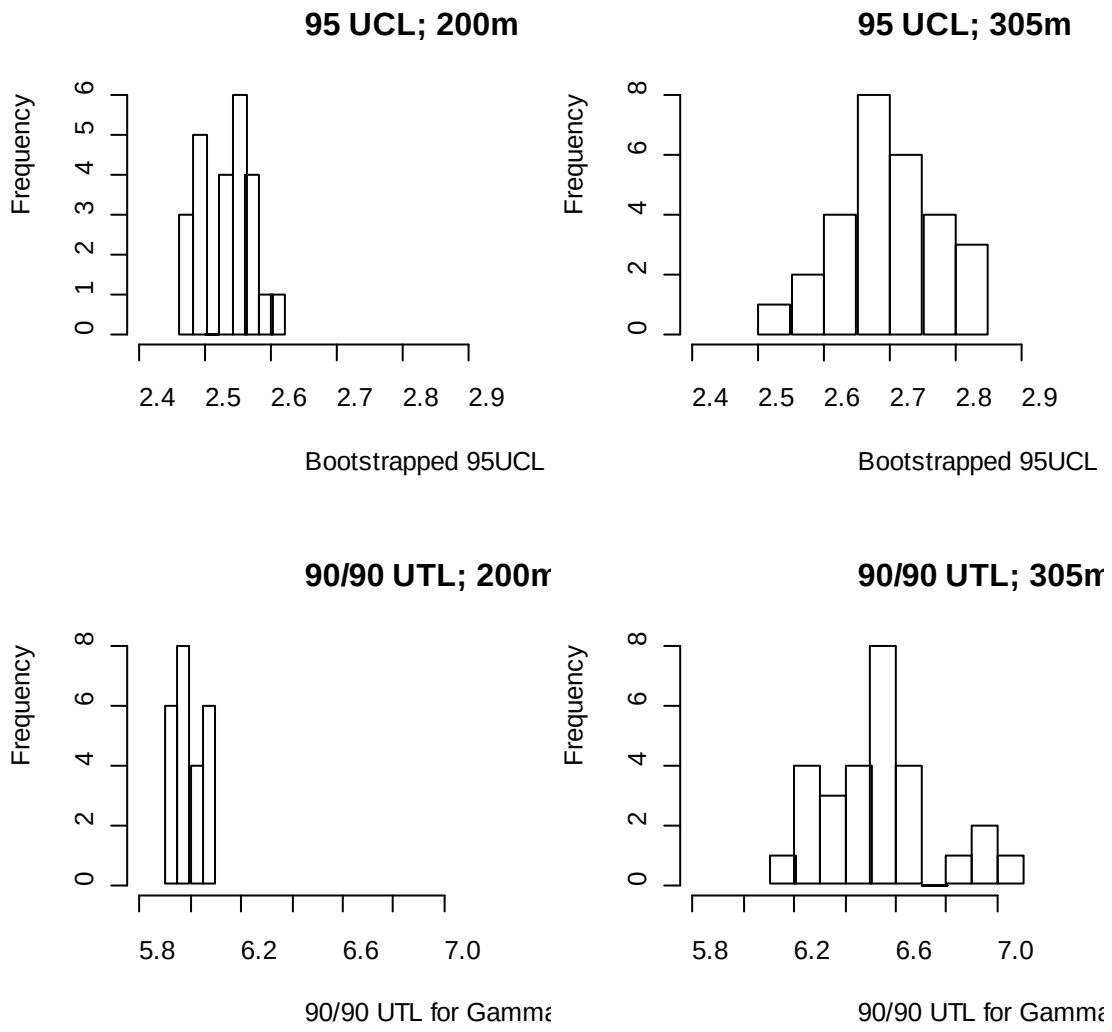

Figure 7. Distribution of 95UCL and 90/90 UTL for TEQ values (ng/kg, dry wt) for the permutations of independent samples in Port Gardner using 200m autocorrelation range (left side), or 305m autocorrelation range (right side).

4.2.3 Port Gardner Conclusions and Recommendations

The autocorrelation investigation and the trend surface regression models indicate that there is trend and/or patchiness in Port Gardner background concentrations of TEQ values (ng/kg, dry weight). Using an independent subset of the existing data will describe the area sampled (the union of the circles around our points sampled), but there are large areas un-sampled with uncertainty about what concentrations might be found there. The differences in the results for the autocorrelation range of 200m vs. 305m indicate the patchy nature of TEQ concentrations in at least some of the areas (although this could only be tested for the dense sample cluster near Weyerhaeuser). This means that interpolation could lead to erroneous conclusions about area averages.

For sample adequacy, we consider the following:

Temporal consistency: The data within background areas are collected from 2004 to 2010. The data from 2004 are all from the Weyerhaeuser sampling, which were also found to be highly influential in the subsetting exercise. Additional sampling in this area may be desirable to describe more current conditions in the area.

Number of samples: The existing data are bimodal, and are not well fit by any parametric distribution (ProUCL). For this data set we can't generate any assumptions about sample size adequacy.

Spatial Extent: The best trend surface was a 3rd order polynomial, indicating patchiness with low concentrations near Weyerhaeuser; higher concentrations near the DMMP disposal site. This site would benefit from additional samples placed out beyond the DMMP disposal site (to see how far out those concentrations extend), and the area between Jetty Island and the disposal site. It may also be desirable to sample the area north and northwest of the current northern boundary of the existing data in order to capture what influence the Snohomish River may have on the bay concentrations.

Sampling Density: Based on the autocorrelation tests, we recommend samples at least 400m – 500m apart. This sampling interval will miss some of the small scale spatial variability that is present, but would allow efficient description of a larger area. In addition, a grid spacing of 500m is expected to achieve independent samples. Any new samples would be placed as evenly as possible within the desired boundaries for the defined background population, and at least 500m away from any other new or existing samples.

- Option #1: Modify the spatial boundary for regional background to exclude everything north to northwest of the line drawn from the southern point of Jetty Island out into Possession Sound. Take at least 10 more samples within any of the available grid squares (500m grid), trying to achieve uniform distribution of samples throughout the area and minimum separation between samples of 500m. Pros: A smaller boundary allows a greater sampling density within the area considered representative of regional background. Excluding the areas on the Snohomish River delta may be justified if the project locations are not heavily influenced by the river. Cons: the spatial boundary may be too limited.

- Option #2: Modify the spatial boundary to exclude the deeper subtidal areas, and include only the areas strongly influenced by the Snohomish River. Pros: A smaller boundary allows a greater sampling density within the area considered representative of local background. Excluding the areas outside of the Snohomish River influence may be justified if the project locations are primarily influenced by the river. Cons: the spatial boundary may be too limited.

- Option #3: Combine Options 1 and 2 to describe a larger background area. Sample in as many grid squares as is affordable, increasing the grid size to 1km (exchange small scale accuracy for broad scale information). Try to achieve a uniform distribution throughout the area and minimum separation between all new and existing data of 0.5-1km. Pros: A broader area is defined that allows the description of an overall background average; if separate populations are present near

the mouth of the Snohomish River vs. subtidal Possession Sound, it may be apparent by these data. Cons: Cost.

4.3 Bellingham Bay Case Study

Ecology staff used BPJ to identify samples that were too near to point sources, or were from a different area-of-influence and therefore considered to be inappropriate for comparison to Bellingham Bay site concentrations. The sampling locations are shown in Figure 8; the grid overlaid on the map is a 0.5 km square grid used simply to illustrate the scale of the distance between samples. The latitude and longitude for the sampling locations and their associated TEQ concentrations (ng/kg, dry weight) are provided in Appendix Table A-3.

4.3.1 Trend and Autocorrelation in Bellingham Bay

For the Bellingham Bay data set, the second-order polynomial provides the best fit (the model with the lowest AICc) for the trend surface model (Figure 9).

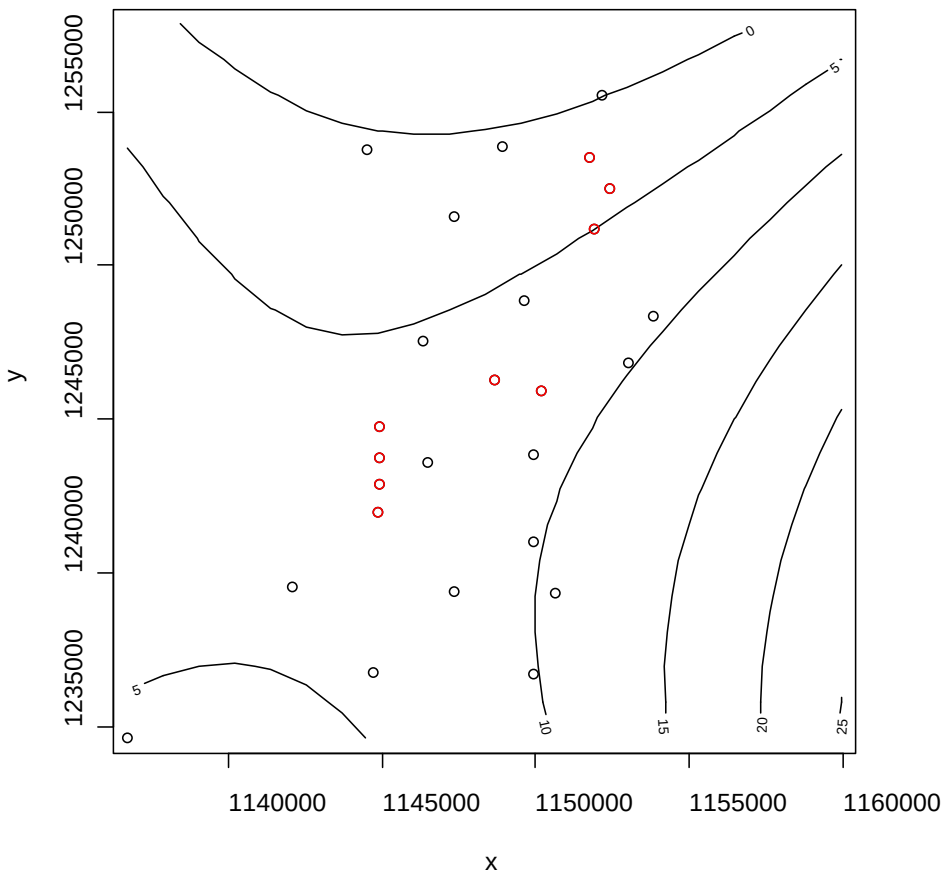

Figure 9. Bellingham Bay sample locations with contours based on second-order polynomial fit. Samples that are <472 m from other samples are highlighted in red.

The maximum distance between the six closest pairs of points is 472m, so the bin sizes are set to multiples of 472m. This is the smallest autocorrelation range that we can test with these data. There may be autocorrelation present at smaller distances, but we cannot test whether the correlation is significant because of insufficient numbers of data pairs available for that distance. The correlation results after removing the trend from the data are shown in Table 3.

Table 3. Autocorrelation results for Bellingham Bay data.

Bin Endpoints (m)	N	Pearson's Correlation Coefficient	one-tailed p-value for parametric test
0 - 472	6	0.315	0.543
472 - 945	23	-0.118	n/a
945 - 1416	41	-0.0704	n/a
1416 - 1890	42	-0.218	n/a

There is no evidence of autocorrelation in the 0-472m range. There may be autocorrelation at smaller ranges, but it is not a testable hypothesis on this dataset, so we could assume that the data set is roughly uncorrelated at the distances that were sampled. However, we don't want to underestimate autocorrelation, so we also test the estimation results by subsampling the data based on a minimum separation distance of 472m (round up to 500m).

4.3.2 Method 1 applied to Bellingham Bay

Subsets of independent samples were generated from the full data set, as described above, using an autocorrelation range of 500m, slightly larger than the smallest autocorrelation range that could be tested.

There were 26 samples from acceptable non-site affected stations in Bellingham Bay. Of these 26 samples, 16 were more than 500m away from any other samples. The remaining ten samples were grouped into three clusters of two or more samples each. All permutations were constructed of independent samples within each of these clusters and combined with the other 16 samples (12 possible permutations). The distributions of summary statistics for these permutations are shown in Figure 10. The red lines indicate the values calculated by assuming that the data are roughly uncorrelated at the distances that were sampled. We can see that the full data set (n=26) produced a lower 95 UCL on the mean (7.3 ng/kg TEQ, dry weight) than some of the random permutations. For the permutations, 95UCL values range from 7.0 to 7.7 ng/kg TEQ. For the 90/90 UTL, permutation values ranged from 14.5 to 15.9, and the observed data had a value of 14.6 ng/kg TEQ.

One of the clusters of samples had substantial variability in the reported TEQ values: 1.5, 1.6, and 6.3 ng/kg. The first two concentrations were reported for samples from the 0-12cm horizon; the last for a sample from the 0-55cm horizon. This was the only sample included in this background data set that was collected beyond the 0-12cm depth horizon. Permutation results excluding this deeper horizon sample had

95 UCL values ranging from 7.0 to 7.4 ng/kg, and 90/90 UTL values ranging from 14.5 to 15.3 ng/kg. So, this sample definitely had an effect on the upper range of the estimates, but it's unknown whether it reflects greater contamination at depth or overall small scale spatial variability in surface concentrations.

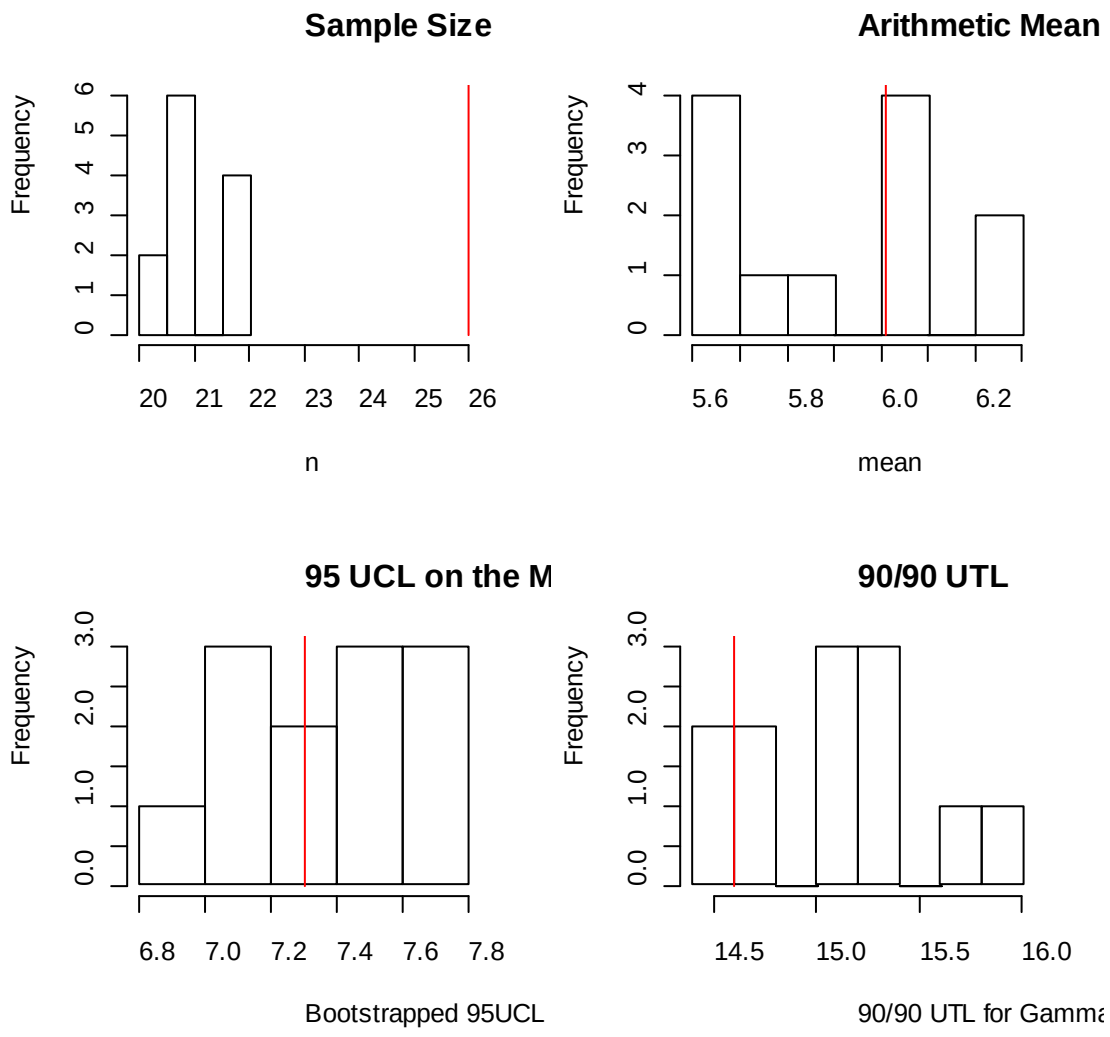

Figure 10. Distribution of summary statistics TEQ values (ng/kg, dry weight) for the permutations of independent samples (>500m) at Bellingham Bay. Red lines indicate the values for the observed data set.

4.3.3 Bellingham Bay Conclusions and Recommendations

The trend surface regression models indicate that there is significant trend in Bellingham Bay background concentrations. The autocorrelation investigation indicated that the data could be considered independent as sampled, but there are large areas un-sampled with uncertainty about what concentrations might be found there.

Given the strong appearance of trends, and potentially two competing trends (one from the southern shoreline of Bellingham Bay, and another from the Nooksack River in the north), this site would benefit from additional samples. Where the boundary is drawn depends on BPJ regarding the relevance of the Nooksack River influence on the Regional Background concentrations. An independent sampling interval would be 500m, but the large area that needs to be sampled justifies using a larger sampling interval (e.g., 1000m) if needed.

For sample adequacy, we consider the following:

Temporal consistency: The data within the background area are collected from 2007 to 2010, so time period is probably not an issue and all these data are useable.

Number of Samples: The existing data are not significantly different from either the normal or the gamma distributions (ProUCL, alpha=0.05). For this sample size calculation, we use the gamma distribution because it allows for more potential skewness in the distribution and a more conservative sample size calculation. We plot the number of samples vs. the width of the gamma confidence interval on the mean (Figure 11). The figure shows that our sample size of 26 provides a UCL width that is 32% of the mean for the area sampled; we're close to the part of the curve where it's starting to flatten out for our sampled population indicating incrementally smaller advantage from each additional sample. Doubling the sample size is expected to decrease the UCL half-width to about 20%, assuming that the mean and the variance stay the same. This assumption may not be realistic given that there is a trend in these data – samples collected from a different area will affect both the mean and the variance, so this graph provides simply a ballpark estimate of expected sample size adequacy.

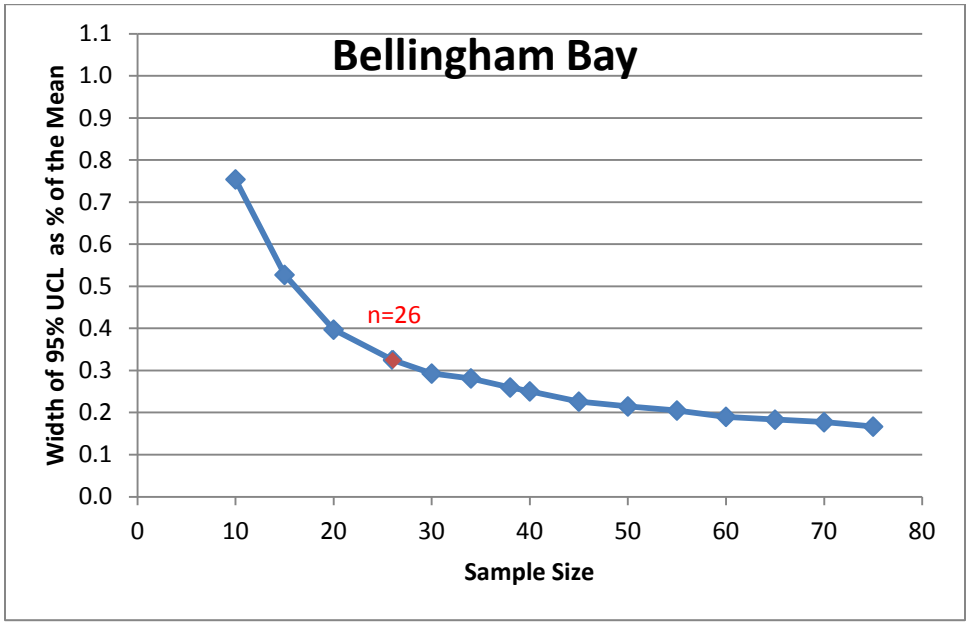

Figure 11. Sample size vs. precision of the mean using Bellingham Bay data, fit with a gamma distribution.

Spatial Extent: The best trend surface was a 2nd order polynomial, showing a strong trend decreasing away from the Bellingham shoreline along the SE portion of the Bay, and a weaker trend that

decreases approaching the northern portion of the Bay and the Nooksack River delta. Where the boundary for background is drawn depends on BPJ regarding the relevance of the Nooksack River influence on the Regional Background concentrations.

Sampling Density: Based on the autocorrelation tests, we recommend samples no closer than 0.5km apart to get a data set of independent samples.

- Option #1: Draw the spatial boundary to be just outside of the existing sampling locations, and take an additional 10 samples within any of the available 500m grid squares within that boundary. As much as possible try to achieve spatial evenness, and a minimum separation of 500m. Pros: Maximizes the use of the existing data and fills some data gaps for this background boundary. Cons: If the northern area influenced by the Nooksack River is more of a local natural background, this data set will be a combination of two blending populations.

- Option #2: Draw the spatial boundary to exclude some of the existing locations in the northern portion of the bay where the Nooksack River may be influencing concentrations. Take an additional 10 samples within any of the available 500m grid squares within this area, trying to achieve spatial evenness and a minimum separation of 500m. Pros: Same cost as Option 1, but a smaller boundary allows a greater sampling density within the area considered representative of regional background. Excluding the areas of the bay with strong Nooksack River influence may be justified if the project locations are more strongly regionally influenced similar to what's found in the southern portion of the Bay. Cons: the spatial boundary may be too limited (encompasses an area generally within 3km of the shoreline) and therefore may be focused too much on the upper concentration end of the trend.

- Option #3: Draw the spatial boundary further out into the Bay to try to identify where the two trends meet. The grid size could be increased to 1km (exchange small scale accuracy for broad scale information). Try to achieve a uniform distribution throughout the area and minimum separation of 0.5-1km. Pros: A broader area is defined that allows better understanding of the two local influences (i.e., the river and the urban area), and therefore a better description of an overall background average. Cons: Cost, and sampling overkill if the regional background is what's needed for project comparison.

5.0 General Recommendations

The approach used by agencies (US EPA, OR DEQ, WA DOE, and ACOE) for describing background involves the initial definition of the population. Given a narrative description of Regional Background, or Local Natural Background, the spatial boundaries for the appropriate background are a site-specific question and must be drawn using existing data from the area, information about fate and transport of contamination from the site, regional influences, as well as best professional judgment.

Once the data within the presumed background area have been compiled, it is fairly simple to fit a selection of modeled trend surfaces, and look for autocorrelation in the residuals. The trend surface provides some information about spatial variability and local patchiness in the concentrations, which can assist in helping determine the best locations for additional sampling. The presence of observable trends also indicates that a random sample (or even a random subset of the existing data) may inadequately describe background areas that have not been uniformly sampled. Treating a non-uniform sample from a trending population as if it were i.i.d. can result in biased estimates of the mean and the 95UCL on the mean. Existing data may be insufficient to detect the presence of a trend (i.e., just because we don't detect a trend doesn't mean it's not there). The optimal sampling design in the presence of trends is one that achieves spatial evenness. Since there is no harm in assuming that trends are present, the optimal design should always attempt to achieve uniform spatial coverage using systematic random samples, or more complex designs such as GRTS sampling (US EPA 2011).

The autocorrelation investigation helps uncover the magnitude of spatial autocorrelation in existing background data sets. Using the autocorrelation range estimate, we can assume that samples outside of this distance range can be considered effectively independent. The sampling density of the existing data set may be insufficient to measure the true autocorrelation range, but it should be sufficient to provide an approximate minimum separation distance to define the grid size for future sampling. If the existing data set does contain dense clusters of samples, then the autocorrelation investigation allows those data to be treated appropriately for the calculation of background summary statistics by selecting independent subsets of the data (i.e., *Method 1*).

Finally, we emphasize that *Method 1* described herein only allows description of the areas that are directly sampled (the union of circles around the sampled data points). Any extrapolation beyond, or interpolation between the sampled locations is avoided. Spatial modeling (*Method 2*) can provide estimates for interpolated concentrations across a broader area but has the disadvantages that it can be complex, requires more assumptions about the behavior of the data, and cannot be done adequately when the data are sparse.

6.0 References

DMMP. 2011. "DMMP Background 8.4.3.pdf". Excerpt from an unknown Puget Sound DMMP document, provided by Chance Asher, WDOE, June 2011.

Journel A. G. and Ch. J. Huijbregts. 1978. Mining Geostatistics. Academic Press.

NewFields. 2011. Memo to Laura Inouye, Department of Ecology, "Approaches for Determining Background in Sediments - Case Study Review," 30 June 2011.

R Development Core Team. 2011. "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

U.S. EPA. 1995. Establishing Background Levels. Quick Reference Fact Sheet. EPA/540/F-94/030, OSWER 9285.7-19FS. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington DC. http://www.epa.gov/superfund/sites/npl/hrsres/fact/bglevels.pdf

U.S. EPA. 2002. Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites. EPA 540/R-01/003, OSWER 9285.7-41. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC.

U.S. EPA. 2010. ProUCL Version 4.00.05 User Guide (Draft). U.S. Environmental Protection Agency, Office of Research and Development, Washington D.C. EPA/600/R-07/041. http://www.epa.gov/esd/tsc/ProUCL_v4.00.05/ProUCL_v4.00.05_tech_guide%28draft%29.pdf

U.S. EPA. 2011. *Aquatic Resources Monitoring. Specific Design Information – Illustrative Examples.* http://www.epa.gov/nheerl/arm/designing/design_intro.htm.

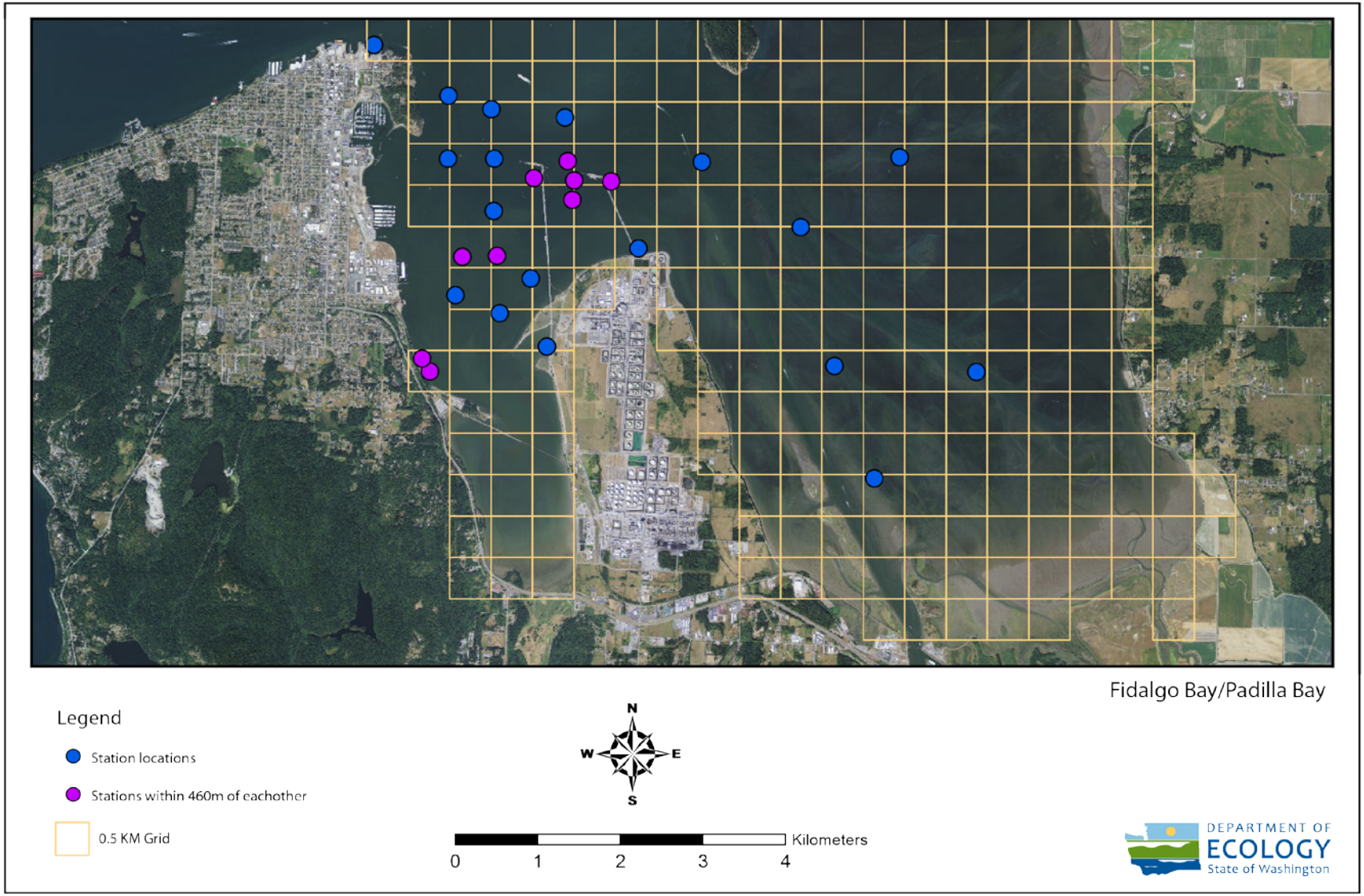

Figure 1. Map of Fidalgo Bay case study site, showing locations of existing data.

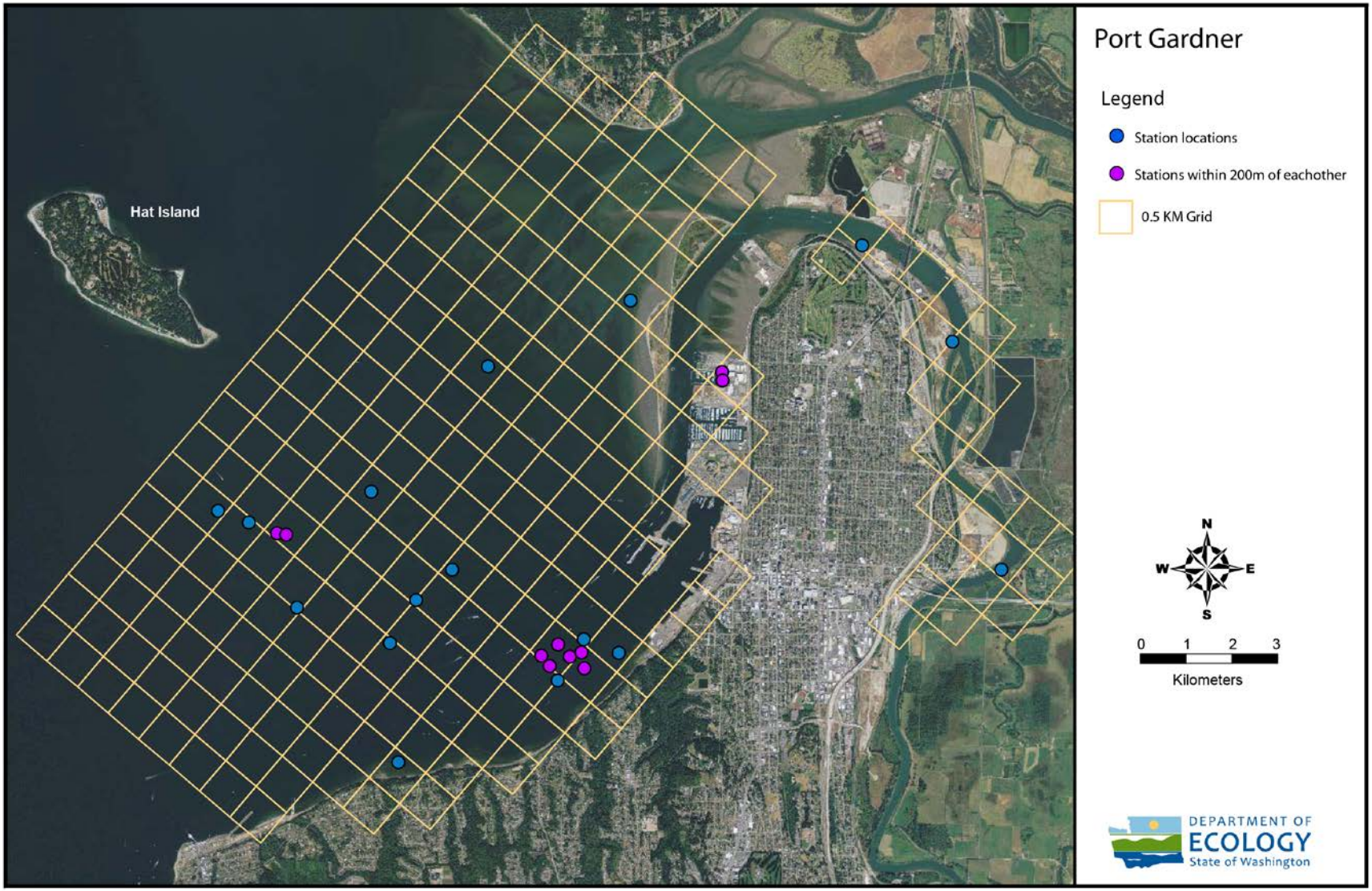

Figure 5. Map of Port Gardner case study site, showing locations of existing data.

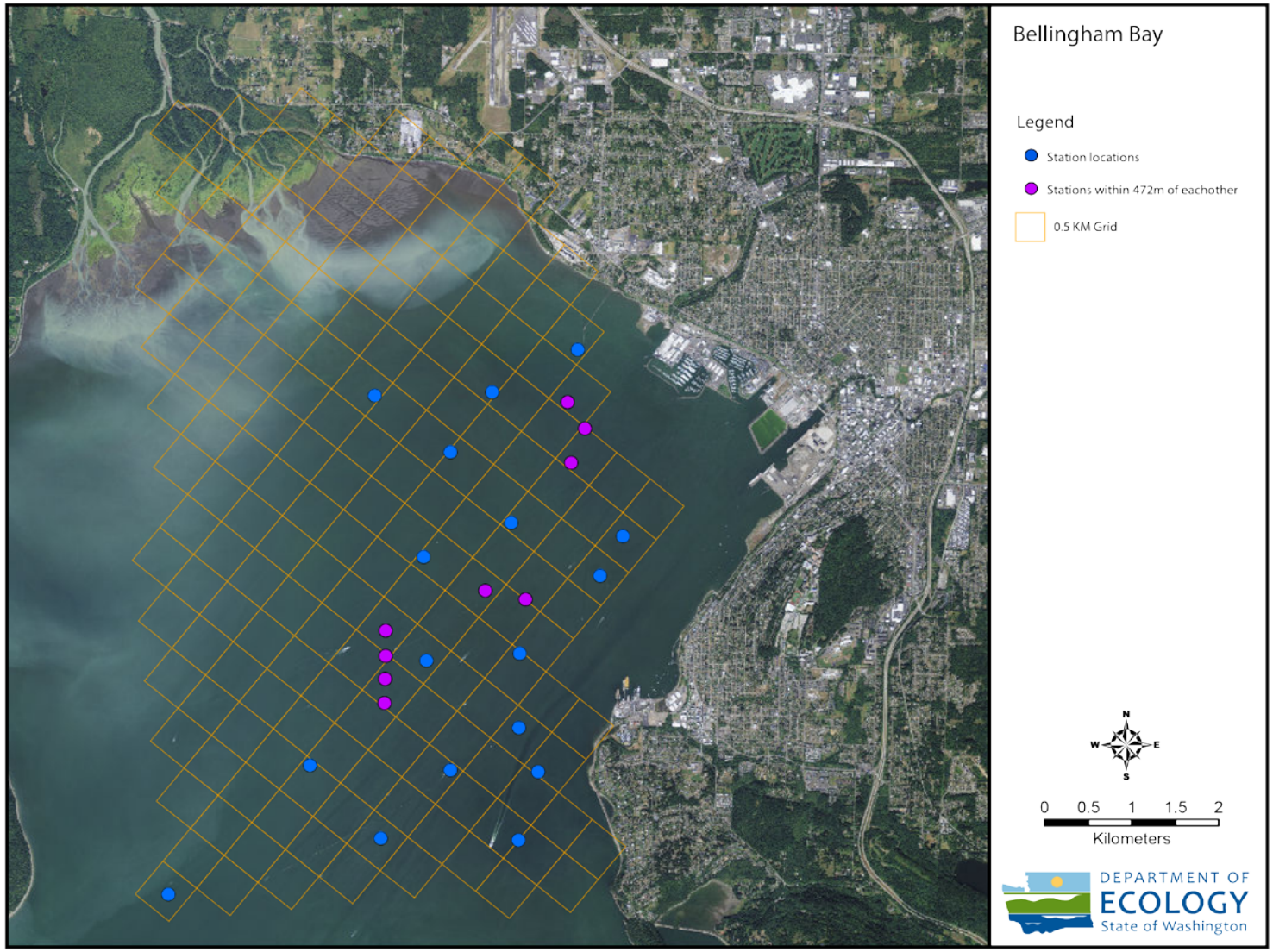

Figure 8. Map of Bellingham Bay case study site, showing locations of existing data.

Table A-1. Fidalgo Bay Case Study TEQ Data for Marine Sediments used as Regional Background

StudyID	LocationID	Study Location Name	FieldActivity StartDate	SampleID	Latitude	Longitude	Study_Type	Location_Setting	TEQ Conc (pptr)
FBCPDX48	FSID6858-PB-10	PADILLABAY-10	6/8/2010	SDS-PB-10	48.476283	-122.5225	SiteInvestigation	Intertidal	0.56
FBCPDX48	FSID6858-CT-05	CLAMCOLLECTIONSITE-05	6/14/2010	SDS-CT-05	48.488004	-122.5969	SiteInvestigation	Intertidal	1.3
FIDALG08	FB-A3-42	FB-A3-42	9/4/2007	FB-A3-42	48.486639	-122.5956	InitialInvestigation	Estuary	3.4
FBCPDX48	FSID6858-PB-08	PADILLABAY-08	6/8/2010	SDS-PB-08	48.488383	-122.5295	SiteInvestigation	Intertidal	0.24
FBCPDX48	FSID6858-PB-09	PADILLABAY-09	6/8/2010	SDS-PB-09	48.48815	-122.5063	SiteInvestigation	Intertidal	0.13
FBCPDX48	FSID6858-FB-10	FIDALGOBAY-10	6/8/2010	SDS-FB-10	48.493153	-122.5844	SiteInvestigation	Intertidal	2
FIDALG08	FB-A3-41	FB-A3-41	9/4/2007	FB-A3-41	48.4897	-122.5766	InitialInvestigation	Intertidal	3.8
FBCPDX48	FSID6858-CPD-12	CUSTOMPLYWOODMILL-12	6/10/2010	SDS-CPD-12	48.495	-122.5918	SiteInvestigation	Intertidal	1.7
FIDALG08	FB-A2-38	FB-A2-38	8/30/2007	FB-A2-38	48.497	-122.5795	InitialInvestigation	Subtidal	1.9
FIDALG08	FB-A3-25	FB-A3-25	9/5/2007	FB-A3-25	48.4992	-122.5908	InitialInvestigation	Subtidal	1.9
FBCPDX48	FSID6858-FB-09	FIDALGOBAY-09	6/8/2010	SDS-FB-09	48.4994	-122.5851	SiteInvestigation	Intertidal	1.3
FIDALG08	FB-A2-35	FB-A2-35	8/30/2007	FB-A2-35	48.5006	-122.562	InitialInvestigation	Intertidal	0.72
FBCPDX48	FSID6858-PB-07	PADILLABAY-07	6/7/2010	SDS-PB-07	48.503367	-122.5356	SiteInvestigation	Intertidal	0.22
FBCPDX48	FSID6858-FB-07	FIDALGOBAY-07	6/8/2010	SDS-FB-07	48.504233	-122.5859	SiteInvestigation	Intertidal	0.79
FBCPDX48	FSID6858-FB-08	FIDALGOBAY-08	6/8/2010	SDS-FB-08	48.50565	-122.573	SiteInvestigation	Intertidal	0.67
FBCPDX48	FSID6858-FB-04	FIDALGOBAY-04	6/8/2010	SDS-FB-04	48.509783	-122.5935	SiteInvestigation	Intertidal	1.1
FBCPDX48	FSID6858-FB-05	FIDALGOBAY-05	6/8/2010	SDS-FB-05	48.509917	-122.586	SiteInvestigation	Intertidal	0.41
FBCPDX48	FSID6858-FB-06	FIDALGOBAY-06	6/8/2010	SDS-FB-06	48.510183	-122.5742	SiteInvestigation	Intertidal	0.33
FIDALG08	FB-A2-03	FB-A2-03	8/30/2007	FB-A2-03	48.5079	-122.5794	InitialInvestigation	Subtidal	1.8
FIDALG08	FB-A2-06	FB-A2-06	8/30/2007	FB-A2-06	48.5078	-122.5668	InitialInvestigation	Subtidal	1.4
FIDALG08	FB-A2-05	FB-A2-05	8/31/2007	FB-A2-05	48.5078	-122.5728	InitialInvestigation	Subtidal	2.7
FBCPDX48	FSID6858-PB-05	PADILLABAY-05	6/7/2010	SDS-PB-05	48.510167	-122.552	SiteInvestigation	Intertidal	0.57
FBCPDX48	FSID6858-PB-06	PADILLABAY-06	6/7/2010	SDS-PB-06	48.5112	-122.5197	SiteInvestigation	Intertidal	0.13
FBCPDX48	FSID6858-FB-02	FIDALGOBAY-02	6/8/2010	SDS-FB-02	48.515283	-122.5866	SiteInvestigation	Intertidal	0.51
FBCPDX48	FSID6858-FB-03	FIDALGOBAY-03	6/8/2010	SDS-FB-03	48.514567	-122.5746	SiteInvestigation	Intertidal	0.46
FBCPDX48	FSID6858-FB-01	FIDALGOBAY-01	6/8/2010	SDS-FB-01	48.516633	-122.5937	SiteInvestigation	Intertidal	0.31
FIDALG08	FB-A4-20	FB-A4-20	9/5/2007	FB-A4-20	48.5219	-122.6061	InitialInvestigation	Subtidal	1.4

Dioxin data downloaded from EIM. TEQs were calculated using TEFs from WAC Tables, found at: http://apps.leg.wa.gov/wac/default.aspx?cite=173-340-900. TEQs were calculated using substitution of non-detects at one-half the detection limit.

Table A-2. Port Gardner Case Study TEQ Data for Marine Sediments used as Regional Background

StudyID	LocationID	Study Location Name	FieldActivity StartDate	SampleID	Latitude	Longitude	Location_Setting	TEQ Conc (pptr)
DMMP_Dioxin_2005-07	DMMP-PGT15	PGT15	6/29/2006	PGT15-A	47.98630	-122.3020	SUBTIDAL	4.30
DMMP_Dioxin_2005-07	DMMP-PGT13	PGT13	6/29/2006	PGT13-A	47.98505	-122.2968	SUBTIDAL	4.20
DMMP_Dioxin_2005-07	DMMP-PGT11	PGT11	6/29/2006	PGT11-A	47.98392	-122.2921	SUBTIDAL	4.40
DMMP_Dioxin_2005-07	DMMP-PGP08_1	PGP08_1	6/29/2006	PGP08_10cm	47.98380	-122.2905	SUBTIDAL	3.90
DMMP_Dioxin_2005-07	DMMP-PGP07_1	PGP07_1	6/30/2006	PGP07_10cm	47.97562	-122.2885	SUBTIDAL	3.80
DMMP_Dioxin_2005-07	DMMP-PGP01_1	PGP01_1	6/30/2006	PGP01_10cm	47.98880	-122.2765	SUBTIDAL	5.00
DMMP_Dioxin_2005-07	DMMP-PGB01	PGB01	6/29/2006	PGB01_10cm	47.97192	-122.2728	SUBTIDAL	3.40
PortGardner_08	A1-46B	A1-46B	9/4/2008	A1-46B-S	47.95856	-122.2710	ESTUARY	0.18
DMMP_Dioxin_2005-07	DMMP-PGP09_1	PGP09_1	6/29/2006	PGP09_10cm	47.97679	-122.2686	SUBTIDAL	3.20
DMMP_Dioxin_2005-07	DMMP-PGB09_1	PGB09_1	6/30/2006	PGB09_10cm	47.98029	-122.2627	SUBTIDAL	3.00
PortGardner_08	A2-02	A2-02	9/4/2008	A2-02-S	48.00314	-122.2575	ESTUARY	0.18
KIMCLK04	KIMCLK04AKC-7	AKC-7	2/26/2004	AKC-7SD	47.97088	-122.2476	Subtidal	0.66
KIMCLK04	KIMCLK04AKC-3	AKC-3	2/26/2004	AKC-3SD	47.96972	-122.2461	Subtidal	0.28
KIMCLK04	KIMCLK04AKC-5	AKC-5	2/26/2004	AKC-5SD	47.97218	-122.2447	Subtidal	0.51
KIMCLK04	KIMCLK04AKC-6	AKC-6	2/26/2004	AKC-6SD	47.96814	-122.2447	Subtidal	1.20
KIMCLK04	KIMCLK04AKC-2	AKC-2	2/26/2004	AKC-2SD	47.97086	-122.2427	Subtidal	0.61
KIMCLK04	KIMCLK04AKC-1	AKC-1	2/26/2004	AKC-1SD	47.97131	-122.2408	Subtidal	0.42
KIMCLK04	KIMCLK04AKC-8	AKC-8	2/26/2004	AKC-8SD	47.97282	-122.2405	Subtidal	0.72
KIMCLK04	KIMCLK04AKC-4	AKC-4	2/26/2004	AKC-4SD	47.96955	-122.2403	Subtidal	0.21
PortGardner_08	A1-31B	A1-31B	9/4/2008	A1-31B-S	47.97136	-122.2346	ESTUARY	0.18
PortGardner_08	A2-08	A2-08	9/4/2008	A2-08-S	48.01088	-122.2340	ESTUARY	0.26
AODE6677	AO6677-462.1	A/H-SED-1	12/10/2010	SED-1	48.00310	-122.2184	Subtidal	2.55
AODE6677	AO6677-465	A/H-SED-4	12/10/2010	SED-4	48.00217	-122.2183	SUBTIDAL	2.10
PortGardner_08	A2-30	A2-30	9/12/2008	A2-30-S	48.01762	-122.1954	ESTUARY	0.42
PortGardner_08	A2-32	A2-32	9/4/2008	A2-32-S	48.00702	-122.1800	ESTUARY	0.16
PortGardner_08	A2-37B	A2-37B	9/4/2008	A2-37B-S	47.98164	-122.1710	ESTUARY	0.18

Dioxin data downloaded from EIM. TEQs were calculated using TEFs from WAC Tables, found at: http://apps.leg.wa.gov/wac/default.aspx?cite=173-340-900. TEQs were calculated using substitution of non-detects at one-half the detection limit.

Table A-3. Bellingham Bay Case Study TEQ Data for Marine Sediments used as Regional Background

StudyID	LocationID	Study Location Name	FieldActivity StartDate	SampleID	Upper Depth (cm)	Lower Depth (cm)	Latitude	Longitude	Setting	TEQ Conc (pptr)
DMMP_Dioxin_2005-07	DMMP-BBB04	BBB04	7/19/2007	BBB04	0	10	48.6998	-122.5846	Subtidal	4.3 J
DMMP_Dioxin_2005-07	DMMP-BBP04	BBP04	7/19/2007	BBP04	0	10	48.7137	-122.5631	Subtidal	5.2 J
DMMP_Dioxin_2005-07	DMMP-BBP01	BBP01	7/19/2007	BBP01	0	10	48.7062	-122.5517	Subtidal	5.5 J
DMMP_Dioxin_2005-07	DMMP-BBP02	BBP02	7/19/2007	BBP02	0	10	48.7136	-122.5411	Subtidal	8.5 J
BELSEDDF	BBDIOX-10	BBDIOX-10	6/9/2010	BBDIOX-10	0	12	48.7064	-122.5303	Subtidal	11 J
DMMP_Dioxin_2005-07	DMMP-BBT05	BBT05	7/20/2007	BBT05	0	10	48.7254	-122.5517	Subtidal	7.2 J
DMMP_Dioxin_2005-07	DMMP-BBP03	BBP03	7/19/2007	BBP03	0	10	48.7204	-122.5517	Subtidal	7 J
DMMP_Dioxin_2005-07	DMMP-BBT04	BBT04	7/19/2007	BBT04	0	10	48.7230	-122.5517	Subtidal	7 J
DMMP_Dioxin_2005-07	DMMP-BBT06	BBT06	7/20/2007	BBT06	0	10	48.7281	-122.5518	Subtidal	6.8 J
BELSEDDF	UWI 32	UWI 32	6/10/2010	UWI 32	0	12	48.7250	-122.5453	Subtidal	2.6 J
BELSEDDF	BBDIOX-11	BBDIOX-11	6/9/2010	BBDIOX-11	0	12	48.7182	-122.5307	Subtidal	6.7 J
DMMP_Dioxin_2005-07	DMMP-BBB02	BBB02	7/20/2007	BBB02	0	10	48.7136	-122.5275	Subtidal	10 J
BELSEDDF	BBDIOX-9	BBDIOX-9	6/10/2010	BBDIOX-9	0	12	48.7260	-122.5309	Subtidal	10 J
BELSEDDF	UWI 277	UWI 277	6/9/2010	UWI 277	0	12	48.7359	-122.5462	Subtidal	5.7 J
Bellinghambay08	HART17_BBDXSS05	Bellingham Bay Dioxin BBDx-SS-05	9/18/2008	BBDX-SS-05	0	12	48.7326	-122.5365	Subtidal	12
BELSEDDF	BBDIOX-1A	BBDIOX-1A	6/15/2010	BBDIOX-1A	0	12	48.7317	-122.5302	Subtidal	11 J
BELSEDDF	BBDIOX-3A	BBDIOX-3A	6/15/2010	BBDIOX-3A	0	12	48.7527	-122.5545	Subtidal	0.57 J
BELSEDDF	BBDIOX-4	BBDIOX-4	6/11/2010	BBDIOX-4	0	12	48.7470	-122.5425	Subtidal	1.7 J
BELSEDDF	BBDIOX-6	BBDIOX-6	6/11/2010	BBDIOX-6	0	12	48.7397	-122.5327	Subtidal	3 J
Bellinghambay08	HART17_BBDXSS04	Bellingham Bay Dioxin BBDx-SS-04	9/19/2008	BBDX-SS-04	0	12	48.7344	-122.5186	Subtidal	13 J
BELSEDDF	UWI 29	UWI 29	6/9/2010	UWI 29	0	11	48.7386	-122.5153	Subtidal	6 J
BELSEDDF	UWI 35	UWI 35	6/11/2010	UWI 35	0	10	48.7534	-122.5363	Subtidal	1.4 J
BELSEDDF	BBDIOX-5	BBDIOX-5	6/15/2010	BBDIOX-5	0	12	48.7462	-122.5236	Subtidal	1.6 J
DMMP O&M Squalicum	Sq-15	Sq-15	9/7/2010	Sq-15	0	55	48.7498	-122.5216		6.29 J
Bellinghambay08	HART17_BBDXSS01	Bellingham Bay Dioxin BBDx-SS-01	9/19/2008	BBDX-SS-01	0	12	48.7526	-122.5244	Subtidal	1.5
BELSEDDF	BBDIOX-2	BBDIOX-2	6/10/2010	BBDIOX-2	0	9	48.7581	-122.5231	Intertidal	0.7 J

Dioxin data downloaded from EIM. TEQs were calculated using TEFs from WAC Tables, found at: http://apps.leg.wa.gov/wac/default.aspx?cite=173-340-900.
TEQs were calculated using substitution of non-detects at one-half the detection limit.