Data Analysis Report: Baseline Assessment of Selenium Concentrations at Benton Lakes NWR (2013-2014)

March 10, 2015

Prepared by Jay Rosencrantz, Kathryn M. Irvine PI

This statistical report is in support of Benton Lakes NWR monitoring efforts, and it was supported by the Interagency agreement IAA No. 4500036892 between USGS PI. K. Irvine and Vanessa Fields USFWS. This report is for internal purposes only and is not for external distribution.

Contents

1 Data Collection

1.1 Data Collection in 2013

Surveys were conducted in the inter-unit canal and units 2, 4C, 5, and 6 during the summer of 2013. The field methods and sampling design are described in the "Selenium Sampling Protocol." Briefly, we used a stratified GRTS (generalized random-tessellation stratified design; Stevens and Olsen 2004) unequal probability sample of points within an areal frame (the basin). The strata were delineated management units and within each strata water zones were sampled with unequal probabilities.

For both macroinvertebrate and water surveys, the sampling design was simplified compared to that for sediment. Points were selected for water and macroinvertebrate specimen collections within only the flooded zones in each management unit from the ordered list of points generated by the GRTS algorithm. This was a stratified GRTS sample across the basin for these trophic levels. The sediment GRTS point locations were selected using unequal probabilities within a management unit. The sampling design was a stratified GRTS design with unequal probability across zones within a stratum. The estimates for selenium concentration reflect these different sampling designs. The unequal probability sampling of selenium within a management unit was chosen for two main reasons: 1) to ensure water, macroinvertebrate, and sediment collections could be co-located at points with standing water; 2) based on the mechanisms of selenium contamination and accumulation in sediments, the areas with more persistent standing water should have, on average, higher concentrations of selenium. The first reason is to allow simpler statistical analyses to be used to investigate associations of selenium concentration among trophic levels, if of interest. From a design standpoint, the second reason suggests that specifying larger inclusion probabilities for those areas with more persistent water should result in smaller estimated variances (Thompson 2002).

1.2 Data Collection in 2014

Surveys were conducted in units 1, 3, 4A, and 4B during the summer of 2014. Similar to 2013, points were generated by the GRTS algorithm described. However, in 2014 only sediment samples were taken. As noted in 2013, there appears to be very different selenium concentrations found in sediment samples which contained a large proportion of root matter. This calls into question the use of data from dry root samples as representative of selenium concentrations within wet sediment. Based on initial findings in 2013, roots tend to have higher Se concentrations, resulting in a biased-high estimate of Se concentrations for sediment. In 2014, sediment samples were determined in the field to be either mostly root or sediment matter, in order to effectively split estimation into two indicators (sediment and root). In addition, for data analysis, the 2013 data has been modified to reflect this change in thinking.

1.3 Target Population Definitions

In 2013, the target population for macroinvertebrates and water was areas within a management unit that contained standing water. The target population for sediment was essentially the remaining areas defined within the management units. However, the definitions for the target populations changed before data collection started in 2014, due to the observed differences in root and sediment selenium concentrations. In 2014, the target population for macroinvertebrate and water remained the same. However, a refinement was made to the target population for sediment specimens to areas within a management unit that had an upper 2 cm of mineral-based matter and areas outside of monotypic stands of vegetation. After 2013 data collection, roots were identified as an additional specimen type of interest, correspondingly the target population was defined as areas within monotypic stands of vegetation. In the case of stands of vegetation, sediment was recorded as non-target and roots were recorded as target. Careful consideration was given to adjustments made to the 2013 data to reflect this as well.

2 Statistical Methods

2.1 Data Preparation For Analysis

The probabilistic sample of points allows us to make inferences to the entire basin as defined by the sampling frame (sampled population). Also, we are interested in estimating the mean selenium concentration in sediments, water, and macroinvertebrates within each management unit. We will use the spsurvey package in R to calculate the estimates of interest (Kincaid and Olsen 2013). The R-code is provided in Appendix B. This package is the same one used to select the points for potential surveying. However, one modification must happen prior to use of the package for analysis; the initial sampling weights must be adjusted to reflect how sampling *actually* occurred (realized sampling design).

Step 1: Adjusting initial weights for two GRTS draws in inter-unit canal, unit 4C, unit 5, and unit 6.

The issue that occurred with two draws for the same survey was that the siteID generated was no longer a unique identifier for a point location. Prior to analysis this needed to be fixed. Also, the weights generated by the `spsurvey` package do not reflect the whole design as it was done in two separate draws. Simply, the weights (column "wgt") in the csv files "`SampleDrawSe_2013May20.csv`" and "`SampleDrawSe_2013July1.csv`" should not be used in analysis.

The two GRTS draws require an adjustment to the original weights for Inter-unit canal and units 4C, 5, and 6 flooded zones. The adjusted initial weight is $w_{z,h} = \frac{A_{z,h}}{n_1 + n_2}$ where n_1 is the initial sample size specified for draw 1 within strata h and water zone z; n_2 is the second sample size specified in the second GRTS draw; and $A_{z,h}$ is the area of zone z in strata h. The adjusted weights are included in Table 1.

These adjusted weights assume that the original design specification in GRTS requested the total sample size all at once as opposed to in two batches as done in 2013. One consideration is that better spatial dispersion of points would have likely occurred if the sample draw were done in one step. Spatial statistics could be used to explore whether this is the case.

Step 2: Cleaning up evaluation status variables and codes.
There are four columns for evaluation status, consisting of EvalWater, EvalBugs, EvalSed, and EvalRoot to store the evaluation status for a point that varies based on criteria specific to water, macroinvertebrates, sediment, and roots. The column for root evaluation was added post-data collection. For each of the four evaluation statuses, one of six different evaluation statuses were

Table 1: Adjusted weights for initial sample draw prior to data collection and site visits in summer of 2013.

Stratum	Water Zone	Initial Weight	Area	n_1	n_2	Adjusted Weight
Inter-unit Canal	Flooded	1205	61476	51	50	609
Unit 1	Flooded	12253	1115066	91		12253
Unit 1	Intermittent	35097	421166	12		35097
Unit 1	Saturated	24643	542154	22		24643
Unit 2	Flooded	43853	1227897	28		43853
Unit 2	Intermittent	99560	199120	2		99560
Unit 2	Saturated	72432	869190	12		72433
Unit 3	Flooded	232785	2327850	10		232785
Unit 3	Intermittent	339542	1697710	5		339542
Unit 3	Saturated	331104	993312	3		331104
Unit 4A	Flooded	77273	695459	9		77273
Unit 4A	Intermittent	141494	565978	4		141495
Unit 4A	Saturated	87438	349753	4		87438
Unit 4B	Flooded	85987	1031848	12		85987
Unit 4B	Saturated	99143	396571	4		99143
Unit 4C	Flooded	106175	1061746	10	50	17696
Unit 4C	Intermittent	267941	1071764	4		267941
Unit 4C	Saturated	153166	2603827	17		153166
Unit 5	Flooded	223054	2230540	10	50	37176
Unit 5	Intermittent	198681	794724	4		198681
Unit 5	Saturated	254844	509687	2		254844
Unit 6	Flooded	120230	1202295	10	50	20038
Unit 6	Intermittent	255256	1531535	6		255256
Unit 6	Saturated	232824	931298	4		232825

recorded: TS, TN, NT, NN, NE, and NA. Originally, the data was structured differently; an explanation for the corrected data file, "`bnl.se.data.all.clean.csv`", is found in Appendix E. Hereafter, "evaluated" refers to whether a point was determined to have water, bugs, sediment, or roots, *not* whether measurements were actually taken. Evaluation makes a distinction between a target point and a non-target point, where target indicates that the desired specimen type (water, bugs, sediment, roots) was available for collection at the location, while not-target refers to points where the desired type was not available for collection. It is entirely possible for a point to be evaluated for presence/absence of a specimen type, but have no measurements taken. Separation of evaluated and measured is a necessary distinction in order to correctly quantify information in analysis. The first three (TS, TN, NT) evaluation statuses occur when the point is evaluated, and the second three (NN, NE, NA) occur when the point is not evaluated. A detailed explanation of these evaluation statuses can be found in Appendix E.

Particularly of interest are the evaluated points: TS (target and sampled; counts in Table 16 (Appendix A)), TN (target and not sampled; counts in Table 17), and NT (non-target and not sampled; counts in Table 19). In addition, counts for NN (not needed; counts in Table 20) and NE (not evaluated; counts in Table 18) are provided in Appendix A. There was only one true NA point, where it is likely that a data collection error occurred and the point was missed in field collection altogether.

Next we briefly describe several possible adjustments to the GRTS weights to account for the realized sample size and all the evaluation information. See Appendix C for more details.

2.1.1 Weighting based on realized sample size

The weights in Table 1 need to be adjusted to reflect the realized sample size within each zone and stratum. The simplest weight adjustment is to use $\frac{Area}{n*}$ where n^* is the number of sampled points (TS; Table 16) and Area is the original delineated area within a zone and stratum, as used in the original GRTS draw (Stevens and Olsen 2003). However, this does not account for over-coverage frame errors.

2.1.2 Weighting based on estimated area adjustment

Another option for weighting adjustment is to reduce the frame area originally delineated to an estimate of the area within the sampling frame that is target population. This can be done by calculating the weights as $\frac{\widehat{Area}}{n^*}$ where n^* is the number of sampled points (TS), and $\widehat{Area}$ is the adjusted frame area, e.g., the flooded zone area. To approximate the area that was target, the original area is multiplied by the proportion of points deemed target (TS or TN) out of the number of points accessible and evaluated, as $\widehat{Area} = Area * \left\{ {(n^{TS} + n^{TN})}/{(n^{TS} + n^{TN} + n^{NT})} \right\}$. For water and bug samples, there are no TN points, so the formula is simplified some in those cases.

2.1.3 Weighting based on area adjustment using GIS.

A final option for weighting adjustment is to reduce the frame area originally delineated by calculating the weights as ${\text{Area}^*}/{n*}$ where n^* is the number of sampled points (TS), and Area* is the specified area obtained from GIS mapping. This is likely the most accurate assessment of the area of interest. However, this Area* can only be used for sediment and root estimation, since it is simply the area with sediment (not roots) or the area with roots (from vegetation stands).

The criteria for a site to be inaccessible and not evaluated were the same across types (water, macroinvertebrate, sediment), so an estimate of inaccessible area could be calculated. If we assume inaccessible areas are similar in Se concentrations compared to those areas that could be accessed for field collection, drawing inferences to those areas is legitimate. Statistically, this is assuming the mechanism of "non-response" is completely at random and not related to the response variable (here, selenium concentration). In the reporting here, no adjustment for non-response or inaccessible areas was done.

2.2 Statistical methods for estimating mean Se concentrations within management units

For both macroinvertebrate and water surveys, the sampling design was simplified compared to that for sediment. Points were selected for water and macroinvertebrate collections only within the flooded zones, so technically a stratified GRTS sample across the basin for these trophic levels. The sediment point locations were selected using an unequal probability draw within a management unit, a stratified GRTS sample with unequal probability of selection across zones within a stratum. The sampling design was not constructed to survey roots; however, the flexibility of the GRTS design allowed for a reasonable assessment of Se concentration within roots when accessible. Using the ordered list allows for a subset of points that were sampled for roots to still be a probabilistic sample within the sampling frame (although not necessarily as spatially balanced as an a priori design for roots).

The final corrected file for analysis using spsurvey is "`SEDataForAnalysisSpSurveyFinal.csv`". This file must include the indicator values (Se concentration for water, bugs, sediment), the site weights for each indicator type, siteID, xcoord, ycoord, stratum, mdcaty, Areas.

The support or area of field collection was approximated by a 1 m diameter circle for water, 30 m diameter circle for macroinvertebrate, and 10.16 cm diameter circle for sediment. These need to be in the same units as the sampling frame (m^2). These values are used in the finite population correction factor (fpc) for the standard errors or in estimates of the population total. In this context, estimating selenium concentration within a management unit, neither of these are needed or of interest. The fpc values were all close to 1, which means they will not decrease the variance estimator. The population is essentially infinite in this case.

To estimate concentration for each stratum separately, subpopulations were specified in the `cont.analysis` function call within spsurvey. The following R-code was used to calculate 90% confidence intervals for Se concentration within sediments for each stratum using the local GRTS estimator (Stevens and Olsen 2003, 2004).

R-code snippet for local GRTS estimators: (`cont.analysis` function is within `spsurvey` package).

```
require(spsurvey)  # loads the package needed

# read the dataset and only use TS sediment values
FixedSeDataAnalysis <- read.csv("SEDataForAnalysisSpSurveyFinal.csv")
FixedSeDataAnalysis <- subset(FixedSeDataAnalysis, EvalSed == "TS")

# creates the site dataframe input
SedSites <- data.frame(siteID = FixedSeDataAnalysis$siteID, Active = rep(TRUE,
    length(FixedSeDataAnalysis$SeSed)))

# data values for indicator or response variable of interest
SedData <- with(FixedSeDataAnalysis, data.frame(siteID = siteID, SeSed))

# specifies the sub-populations for mean estimates desired, here
# management units
subpop <- FixedSeDataAnalysis[, c(2, 9)]
colnames(subpop)[2] <- "ManUnit"
sizes <- with(FixedSeDataAnalysis, tapply(init.area, stratum.x, mean))
SedPopSizes.Adj <- list(ManUnit = as.list(sizes))

# specified the design weights [here used estimated area adjusted
# weights]
AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
```

```
    wgt = adj.wgt.sed, xcoord = xcoord, ycoord = ycoord, support = 0.008107))

# function call conf=90 is the 90% confidence interval with
# vartype=local using local GRTS variance estimator
Sed.Est.Local.adj <- cont.analysis(sites = SedSites, subpop = subpop, design = AreaAdj.Design,
    data.cont = SedData, total = T, popsize = SedPopSizes.Adj, popcorrect = T,
    pcfsize = sum(sizes), vartype = "local", conf = 90, pctval = c(0))

# summarizing output to just the mean and confidence interval
subset(Sed.Est.Local.adj$Pct, Statistic == "Mean")[, c(2, 5:9)]
```

3 Results

As expected unit 1 has the highest concentration, on average, for all trophic levels for sediment and roots (sediment Figures 1 and 2; roots Figures 5 and 6). Note these figures are just raw data summaries and do not incorporate the unequal probability sampling that occurred. Unit 2 also has elevated selenium concentrations for sediment and roots, albeit much smaller than unit 1, as well as the highest concentrations, on average, for macroinvertebrates and water (macroinvertebrates Figure 3; water Figure 4).

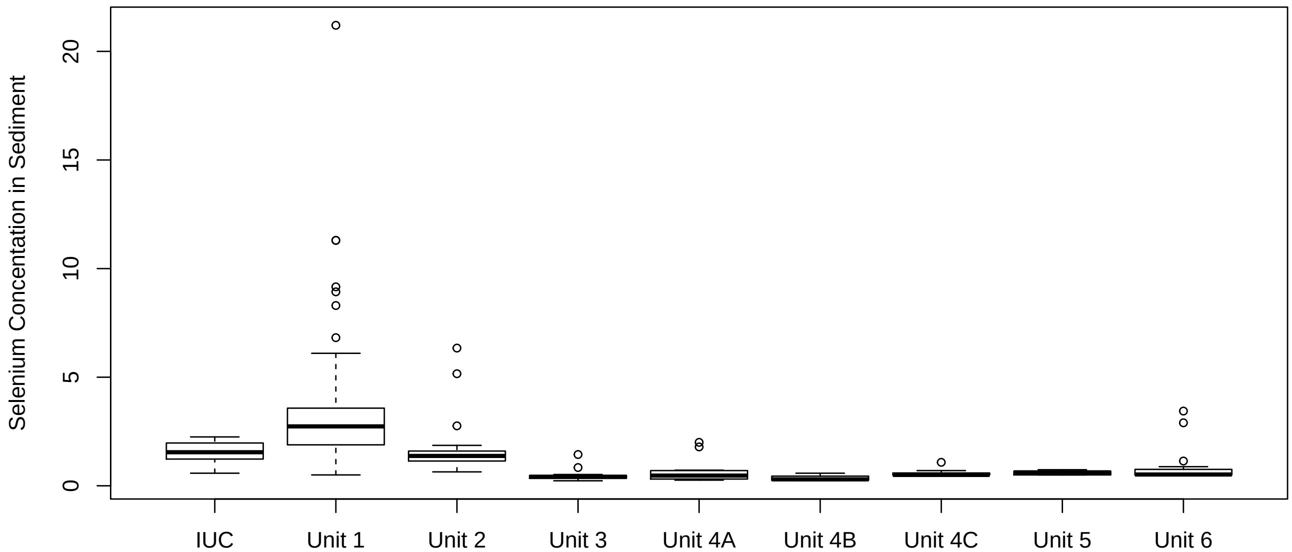


Figure 1: Boxplots for selenium concentration in sediment within each management unit.

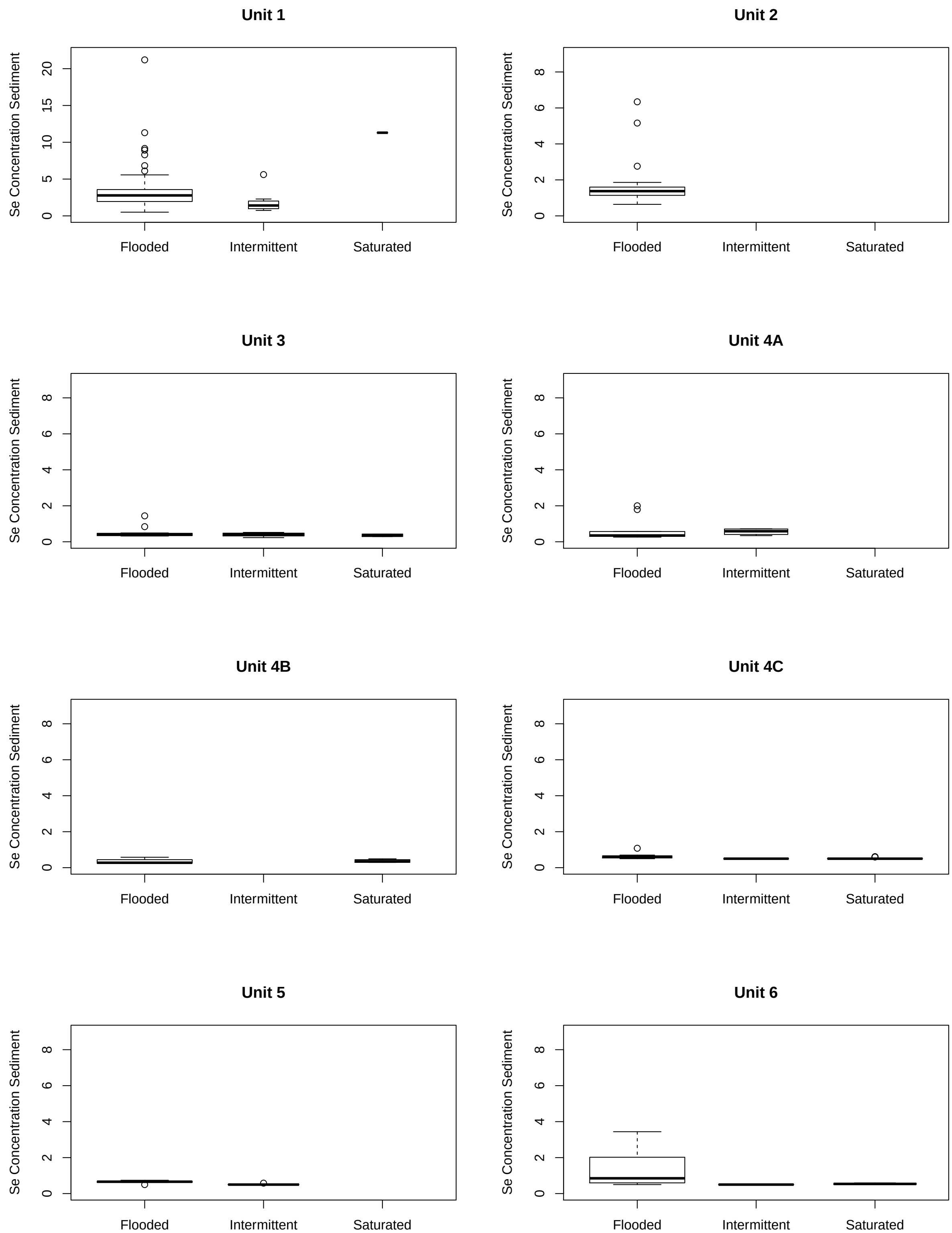


Figure 2: Boxplots of selenium concentration in sediments (non-root) by zone within each unit. Notice the difference in scale between unit 1 and every other unit.

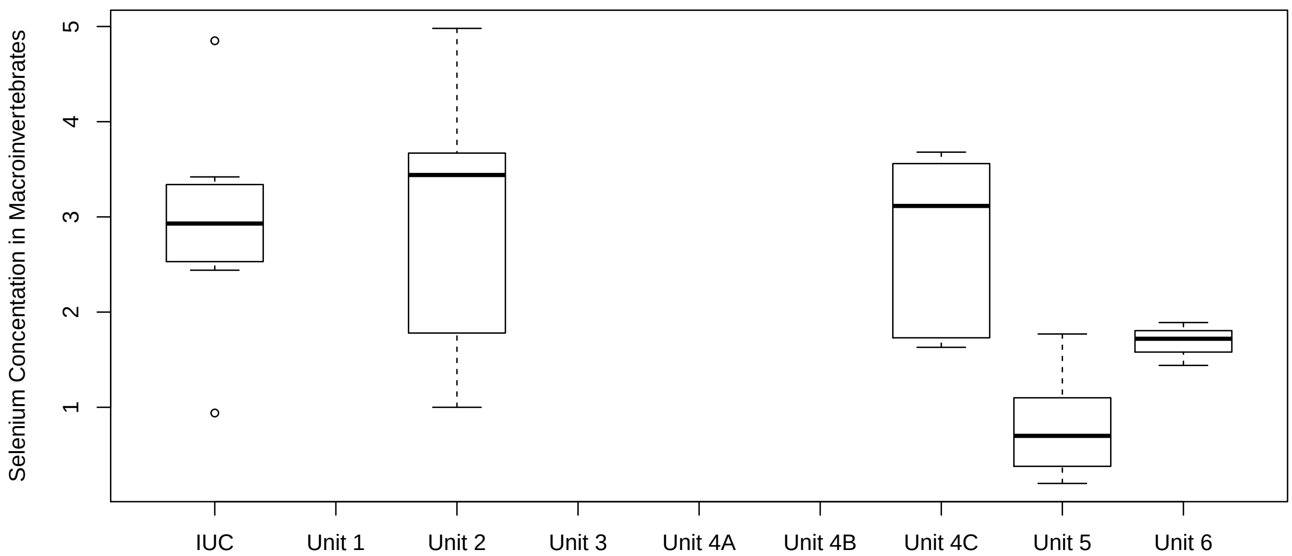


Figure 3: Boxplots for selenium concentration in macroinvertebrates within each management unit.

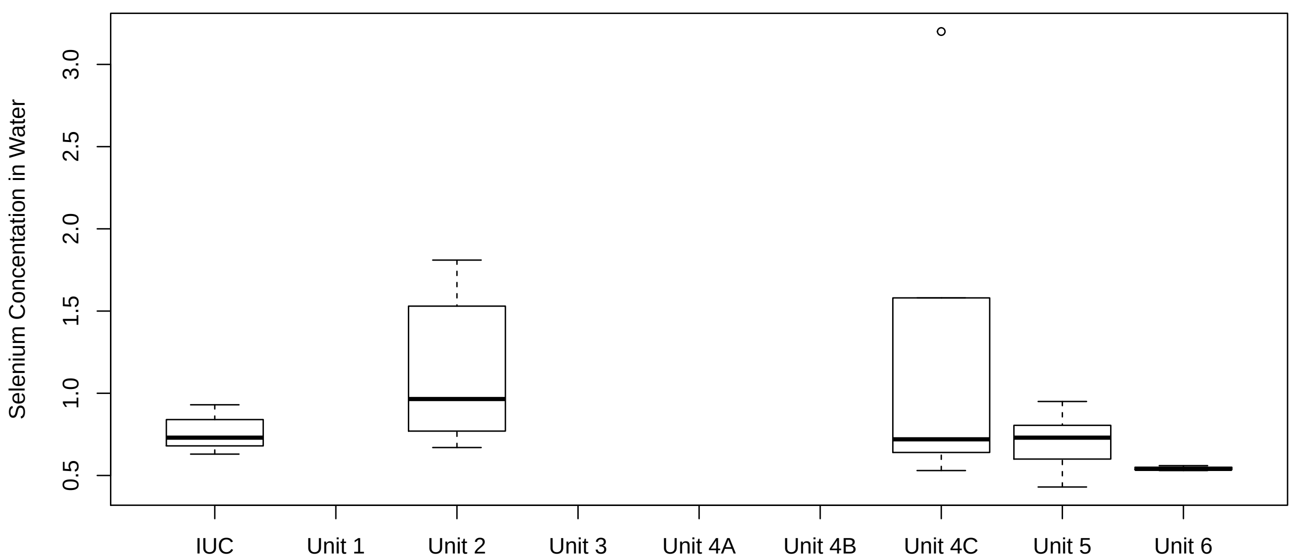


Figure 4: Boxplots for selenium concentration in water within each management unit.

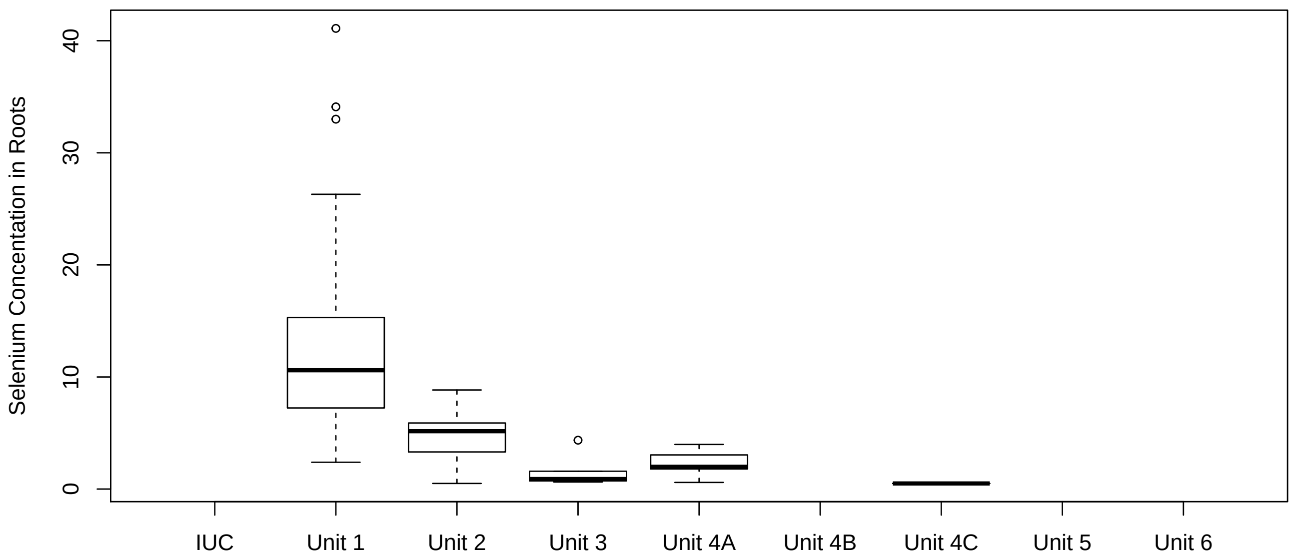


Figure 5: Boxplots for selenium concentration in roots within each management unit.

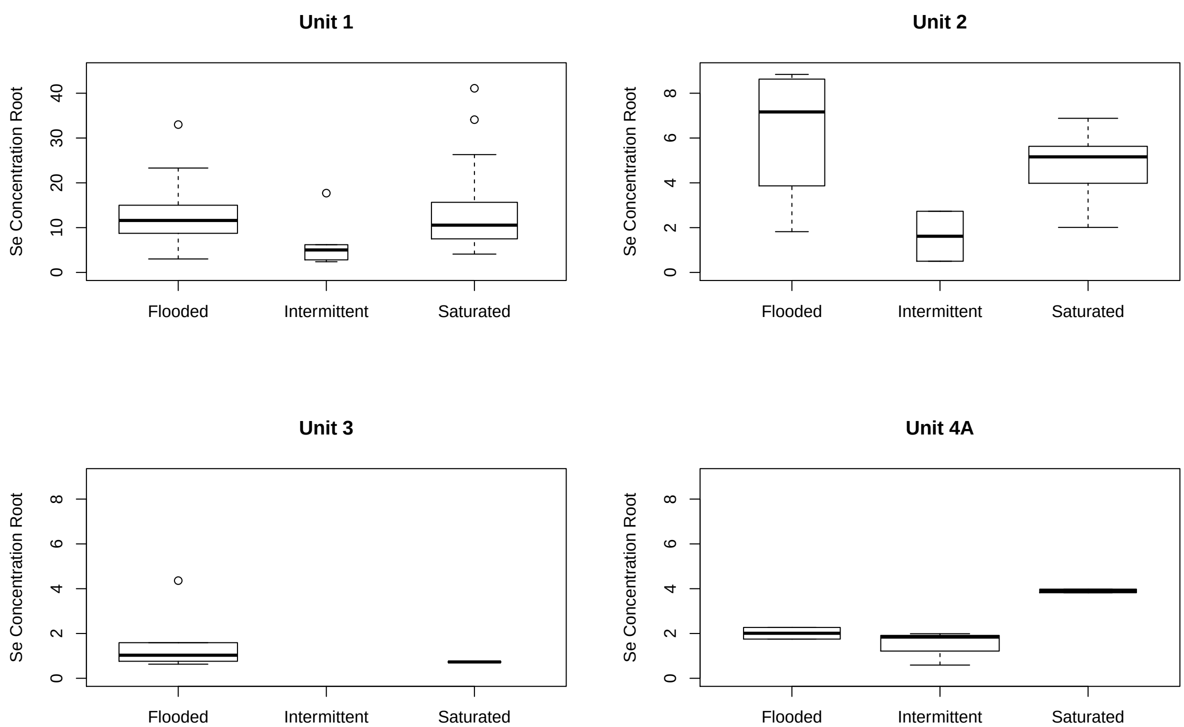


Figure 6: Boxplots of selenium concentration in roots by zone within each unit.

In terms of selenium concentration in macroinvertebrates, the variation was different among the units surveyed in 2013 (Figure 3). Unit 6 had the smallest variability and Unit 2 had the most. For selenium concentration in water, unit 2 and 4C has more variation compared to the inter-unit canal and unit 5; unit 6 had the smallest variation among measurements (Figure 4). For selenium concentration in roots, unit 1 had the most variability, while units 3, 4A, and 4C had very little.

Using the methods described in the previous sections, the mean concentration of selenium within the three trophic levels of sediment (Tables 2, 3, and 4), roots (Tables 5, 6, and 7), water (Table 8), macroinvertebrate (Table 9) were estimated. Multiple tables are provided for sediment and root selenium estimation to reflect the different ways to incorporate weighting adjustments described in Section 2 for these two specimen types. Water and macroinvertebrate selenium estimation did not change using different weight adjustments, so only one table is necessary. This is due to the lack of TS points for water and bugs in intermittent or saturated zones. Naturally, water and bug data can only be collected in flooded zones (areas with standing water) and therefore estimation is based on one zone. Thus for water and invertebrate samples, weights within a management unit are equal, simplifying estimation to a simple random sample estimator. The local GRTS and Horvitz-Thompson variance estimators were calculated for all specimen types, although there was no difference between them. Only the local GRTS variance estimators are displayed in the following tables. The lack of a difference in variance estimators suggests no spatial structure in the data. However, this could be related to how `spsurvey` defines the "local" neighborhood. A thought for future investigation is to explore modifying the local neighborhood using spatial statistics.

Table 2: Sediment mean selenium concentration estimate using weights based only on realized sample size and original area

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Inter-unit Canal	10	1.551	0.171	1.269	1.833	2.25
Unit 1	88	5.147	1.646	2.439	7.856	21.20
Unit 2	20	1.812	0.319	1.287	2.336	6.34
Unit 3	21	0.447	0.052	0.362	0.532	1.44
Unit 4A	13	0.642	0.131	0.427	0.857	2.00
Unit 4B	16	0.352	0.028	0.305	0.398	0.58
Unit 4C	30	0.546	0.018	0.516	0.575	1.08
Unit 5	14	0.616	0.023	0.578	0.653	0.74
Unit 6	19	0.792	0.157	0.533	1.051	3.44

Table 3: Sediment mean selenium concentration estimate using estimated area adjusted weights

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Inter-unit Canal	10	1.551	0.171	1.269	1.833	2.25
Unit 1	88	3.134	0.331	2.589	3.679	21.20
Unit 2	20	1.811	0.319	1.287	2.336	6.34
Unit 3	21	0.436	0.043	0.366	0.506	1.44
Unit 4A	13	0.639	0.127	0.429	0.848	2.00
Unit 4B	16	0.352	0.028	0.305	0.398	0.58
Unit 4C	30	0.547	0.019	0.516	0.578	1.08
Unit 5	14	0.616	0.023	0.578	0.653	0.74
Unit 6	19	0.792	0.157	0.533	1.051	3.44

Table 4: Sediment mean selenium concentration estimate using GIS area adjusted weights

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Inter-unit Canal	10	1.551	0.171	1.269	1.833	2.25
Unit 1	88	3.173	0.424	2.475	3.870	21.20
Unit 2	20	1.811	0.319	1.287	2.336	6.34
Unit 3	21	0.440	0.047	0.363	0.517	1.44
Unit 4A	13	0.631	0.119	0.436	0.826	2.00
Unit 4B	16	0.352	0.028	0.305	0.398	0.58
Unit 4C	30	0.535	0.013	0.513	0.557	1.08
Unit 5	14	0.614	0.023	0.576	0.652	0.74
Unit 6	19	0.775	0.149	0.529	1.021	3.44

Particularly in unit 1, the mean estimate of sediment selenium concentration varies greatly depending on which method is used to adjust the weights. This can be attributed to the presence of a large amount of cattail/garrison's foxtail stands in unit 1, which corresponds to a much smaller adjusted sampling frame area for sediment. Without this correction, the estimate is extremely high from a management perspective (over 5 ug/L). The correction using estimated frame area from the adjustments based off TS, TN, and NT sample points is similar to using the area based on GIS mapping. This similarity indicates that the frame adjustment based off field and office evaluation of GRTS points did a reasonable job of estimating the GIS mapped area that is not within monotypic stands of vegetation.

For the rest of the units, mean selenium concentration and variance estimation changes only marginally using any of the three weight adjustments. These units did not have nearly the amount of root matter from vegetation stands, and so the adjustment to frame area was small.

Table 5: Root mean selenium concentration estimate using weights based only on realized sample size and original area

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Unit 1	63	11.779	1.044	10.062	13.495	41.10
Unit 2	19	5.267	0.863	3.847	6.687	8.84
Unit 3	6	1.392	0.531	0.518	2.265	4.36
Unit 4A	7	2.232	0.369	1.625	2.838	3.98
Unit 4C	2	0.500				0.50

Table 6: Root mean selenium concentration estimate using estimated area adjusted weights

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Unit 1	63	12.405	1.225	10.390	14.420	41.10
Unit 2	19	4.833	0.558	3.915	5.751	8.84
Unit 3	6	1.382	0.527	0.515	2.248	4.36
Unit 4A	7	2.587	0.469	1.816	3.358	3.98
Unit 4C	2	0.500				0.50

Table 7: Root mean selenium concentration estimate using GIS adjusted area weights

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Unit 1	63	13.064	1.265	10.983	15.146	41.10
Unit 2	19	5.290	0.632	4.251	6.330	8.84
Unit 3	6	1.463	0.562	0.539	2.387	4.36
Unit 4A	7	2.756	0.476	1.974	3.539	3.98
Unit 4C	2	0.500				0.50

Roots were sampled in only units 1, 2, 3, 4A, and 4C. Estimates do change some depending on which method is used to calculate the weights. The estimate for unit 1 is extremely high, as expected from the earlier look at the data visually. Unit 4C had only two root selenium measurements; therefore, there mean is likely not representative and the variance estimate is not calculable.

Table 8: Water mean estimate using estimated area adjusted weights

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Inter-unit Canal	5	0.762	0.055	0.672	0.852	0.93
Unit 2	6	1.118	0.185	0.815	1.422	1.81
Unit 4C	6	1.232	0.423	0.536	1.927	3.20
Unit 5	7	0.703	0.068	0.592	0.814	0.95
Unit 6	3	0.543	0.009	0.529	0.558	0.56

Table 9: Invertebrate mean estimate using estimated area adjusted weights

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Inter-unit Canal	10	2.925	0.314	2.408	3.442	4.85
Unit 2	13	2.982	0.327	2.444	3.519	4.98
Unit 4C	6	2.805	0.369	2.197	3.413	3.68
Unit 5	9	0.781	0.171	0.500	1.062	1.77
Unit 6	3	1.683	0.131	1.468	1.899	1.89

3.1 Simple Random Sample Estimator

For demonstration, a simple SRS estimator was used for the sediment samples to explore whether the unequal probability sampling improved inferences (Table 10). This 'pretends' that the observations were selected using just GRTS without unequal probability across zones. For units 3, 4A, 4B, 4C, 5, and 6, the estimated standard errors were smaller using the unequal probability design with adjusted weights based on GIS area adjustment for sediment. This suggested the design provided greater efficiency (smaller variance) compared to just using a simple random sample (or GRTS since basically the local-estimator was the same) ignoring water zones.

Table 10: Sediment mean estimates using simple random sample design

Stratum	n	Mean Est	Std Error
Inter-unit Canal	10	1.551	0.171
Unit 1	88	3.339	0.300
Unit 2	20	1.812	0.319
Unit 3	21	0.459	0.056
Unit 4A	13	0.663	0.158
Unit 4B	16	0.351	0.029
Unit 4C	30	0.551	0.021
Unit 5	14	0.603	0.024
Unit 6	19	0.874	0.191

3.2 Vegetation Type Root Analysis

In addition to overall root Se concentration estimates, the estimation for different vegetation types may be of interest. Table 11 shows the number of each category of vegetation recorded and sampled in the field notes. When there is more than one species listed, both were present.

Table 11: Number of total occurrences of different vegetations

alkali, hardstem bulrush	1
cattail	27
cattail, sandspurry	1
foxtail barley	4
garrison	47
garrison, cattail	4
garrison, sandspurry	1
hardstem bulrush	3
lambsquarter	1
reed canary	1
reed canary grass roots	1
western wheatgrass, forb	1
western wheatgrass, kentucky bluegrass	1

Most of the vegetation types were not often seen, recording from 1 to 4 occurrences. Across many units, these sample sizes are too small to realistically obtain any reasonable estimate of a mean or variance. Instead, we will focus on the two species which were common in root samples: cattails and garrison's foxtail. There are 32 total occurrences of cattails, and 52 garrison's foxtails. Tables 12 and 13 provide a breakdown of cattails and garrison's foxtail across units and zones.

Table 12: Cattail occurrences across stratum and zones

	Flooded	Intermittent	Saturated
Unit 1	25	1	2
Unit 2	2	0	0
Unit 3	2	0	0

Table 13: Garrison's foxtail occurrences across stratum and zones

	Flooded	Intermittent	Saturated
Unit 1	7	4	17
Unit 2	1	1	12
Unit 3	3	0	1
Unit 4A	2	2	2

As an **exploratory analysis**, we will focus on estimation of root selenium concentration by species using only the number of target sampled points, with no area adjustment. An area adjustment based off field notes is not reasonable considering the data was not originally collected with vegetation species recorded as TS, TN, or NT. Areas from GIS mapping are available for cattail and garrison's foxtail; however, it may not be reasonable to utilize these values. For example, according to the GIS mapping, there were no cattail stands large enough to warrant an area estimate in the saturated zone of unit 1, but there were two cattail samples recorded there. Using the area based on GIS would ignore these samples entirely in the calculation of mean selenium concentration for unit 1. Tables 14 and 15 provide estimates of Se concentrations for cattails and garrison's foxtail using no area adjustment. Simply put, these estimates are meant to provide a reasonable exploration of potential mean Se concentration differences between species and units, not a concrete understanding of these differences. The files used for analysis of selenium concentration mean estimation in cattails and garrison's foxtail are "`SEDataForAnalysisSpSurveyCat.csv`" and "`SEDataForAnalysisSpSurveyGar.csv`", respectively.

Table 14: Cattail root mean estimates using weights based only on realized sample size and original area

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Unit 1	28	19.877	3.104	14.772	24.982	41.10
Unit 2	2	5.330				8.84
Unit 3	2	2.975				4.36

Table 15: Garrison's foxtail roots mean estimates using weights based only on real sample size and original area

Stratum	n	Mean Est	Std Error	LCB90Pct	UCB90Pct	Max
Unit 1	28	10.305	1.095	8.503	12.106	34.10
Unit 2	14	5.075	0.635	4.030	6.120	6.88
Unit 3	4	1.849	0.814	0.511	3.187	4.36
Unit 4A	6	2.167	0.434	1.453	2.882	3.98

4 Discussion

The survey conducted during the summers of 2013 and 2014 in Benton Lakes NWR provides a much needed baseline assessment of selenium contamination for inter-unit canal and units 1, 2, 3, 4, 4A, 4B, 4C, 5, and 6 within the basin. Based on these data, there was no benefit to using the local-GRTS variance estimator for any of the trophic levels as they were the same as using a Horvitz-Thompson estimator. With a higher sampling intensity, which would result in points closer together in space, the GRTS local estimator may have smaller variance compared to the H-T estimator.

There appeared to be a minor reduction in the standard errors for sediment selenium concentration in units 3, 4A, 4B, 4C, 5, and 6 with the unequal probability selection across zones. For long-term monitoring, the trade-off between increased sampling design complexity and minimal reduction in standard errors should be considered. Any statistical model used to analyze data under an unequal probability design should investigate and likely include the design weights; this could be an unnecessary burden analytically if the standard errors for status estimation are not dramatically improved. A more thorough investigation into whether concentrations are positively correlated with inclusion probabilities should be done in 2015 to explore sampling design options for long-term monitoring.

The sample size calculations presented in the sampling protocol should be revisited once the entire basin has been surveyed (field collection in 2014) in 2015. A comparison of standard deviations calculated from the baseline data gathered in 2013-2014 to those estimated (based on historic datasets) and used to determine needed sampling effort in 2013-2014 should be done. Also, an investigation into whether the sampling effort is sufficient to detect changes in selenium concentration as needed to guide management can be explored. An important and informative finding was that the type of sediment specimen matters when measuring selenium concentration. The roots had, on average, higher Se concentrations and potentially biased the estimates if assumed to be sediment only concentrations.

Another consideration is the data management required to generate appropriate design-based estimators of selenium concentration. The site IDs with multiple draws needed to be corrected as they were no longer unique. Also, the initial design weights needed to be adjusted for the two GRTS draws. Importantly, the initial weights need to be adjusted to reflect the realized sample size and design. The weighting adjustments incorporate the information from points that were evaluated regarding target population status (e.g., TN,TS,NT). These adjustments are needed to reflect and correct for the imperfect sampling frame that is available and used for the sampling design. There was no adjustment made for inaccessible areas.

Moving forward, if an estimate for the selenium concentration within roots of different vegetation types (species) is of interest, then the sampling design should be adjusted to better represent this specimen type. The original sampling design was assuming only water, macroinvertebrates, and sediment selenium concentrations were of interest. The finding that selenium concentrations based on collecting mainly roots was substantially different than collecting mainly sediment should be considered. The sampling frame and design should be re-evaluated given the wealth of information provided by the baseline assessment conducted in 2013-3014.

5 References

1. "Aquatic Resource Monitoring." EPA. Environmental Protection Agency, 6 Apr. 2011. Web.

2. Kincaid, T. M. and Olsen, A. R. 2013. spsurvey: Spatial Survey Design and Analysis. R package version 2.6. URL: http://www.epa.gov/nheerl/arm/.

3. Olsen, Anthony R., Blaine D. Snyder, Leanne L. Stahl, and Jennifer L. Pitt. "Survey Design for Lakes and Reservoirs in the United States to Assess Contaminants in Fish Tissue." Environmental Monitoring and Assessment 150.1-4 (2009): 91-100.

4. Rodhouse, Thomas J., Kerri T. Vierling, and Kathryn M. Irvine. "A Practical Sampling Design For Acoustic Surveys of Bats." The Journal of Wildlife Management 75.5 (2011): 1094-1102.

5. Stevens, Jr D.L and Olsen, A.R. 2004. Spatially balanced sampling of natural resources. JASA 99(465): 262-278.

6. Stevens, Jr D.L and Olsen, A.R. 2003. Variance estimation for spatially balanced samples of environmental resources. Environmetrics 14:593-610

7. Thompson, S.K. 2002. Sampling, Second Edition. Wiley Series in Probability and Statistics

6 Appendix A: Summary of Realized Sampling Design

Table 16: TS - realized sample size for sediment, water, macroinvertebrates, and roots by stratum and zone.

SEDIMENT	Flooded	Intermittent	Saturated
Inter-unit Canal	10	0	0
Unit 1	79	8	1
Unit 2	20	0	0
Unit 3	11	8	2
Unit 4A	9	4	0
Unit 4B	12	0	4
Unit 4C	8	7	15
Unit 5	9	5	0
Unit 6	8	5	6

WATER	Flooded	Intermittent	Saturated
Inter-unit Canal	5	0	0
Unit 2	6	0	0
Unit 4C	6	0	0
Unit 5	7	0	0
Unit 6	3	0	0

MACROINVERTEBRATES	Flooded	Intermittent	Saturated
Inter-unit Canal	10	0	0
Unit 2	13	0	0
Unit 4C	6	0	0
Unit 5	9	0	0
Unit 6	3	0	0

ROOTS	Flooded	Intermittent	Saturated
Unit 1	38	5	20
Unit 2	4	2	13
Unit 3	5	0	1
Unit 4A	2	3	2
Unit 4C	0	0	2

Table 17: TN - number of target points which were not sampled for sediment and roots by stratum and zone.

SEDIMENT	Flooded	Intermittent	Saturated
Unit 2	3	0	0
Unit 4C	56	0	0
Unit 5	63	0	0
Unit 6	58	0	0

ROOTS	Flooded	Intermittent	Saturated
Unit 1	14	0	8
Unit 2	6	0	5
Unit 3	5	0	1
Unit 4A	6	0	3

Table 18: NE - number of points not evaluated for water and macroinvertebrates by stratum and zone.

WATER	Flooded	Intermittent	Saturated
Unit 2	1	0	0

MACROINVERTEBRATES	Flooded	Intermittent	Saturated
Unit 2	2	0	0

Table 19: NT - number of points deemed non-target for sediment, water, macroinvertebrates, and roots by stratum and zone.

SEDIMENT	Flooded	Intermittent	Saturated
Unit 1	52	5	28
Unit 2	10	3	18
Unit 3	10	0	2
Unit 4A	8	3	5
Unit 4C	0	0	2

WATER	Flooded	Intermittent	Saturated
Unit 2	3	1	0
Unit 4C	58	0	0
Unit 5	63	0	0
Unit 6	63	0	0

MACROINVERTEBRATES	Flooded	Intermittent	Saturated
Inter-unit Canal	10	0	0
Unit 2	13	0	0
Unit 4C	6	0	0
Unit 5	9	0	0
Unit 6	3	0	0

ROOTS	Flooded	Intermittent	Saturated
Inter-unit Canal	10	0	0
Unit 1	79	8	1
Unit 2	23	1	0
Unit 3	11	8	2
Unit 4A	9	4	0
Unit 4B	12	0	4
Unit 4C	64	7	15
Unit 5	72	5	0
Unit 6	66	5	6

Table 20: NN - number of points not-needed for sediment, water, macroinvertebrates, and roots by stratum and zone.

SEDIMENT	Flooded	Intermittent	Saturated
Inter-unit Canal	101	0	0
Unit 1	0	1	1
Unit 2	6	2	0
Unit 3	0	2	0
Unit 4B	8	0	3
Unit 4C	0	3	11
Unit 5	0	3	2
Unit 6	0	6	2

WATER	Flooded	Intermittent	Saturated
Inter-unit Canal	106	0	0
Unit 1	131	14	30
Unit 2	29	4	18
Unit 3	21	10	4
Unit 4A	17	7	5
Unit 4B	20	0	7
Unit 4C	0	10	28
Unit 5	2	8	2
Unit 6	0	11	8

MACROINVERTEBRATES	Flooded	Intermittent	Saturated
Inter-unit Canal	101	0	0
Unit 1	131	14	30
Unit 2	21	4	18
Unit 3	21	10	4
Unit 4A	17	7	5
Unit 4B	20	0	7
Unit 4C	0	10	28
Unit 5	0	8	2
Unit 6	0	11	8

ROOTS	Flooded	Intermittent	Saturated
Inter-unit Canal	101	0	0
Unit 1	0	1	1
Unit 2	6	2	0
Unit 3	0	2	0
Unit 4B	8	0	3
Unit 4C	0	3	11
Unit 5	0	3	2
Unit 6	0	6	2

7 Appendix B: R-code used for data preparation and analysis

```
SE.Data <- read.csv("bnl.se.data.all.clean.csv", header = T)
SE.Data <- subset(SE.Data, stratum != "Unit 4A-T")
SE.Data <- SE.Data[-((nrow(SE.Data) - 3):nrow(SE.Data)), ]
SE.Data$SeSed <- as.numeric(ifelse(SE.Data$SeSed == "noData", NA, as.character(SE.Data$SeSed)))
SE.Data$SeWater <- as.numeric(ifelse(SE.Data$SeWater == "noData", NA, as.character(SE.Data$SeWater)))
SE.Data$SeBugs <- as.numeric(ifelse(SE.Data$SeBugs == "noData", NA, as.character(SE.Data$SeBugs)))
SE.Data$SeRoots <- as.numeric(ifelse(SE.Data$SeRoots == "noData", NA, as.character(SE.Data$SeRoots)))
SE.Data <- subset(SE.Data, EvalSed != "NA")

####################################### SEDIMENT ESTIMATES ############
FixedSeDataAnalysis <- read.csv("SEDataForAnalysisSpSurveyFinal.csv")
FixedSeDataAnalysis <- subset(FixedSeDataAnalysis, EvalSed == "TS")
SedSites <- data.frame(siteID = FixedSeDataAnalysis$siteID, Active = rep(TRUE,
    length(FixedSeDataAnalysis$SeSed)))

AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = non.area.adj.wgt.sed, xcoord = xcoord, ycoord = ycoord, support = 0.008107))

SedData <- with(FixedSeDataAnalysis, data.frame(siteID = siteID, SeSed))
subpop <- FixedSeDataAnalysis[, c(2, 9)]
colnames(subpop)[2] <- "ManUnit"

sizes <- with(FixedSeDataAnalysis, tapply(init.area, stratum.x, mean))

SedPopSizes.Adj <- list(ManUnit = as.list(sizes))

# using non area adjusted weights
Sed.Est.Local.init <- cont.analysis(sites = SedSites, subpop = subpop,
    design = AreaAdj.Design, data.cont = SedData, total = T, popsize = SedPopSizes.Adj,
    popcorrect = T, pcfsize = sum(sizes), vartype = "local", conf = 90,
    pctval = c(0))


# adjusting using estimated area
AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = adj.wgt.sed, xcoord = xcoord, ycoord = ycoord, support = 0.008107))

Sed.Est.Local.adj <- cont.analysis(sites = SedSites, subpop = subpop, design = AreaAdj.Design,
    data.cont = SedData, total = T, popsize = SedPopSizes.Adj, popcorrect = T,
    pcfsize = sum(sizes), vartype = "local", conf = 90, pctval = c(0))

# adjusting using real area
AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = real.wgt.sed, xcoord = xcoord, ycoord = ycoord, support = 0.008107))

Sed.Est.Local.real <- cont.analysis(sites = SedSites, subpop = subpop,
    design = AreaAdj.Design, data.cont = SedData, total = T, popsize = SedPopSizes.Adj,
    popcorrect = T, pcfsize = sum(sizes), vartype = "local", conf = 90,
    pctval = c(0))


# ignoring unequal wts and assuming JUST A SRS within each stratum
```

```r
# [WRONG!!]

Simple.SRS.sed <- ddply(FixedSeDataAnalysis, ~stratum.x, summarize, SampleSize = length(SeSed),
    Mean = mean(SeSed), SE = sd(SeSed)/sqrt(length(SeSed)))
maxes.sed <- as.matrix(with(FixedSeDataAnalysis, tapply(SeSed, stratum.x,
    max)))

####################################### ROOT ESTIMATES ############
FixedSeDataAnalysis <- read.csv("SEDataForAnalysisSpSurveyFinal.csv")
FixedSeDataAnalysis <- subset(FixedSeDataAnalysis, EvalRoot == "TS")
RootSites <- data.frame(siteID = FixedSeDataAnalysis$siteID, Active = rep(TRUE,
    length(FixedSeDataAnalysis$SeRoots)))

AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = non.area.adj.wgt.root, xcoord = xcoord, ycoord = ycoord, support = 0.008107))

RootData <- with(FixedSeDataAnalysis, data.frame(siteID = siteID, SeRoots))
subpop <- FixedSeDataAnalysis[, c(2, 9)]
colnames(subpop)[2] <- "ManUnit"

sizes <- with(FixedSeDataAnalysis, tapply(init.area, stratum.x, mean))
sizes <- sizes[c(2, 3, 4, 5, 7)]

RootPopSizes.Adj <- list(ManUnit = as.list(sizes))

# using non area adjusted weights
Root.Est.Local.init <- cont.analysis(sites = RootSites, subpop = subpop,
    design = AreaAdj.Design, data.cont = RootData, total = T, popsize = RootPopSizes.Adj,
    popcorrect = T, pcfsize = sum(sizes), vartype = "local", conf = 90,
    pctval = c(0))


# adjusting using estimated area
AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = adj.wgt.root, xcoord = xcoord, ycoord = ycoord, support = 0.008107))
Root.Est.Local.adj <- cont.analysis(sites = RootSites, subpop = subpop,
    design = AreaAdj.Design, data.cont = RootData, total = T, popsize = RootPopSizes.Adj,
    popcorrect = T, pcfsize = sum(sizes), vartype = "local", conf = 90,
    pctval = c(0))

# adjusting using real area
AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = real.wgt.root, xcoord = xcoord, ycoord = ycoord, support = 0.008107))
Root.Est.Local.real <- cont.analysis(sites = RootSites, subpop = subpop,
    design = AreaAdj.Design, data.cont = RootData, total = T, popsize = RootPopSizes.Adj,
    popcorrect = T, pcfsize = sum(sizes), vartype = "local", conf = 90,
    pctval = c(0))


# ignoring unequal wts and assuming JUST A SRS within each stratum
# [WRONG!!]

Simple.SRS.root <- ddply(FixedSeDataAnalysis, ~stratum.x, summarize, SampleSize = length(SeRoots),
    Mean = mean(SeRoots), SE = sd(SeRoots)/sqrt(length(SeRoots)))
maxes.root <- as.matrix(with(FixedSeDataAnalysis, tapply(SeRoots, stratum.x,
    max)))
```

```r
######################################## WATER ESTIMATES ############
FixedSeDataAnalysis <- read.csv("SEDataForAnalysisSpSurveyFinal.csv")
FixedSeDataAnalysis <- subset(FixedSeDataAnalysis, EvalWater == "TS")
WaterSites <- data.frame(siteID = FixedSeDataAnalysis$siteID, Active = rep(TRUE,
    length(FixedSeDataAnalysis$SeWater)))

AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = non.area.adj.wgt.water, xcoord = xcoord, ycoord = ycoord, support = 0.008107))

WaterData <- with(FixedSeDataAnalysis, data.frame(siteID = siteID, SeWater))
subpop <- FixedSeDataAnalysis[, c(2, 9)]
colnames(subpop)[2] <- "ManUnit"

sizes <- with(FixedSeDataAnalysis, tapply(init.area, stratum.x, mean))
sizes <- sizes[c(1, 3, 7, 8, 9)]

WaterPopSizes.Adj <- list(ManUnit = as.list(sizes))

# using non area adjusted weights
Water.Est.Local.init <- cont.analysis(sites = WaterSites, subpop = subpop,
    design = AreaAdj.Design, data.cont = WaterData, total = T, popsize = WaterPopSizes.Adj,
    popcorrect = T, pcfsize = sum(sizes), vartype = "local", conf = 90,
    pctval = c(0))


# adjusting using estimated area
AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = adj.wgt.water, xcoord = xcoord, ycoord = ycoord, support = 0.008107))
Water.Est.Local.adj <- cont.analysis(sites = WaterSites, subpop = subpop,
    design = AreaAdj.Design, data.cont = WaterData, total = T, popsize = WaterPopSizes.Adj,
    popcorrect = T, pcfsize = sum(sizes), vartype = "local", conf = 90,
    pctval = c(0))


# ignoring unequal wts and assuming JUST A SRS within each stratum
# [WRONG!!]

Simple.SRS.water <- ddply(FixedSeDataAnalysis, ~stratum.x, summarize, SampleSize = length(SeWater),
    Mean = mean(SeWater), SE = sd(SeWater)/sqrt(length(SeWater)))
maxes.water <- as.matrix(with(FixedSeDataAnalysis, tapply(SeWater, stratum.x,
    max)))


######################################## BUGS ESTIMATES #############
FixedSeDataAnalysis <- read.csv("SEDataForAnalysisSpSurveyFinal.csv")
FixedSeDataAnalysis <- subset(FixedSeDataAnalysis, EvalBugs == "TS")
BugSites <- data.frame(siteID = FixedSeDataAnalysis$siteID, Active = rep(TRUE,
    length(FixedSeDataAnalysis$SeBugs)))

AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = non.area.adj.wgt.bugs, xcoord = xcoord, ycoord = ycoord, support = 0.008107))

BugData <- with(FixedSeDataAnalysis, data.frame(siteID = siteID, SeBugs))
subpop <- FixedSeDataAnalysis[, c(2, 9)]
colnames(subpop)[2] <- "ManUnit"
```

```r
sizes <- with(FixedSeDataAnalysis, tapply(init.area, stratum.x, mean))
sizes <- sizes[c(1, 3, 7, 8, 9)]

BugPopSizes.Adj <- list(ManUnit = as.list(sizes))

# using non area adjusted weights
Bug.Est.Local.init <- cont.analysis(sites = BugSites, subpop = subpop,
    design = AreaAdj.Design, data.cont = BugData, total = T, popsize = BugPopSizes.Adj,
    popcorrect = T, pcfsize = sum(sizes), vartype = "local", conf = 90,
    pctval = c(0))


# adjusting using estimated area
AreaAdj.Design <- with(FixedSeDataAnalysis, data.frame(siteID = siteID,
    wgt = adj.wgt.bugs, xcoord = xcoord, ycoord = ycoord, support = 0.008107))
Bug.Est.Local.adj <- cont.analysis(sites = BugSites, subpop = subpop, design = AreaAdj.Design,
    data.cont = BugData, total = T, popsize = BugPopSizes.Adj, popcorrect = T,
    pcfsize = sum(sizes), vartype = "local", conf = 90, pctval = c(0))


# ignoring unequal wts and assuming JUST A SRS within each stratum
# [WRONG!!]

Simple.SRS.bug <- ddply(FixedSeDataAnalysis, ~stratum.x, summarize, SampleSize = length(SeBugs),
    Mean = mean(SeBugs), SE = sd(SeBugs)/sqrt(length(SeBugs)))
maxes.bug <- as.matrix(with(FixedSeDataAnalysis, tapply(SeBugs, stratum.x,
    max)))
```

Prepared by Kathryn M. Irvine, USGS and Jay R.

8 Appendix C: Adjusting Design Weights for Over-coverage Frame Errors

A frame error happens commonly in environmental and ecological monitoring. Likely because of the strong reliance on GIS spatial information for defining an areal and/or list frame. A common frame error is over-coverage. This happens when a G.I.S. layer may contain areas that are not aligned with the target population of interest. For example, wetlands have dynamic hydrologic cycles in that basins or units fluctuate between wet and dry annually and seasonally. Alternatively, with water management a unit may be flooded one year, *wet*, and then dry other years due to pumping of water into the unit. However, the only known spatial information is the delineated wetland area at full pool or based on imagery from a different year. If surveying of many trophic levels is of interest (i.e., multiple specimen types) sediments, water, and invertebrates within a wetland, this results in different defined sub-populations corresponding to each trophic level. These different ecological indicators may have over-lapping or disjoint areas of interest for inference that may change over time.

In the case of over-coverage only frame errors, the recommended approach for unbiased estimation is domain estimation (Lessler and Kalsbeek 1992). In this case, the domain is the target population contained within the sampled population (or sampling frame). A probability weighted estimator of the domain mean is

$$\hat{\bar{y}}_{s_d} = \frac{1}{\hat{N}_d} \sum_{s_d} w_i y_i \tag{1}$$

with s_d denoting the sampled sites within the target population, w_i the survey weight for a given observation y_i, and $\hat{N}_d = \sum_{s_d} w_i$ (Sarndal et al. 1992). The survey weight is the inverse of the inclusion probability. The inclusion probability is the probability of a given sample unit being selected into the sample. A survey weight can be thought of as the number of units a sampled unit represents in the target population. For example, a data value with survey weight of 10 means that data value represents 10 other sample units in the target population.

A flexible sampling design that accommodates co-location of ecological indicators and allows for dynamically adding or deleting of sample points is GRTS (Stevens and Olsen 2004). GRTS is a spatially balanced sampling design. For a readable description of this type of sampling design, see Rodhouse, Vierling, Irvine (2011). The utility of the GRTS design is to anticipate the need for over-sample sites due to over-coverage. There is little loss of efficiency if the uncontrolled domain sample size is the same as the controlled domain sample size in the case of no frame errors (p. 393 Sarndal et al. 1992). However, using a dynamic sampling design, such as GRTS, requires adjusting the original survey weight after the design has been implemented. Currently, guidance regarding adjusting the GRTS survey weights is restricted to the case of having sites (points) evaluated and deemed target and data is collected versus deemed not target and no data are collected.

The situation at Benton Lakes NWR is unique, although likely not uncommon. Not all points evaluated and deemed target had data collected, depending on the point's order within the GRTS ordered list. Therefore, the need to create another evaluation code, TN- target and not sampled. This code was only used for sediment and root evaluation criteria. Also, a vegetation map within a GIS was available to determine sampling frame area within monotypic stands of cattails and garrison's which were deemed non-target for sediment collection (a posteriori in 2014). The original sampling design was only optimized to survey sediment, water, and macroinvertebrates. However, we found based on field surveys in 2013 that selenium concentration measurements from mainly roots were much higher compared to mainly sediment materials. Therefore, adjustments were made post-hoc to discriminate mainly root material from mainly sediment collections. This raised challenges regarding redefining the target populations and, consequently, estimating the sampling frame area that was within the target population for roots versus target population for sediment within a unit.

8.1 Adjustment Procedures for Survey Weights

For simplicity of notation, the following ignores stratification and only considers an unequal probability sample within one management unit.

Let $n^{'}$ = the total sample size desired, $n^{'} = \sum_k n_k$ where n_k is the sample size specified in category k (multi-density category), assume the sampling frame is delineated into k categories or areas with total frame area ($A = \sum_k A_k$), the original design weight for location i in category k (denoted as w_{ik}) is the inverse of the inclusion probability, i.e., $w_{ik} = \pi_{ik}^{-1}$ with $\pi_{ik} = \frac{n_k}{A_k}$. After field surveys and desktop evaluation occurred, site (point) locations were recorded as TS, TN, or NT, let $n_{TS,k}$ denote the number of sites target and sampled within category k, $n_{TN,k}$ denote the number of sites target and not sampled within category k, and $n_{NT,k}$ denote the number of sites deemed non-target within category k.

The sampling frame may contain over-coverage errors which means that a fraction of the GRTS points may occur within NT (non-target) areas. Denote the target population area as A_k^* within category k, then $A_k^* \subset A_k$.

For all adjustment procedures, ideally, the sum of the adjusted weights over the number of points deemed TS (target and data collected) should equal the adjusted target population size or area,

$$\sum_{n_{TS,k}} w_{ik}^{adj} = \hat{A}_k. \tag{2}$$

In the setting of Benton Lakes NWR, $\hat{A}_k$ is an estimated area of target population within category k (water zone within a management unit). For example, for selenium concentration in water, the target population or area for inference is the area of standing water in a given year, index period (within a year the time period for data collection), and management unit.

8.1.1 EPA procedure

For simplicity, we drop the sub-script k. The adjustment is made within each k dependent on the number of GRTS points evaluated.

The following is translated from the website
http://www.epa.gov/nheerl/arm/analysispages/analysisadjwts.htm.

It appears that the situation considered on the EPA website is when the sites evaluated fall into only 3 types: TS - target and sampled site, NT - site was non-target (dry, too shallow,...), NN - site not needed (not evaluated). Also, non-response such as inaccessible or access denied sites are assumed to occur at random and independent of site characteristics.

To account for the use of "inverse sampling" or dynamically adding points from the over-sample list (a hallmark of the GRTS design) we calculate adjusted weights, w_i^{adj} for site i, as follows:

$$w_i^{adj} = w_i \frac{A_k}{\sum_{i \in S_k} w_i}$$

where the sum is over the sites evaluated (TS+NT) within category k. So if a site was TS or NT for category k then it is in the set S_k where TS - target and sampled site, NT - site was non-target (dry, too shallow,...), NN - site not needed (not evaluated).

Notice for equal weights (i.e. $w_i = w \; \forall i \in k$) and the number of sites in S_k as $n_{TS} + n_{NT}$, then we have that

$$w_i^{adj} = \left(\frac{A_k}{n_k}\right) \frac{A_k}{(n_{TS} + n_{NT})\frac{A_k}{n_k}} = \frac{A_k}{n_{TS} + n_{NT}}.$$

So the inclusion probability is just $(n_{TS} + n_{NT})/A_k$; adjusted for how many sites are actually evaluated (TS and NT sites) within each category.

This results in an adjusted area that accounts for over-coverage or having a GIS areal frame that contains non-target areas. The sum of the adjusted weights for the TS sites is $= \hat{A}'_k$, the estimated frame area. The estimated frame area is simply

$$\hat{A}'_k = A_k \frac{n_{TS}}{n_{TS} + n_{NT}}. \tag{3}$$

An alternative derivation to the adjusted weights is to calculate the estimated frame area $\hat{A}'_k$ and then for the weights divide by the number of target and sampled sites (realized sample size) as follows: $w_i^{adj} = \frac{\hat{A}'_k}{n_{TS}}$. These turn out to be equivalent to that described on the EPA website when sites are considered TS or NT. However, these are not equivalent when we consider the situation at Benton Lakes NWR, less points had data collected than were available for data collection (the new category TN- target and not sampled).

Copied from the website

Prior to analyzing GRTS Design probability data, it is necessary to ensure that each sampled site is assigned the correct weight. Commonly, weights are adjusted following the completion of the Design Status File to reflect the implementation of the design. Most often the design is not implemented exactly as initially specified, commonly fewer or more sites, thus requiring the weight adjustment.

The initial design weights are assigned during the GRTS design process and individual site weights reflect the amount (stream length, estuarine area, etc.) of the target population represented by the site. For example, if the frame contains 1000 km of stream length and the design sample size is 100, then each site has a design weight of 10. If during the evaluation and sampling phase only the first 90 sites are used, then the initial design weights are adjusted:

adjusted weight(i) = sum(design wts)/sum(used wts)x design weight(i)

These weights are then used when calculating estimates for the Target Population that is included in the Sample Frame.

Estimates based on Target Sampled sites apply to the Sampled Population with no additional assumptions.
Estimates based on Target Sampled sites can apply to the portion of Target Population within the Sample Frame ONLY IF assume that the Access Denied, Target Not-Sampled, etc., sites occurred randomly and independently of site characteristics
Estimates for Target Population NOT ONLY require assumptions above BUT ALSO that portions of Target Population that are not included in the Sample Frame have same characteristics as the Sampled Population

8.1.2 Incorporating Ancillary Information

At Benton Lakes NWR, we had ancillary information from GIS vegetation mapping and additional points that were evaluated and not sampled. Therefore, we need to deviate from the typical GRTS weight adjustment implementation, where only TS and NT points are considered. We considered including TN points as part of the frame area adjustment. The difference in our case lies in the fact that there are now two designations for a target site: sampled and not sampled. Including target but not sampled points in this area adjustment is necessary to avoid underestimation of target area. We calculated the adjusted frame area as

$$\hat{A}'_k = A_k \frac{n_{TS} + n_{TN}}{n_{TS} + n_{NT} + n_{TN}}. \tag{4}$$

The adjusted weights are $w_i^{adj} = \frac{\hat{A}'_k}{n_{TS}}$. These weights still have the desired property of $\sum_{i=1}^{n_{TS}} w_i^{adj} = \hat{A}'_k$.

Consider this simple example: in a certain unit, a field crew samples only 1 target point (1 TS). In addition, they record 1 non-target point (1 NT), and also 98 target points that they did not sample (98 TN) due to time/funding constraints. If we ignore the 98 TN points, we would conclude that the target area is 50% of the original area: $\frac{TS}{(TS+NT)} = \frac{1}{2}$. However, in reality out of 100 total points recorded, 99 of them were target and only one was non-target. Including the 98 TN points, we would conclude that the target area is 99% of the original area: $\frac{(TS+TN)}{(TS+TN+NT)} = \frac{99}{100}$, avoiding a severe underestimate of target area.

If we had used the EPA procedure (naively), we would have

$$w_i^{adj} = \left(\frac{A_k}{n_k}\right) \frac{A_k}{(n_{TS} + n_{NT} + n_{TN})\frac{A_k}{n_k}} = \frac{A_k}{n_{TS} + n_{NT} + n_{TN}}. \tag{5}$$

The $\sum_{i=1}^{n_{TS}} w_i^{adj} \neq \hat{A}'_k$ in this case.

The other way we adjusted the frame area was by using the available GIS vegetation layer. This allowed for distinguishing areas with monotypic stands of vegetation from sediment areas. In the case of water and macroinvertebrates indicators this was not an option because the area or extent of wet area is unknown and had to be estimated. There were no additional points evaluated for wet versus dry, so the typical EPA adjustment was used as described in the previous section.

Table 21: Design weight summary for sediment specimens; Adj Area is adjusted area as described in Section 2.1.2; Area Adj Wts are based on using the adjusted area (Section 2.1.2); Adj Wts are the simple adjustment (Section 2.1.1); GIS Area Adj Wgts are based on using the GIS mapped area (Section 2.1.3). NA were included in the grts draw but with VERY low probability, there were 2 locations in the overdraw list for the first draw, but none in the panelone list.

Stratum	Water Zone	Area	Adj Area	Real Area	Adj Wgts	Area Adj Wgts	GIS Area Adj Wgts
IUC	Flooded	61476	61476	61476	6148	6148	6148
Unit 1	Flooded	1115066	672444	712723	14115	8512	9022
Unit 1	Intermittent	421166	259179	378444	52646	32397	47305
Unit 1	Saturated	542154	18695	40947	542154	18695	40947
Unit 2	Flooded	1227897	855807	712024	61395	42790	35601
Unit 2	Intermittent	199120	0	196964	NA	NA	NA
Unit 2	Saturated	869190	0	19341	NA	NA	NA
Unit 3	Flooded	2327850	1219350	1557277	211623	110850	141571
Unit 3	Intermittent	1697710	1697710	1511327	212214	212214	188916
Unit 3	Saturated	993312	496656	771889	496656	248328	385944
Unit 4A	Flooded	695459	368184	454045	77273	40909	50449
Unit 4A	Intermittent	565978	323416	484063	141494	80854	121016
Unit 4A	Saturated	349753	0	100251	NA	NA	NA
Unit 4B	Flooded	1031848	1031848	1024572	85987	85987	85381
Unit 4B	Saturated	396571	396571	375153	99143	99143	93788
Unit 4C	Flooded	1061746	1061746	607873	132718	132718	75984
Unit 4C	Intermittent	1071764	1071764	1071764	153109	153109	153109
Unit 4C	Saturated	2603827	2297494	2437731	173588	153166	162515
Unit 5	Flooded	2230540	2230540	1950095	247838	247838	216677
Unit 5	Intermittent	794724	794724	741023	158945	158945	148205
Unit 5	Saturated	509687	0	509687	NA	NA	NA
Unit 6	Flooded	1202295	1202295	1080566	150287	150287	135071
Unit 6	Intermittent	1531535	1531535	1519686	306307	306307	303937
Unit 6	Saturated	931298	931298	906332	155216	155216	151055

Table 22: Design weight summary for water specimens; Adj Area is adjusted area as described in Section 2.1.2; Area Adj Wts are based on using the adjusted area (Section 2.1.2); Adj Wts are the simple adjustment (Section 2.1.1). NA entries were not included in the sampling frame.

Stratum	Water Zone	Area	Adj Area	Adj Wgts	Area Adj Wgts
IUC	Flooded	61476	61476	12295	12295
Unit 1	Flooded	1115066	NA	NA	0
Unit 1	Intermittent	421166	NA	NA	0
Unit 1	Saturated	542154	NA	NA	0
Unit 2	Flooded	1227897	818598	204650	136433
Unit 2	Intermittent	199120	0	NA	0
Unit 2	Saturated	869190	NA	NA	0
Unit 3	Flooded	2327850	NA	NA	0
Unit 3	Intermittent	1697710	NA	NA	0
Unit 3	Saturated	993312	NA	NA	0
Unit 4A	Flooded	695459	NA	NA	0
Unit 4A	Intermittent	565978	NA	NA	0
Unit 4A	Saturated	349753	NA	NA	0
Unit 4B	Flooded	1031848	NA	NA	0
Unit 4B	Saturated	396571	NA	NA	0
Unit 4C	Flooded	1061746	99539	176958	16590
Unit 4C	Intermittent	1071764	NA	NA	0
Unit 4C	Saturated	2603827	NA	NA	0
Unit 5	Flooded	2230540	223054	318649	31865
Unit 5	Intermittent	794724	NA	NA	0
Unit 5	Saturated	509687	NA	NA	0
Unit 6	Flooded	1202295	54650	400765	18217
Unit 6	Intermittent	1531535	NA	NA	0
Unit 6	Saturated	931298	NA	NA	0

Table 23: Design weight summary for macroinvertebrate specimens; Adj Area is adjusted area as described in Section 2.1.2; Area Adj Wts are based on using the adjusted area (Section 2.1.2); Adj Wts are the simple adjustment (Section 2.1.1). NA entries were not included in the sampling frame.

Stratum	Water Zone	Area	Adj Area	Adj Wgts	Area Adj Wgts
IUC	Flooded	61476	61476	6148	6148
Unit 1	Flooded	1115066	NA	NA	0
Unit 1	Intermittent	421166	NA	NA	0
Unit 1	Saturated	542154	NA	NA	0
Unit 2	Flooded	1227897	997666	94454	76744
Unit 2	Intermittent	199120	0	NA	0
Unit 2	Saturated	869190	NA	NA	0
Unit 3	Flooded	2327850	NA	NA	0
Unit 3	Intermittent	1697710	NA	NA	0
Unit 3	Saturated	993312	NA	NA	0
Unit 4A	Flooded	695459	NA	NA	0
Unit 4A	Intermittent	565978	NA	NA	0
Unit 4A	Saturated	349753	NA	NA	0
Unit 4B	Flooded	1031848	NA	NA	0
Unit 4B	Saturated	396571	NA	NA	0
Unit 4C	Flooded	1061746	99539	176958	16590
Unit 4C	Intermittent	1071764	NA	NA	0
Unit 4C	Saturated	2603827	NA	NA	0
Unit 5	Flooded	2230540	278818	247838	30980
Unit 5	Intermittent	794724	NA	NA	0
Unit 5	Saturated	509687	NA	NA	0
Unit 6	Flooded	1202295	54650	400765	18217
Unit 6	Intermittent	1531535	NA	NA	0
Unit 6	Saturated	931298	NA	NA	0

Table 24: Design weight summary for root specimens; Adj Area is adjusted area as described in Section 2.1.2; Area Adj Wts are based on using the adjusted area (Section 2.1.2); Adj Wts are the simple adjustment (Section 2.1.1); True Area Adj Wgts are based on using the GIS mapped area (Section 2.1.3).

Stratum	Water Zone	Area	Adj Area	Real Area	Adj Wgts	Area Adj Wgts	GIS Area Adj Wgts
IUC	Flooded	61476	0	0	NA	NA	NA
Unit 1	Flooded	1115066	442622	402343	29344	11648	10588
Unit 1	Intermittent	421166	161987	42722	84233	32397	8544
Unit 1	Saturated	542154	523459	501207	27108	26173	25060
Unit 2	Flooded	1227897	372090	515873	306974	93022	128968
Unit 2	Intermittent	199120	132747	2156	99560	66373	1078
Unit 2	Saturated	869190	869190	849849	66861	66861	65373
Unit 3	Flooded	2327850	1108500	770573	465570	221700	154115
Unit 3	Intermittent	1697710	0	186383	NA	NA	NA
Unit 3	Saturated	993312	496656	221423	993312	496656	221423
Unit 4A	Flooded	695459	327275	241414	347730	163637	120707
Unit 4A	Intermittent	565978	242562	81915	188659	80854	27305
Unit 4A	Saturated	349753	349753	249502	174876	174876	124751
Unit 4B	Flooded	1031848	0	7276	NA	NA	NA
Unit 4B	Saturated	396571	0	21418	NA	NA	NA
Unit 4C	Flooded	1061746	0	453873	NA	NA	NA
Unit 4C	Intermittent	1071764	0	0	NA	NA	NA
Unit 4C	Saturated	2603827	306333	166096	1301914	153166	83048
Unit 5	Flooded	2230540	0	280445	NA	NA	NA
Unit 5	Intermittent	794724	0	53701	NA	NA	NA
Unit 5	Saturated	509687	0	0	NA	NA	NA
Unit 6	Flooded	1202295	0	121729	NA	NA	NA
Unit 6	Intermittent	1531535	0	11849	NA	NA	NA
Unit 6	Saturated	931298	0	24966	NA	NA	NA

8.2 References For Appendix C

1. Lessler and Kalsbeek (1992) Nonsampling errors in surveys. Wiley and Sons

2. Rodhouse, T. J., K. T. Vierling, and K. M. Irvine. (2011). A practical sampling design for acoustic surveys of bats. The Journal of Wildlife Management. DOI: 10.1002/jwmg.151

3. Sarndal et al. (1992) Model assisted survey sampling. Springer Series in Statistics

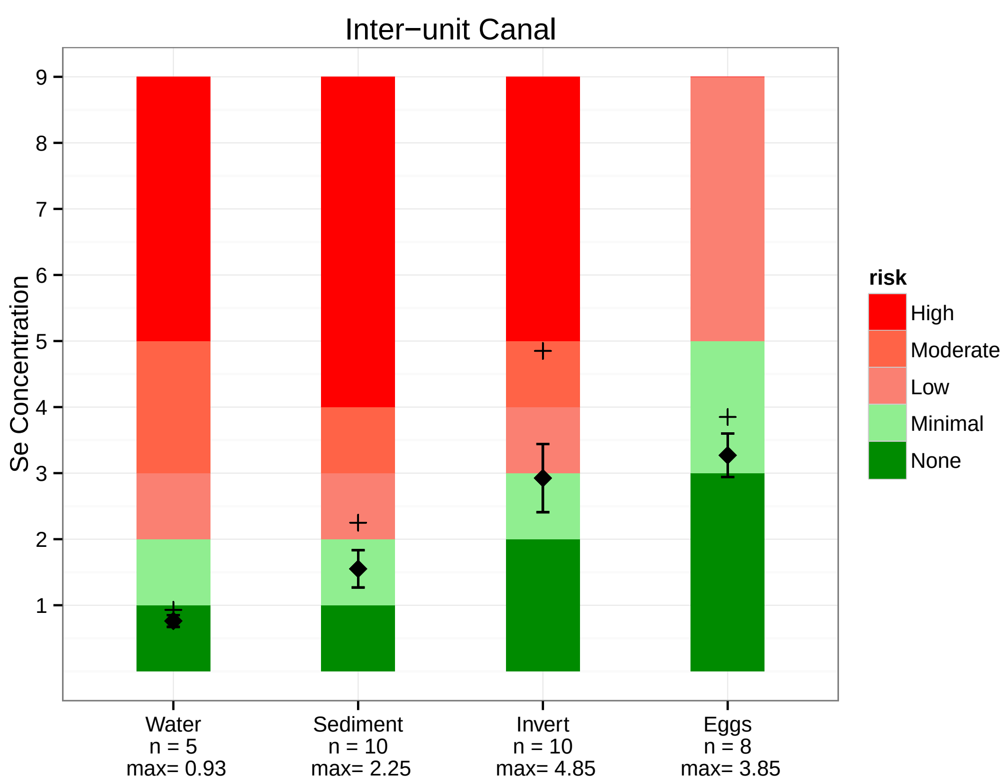
4. Stevens, D. L., Jr. and A. R. Olsen (2004). "Spatially-balanced sampling of natural resources." Journal of American Statistical Association 99(465): 262-278.

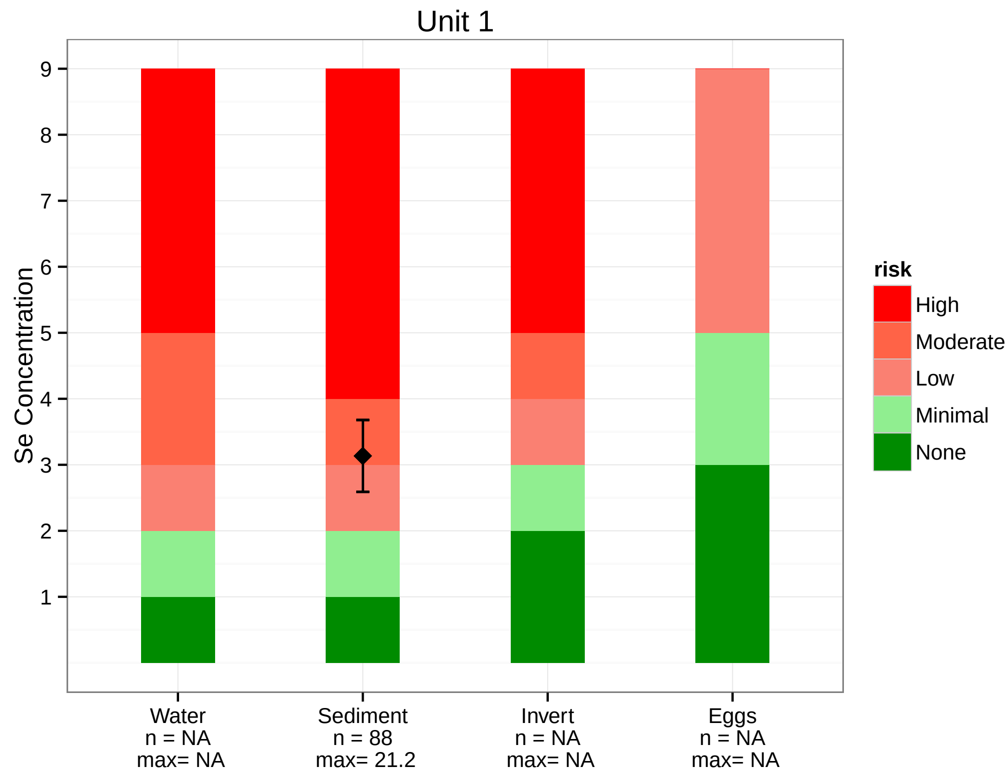
9 Appendix D: Management Risk Plots

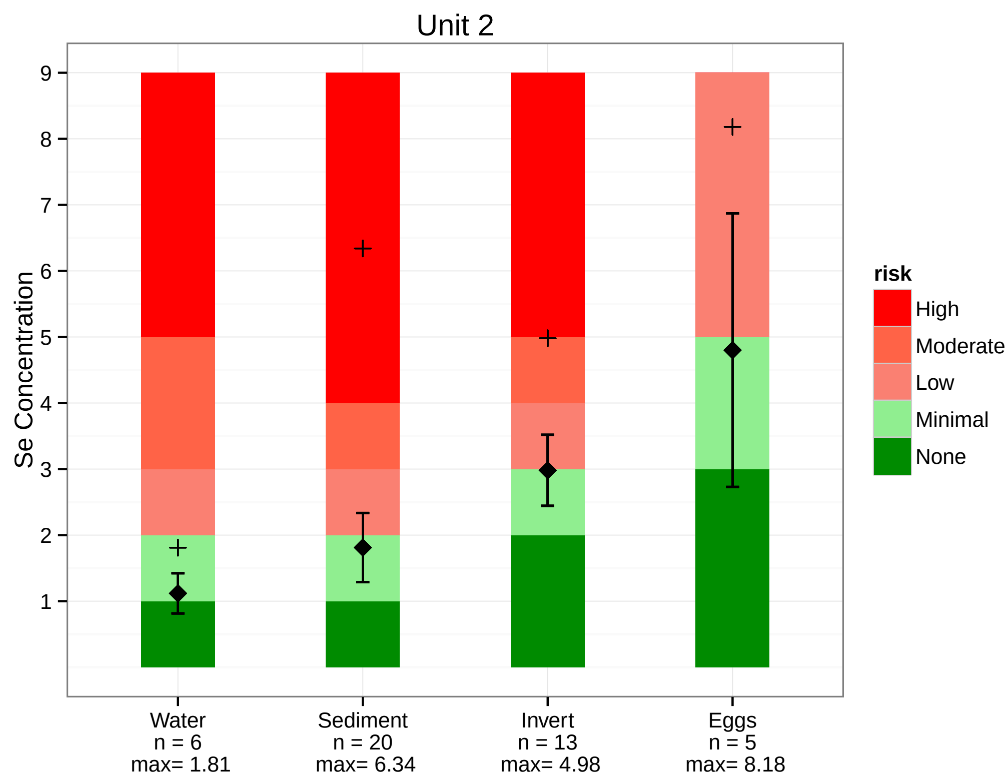
Function use:

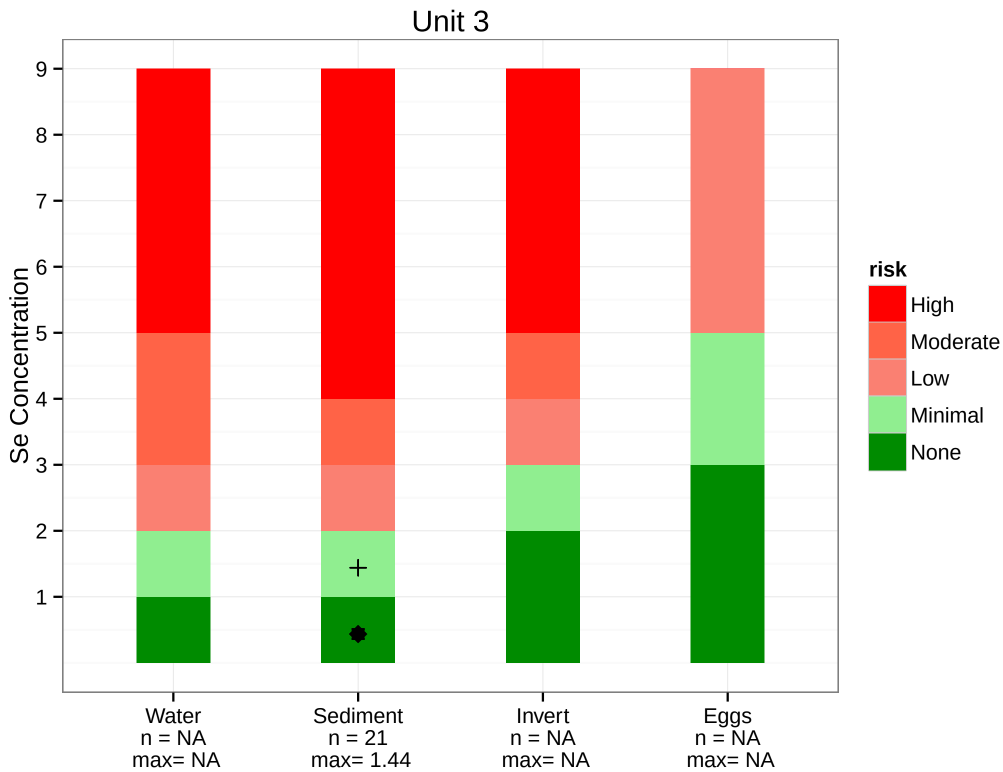
```
# Function takes in the following arguments:
# 1) unit/stratum of interest
# 2) vector of values for eggs, since this is not stored like water,sed,bugs
#    vector is mean, lower CI, upper CI, max, and n
# 3) which estimates should be plotted (bug, sed, water - in that order)
#    for bugs, can be init or adj
#    for sed, can be init, adj, or real
#    for water, can be init or adj

# an example for IUC with all adjusted weights
hazard.plot("Inter-unit Canal",c(3.27,2.94,3.60,3.85,8),c("adj","adj","adj"))
```

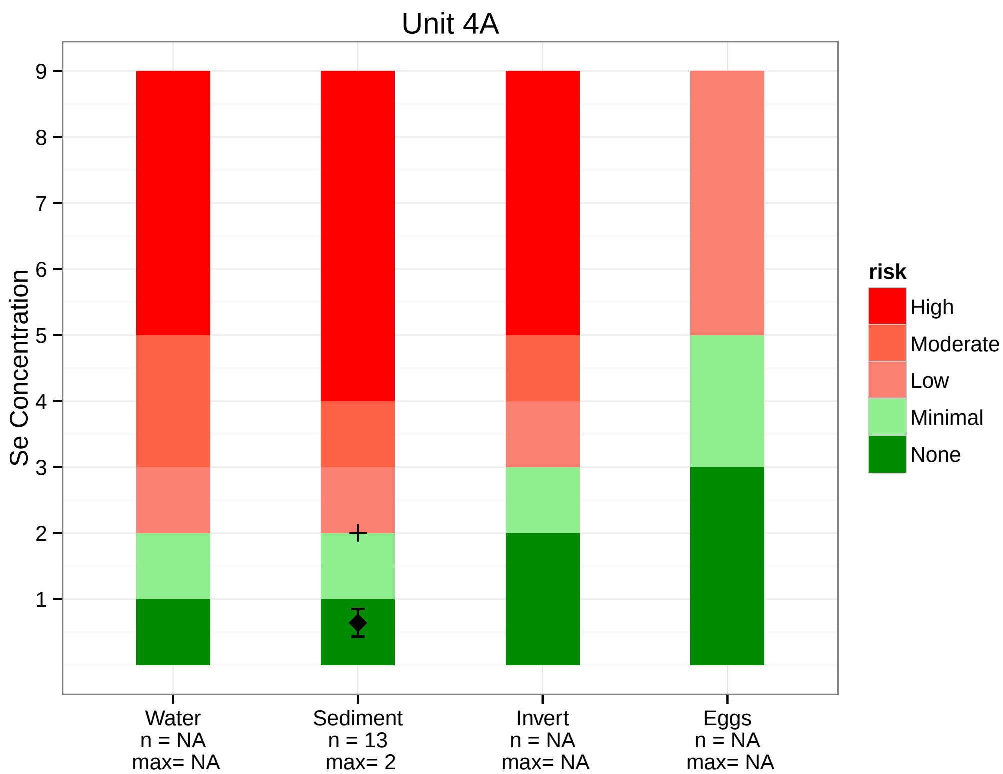
Figure 7: Mean values (◆) with 90% confidence intervals for mean selenium concentrations using estimated-area adjusted weights in water, sediment, macro-invertebrates, and non area adjusted weights for waterfowl eggs for each unit sampled on Benton Lake in 2013. Maximum values are indicated by the (+) symbols. Colored bars represent the associated biological hazard for each trophic level (green colors are below objective, pink-red colors are above objective).

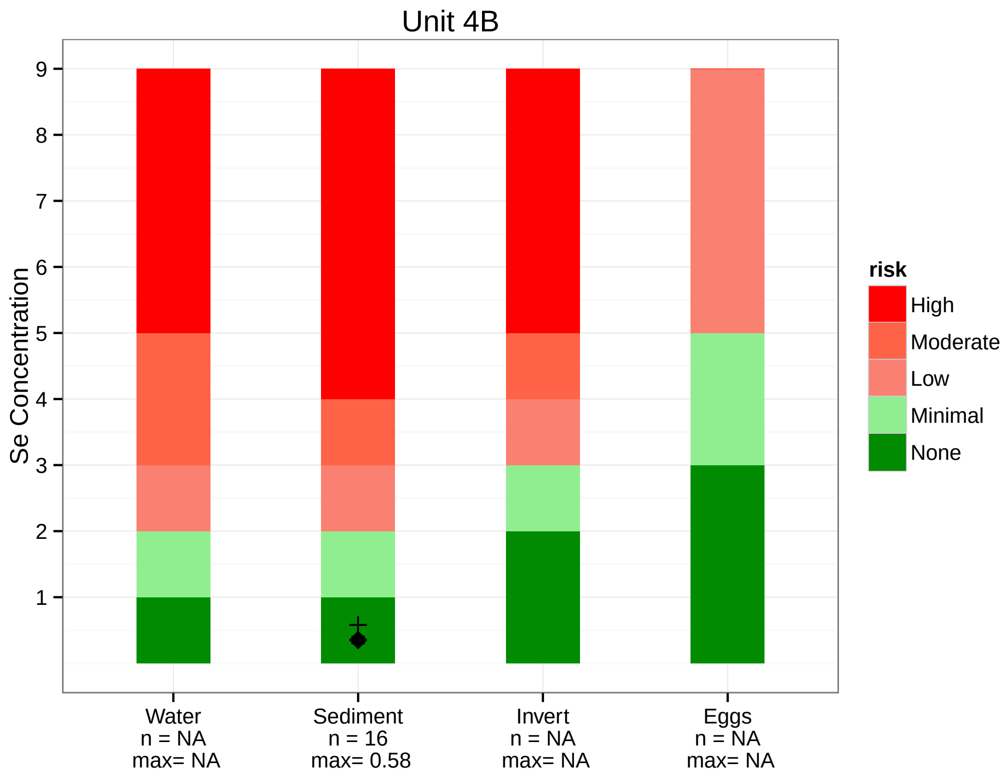



Unit 1
Se Concentration
9
8
7
6
5
4
3
2
1
Water
n = NA
max= NA
Sediment
n = 88
max= 21.2
Invert
n = NA
max= NA
Eggs
n = NA
max= NA
risk
High
Moderate
Low
Minimal
None

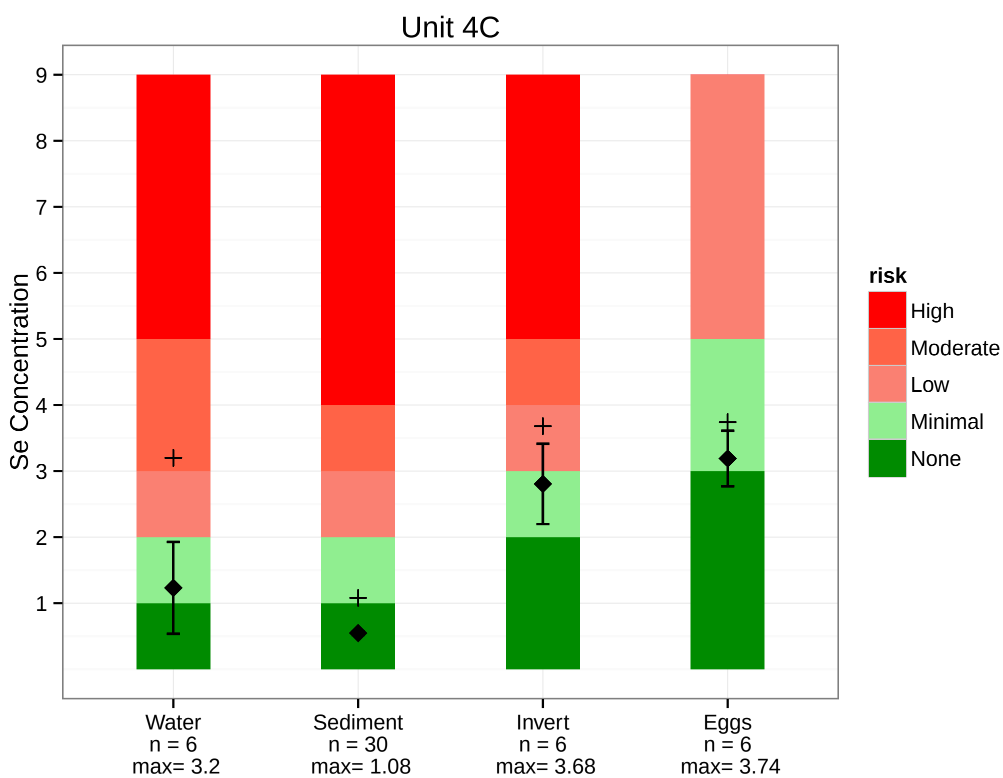

Unit 2
Se Concentration
9
8
7
6
5
4
3
2
1
Water
n = 6
max= 1.81
Sediment
n = 20
max= 6.34
Invert
n = 13
max= 4.98
Eggs
n = 5
max= 8.18
risk
High
Moderate
Low
Minimal
None


Unit 3
Se Concentration
9
8
7
6
5
4
3
2
1
Water
n = NA
max= NA
Sediment
n = 21
max= 1.44
Invert
n = NA
max= NA
Eggs
n = NA
max= NA
risk
High
Moderate
Low
Minimal
None

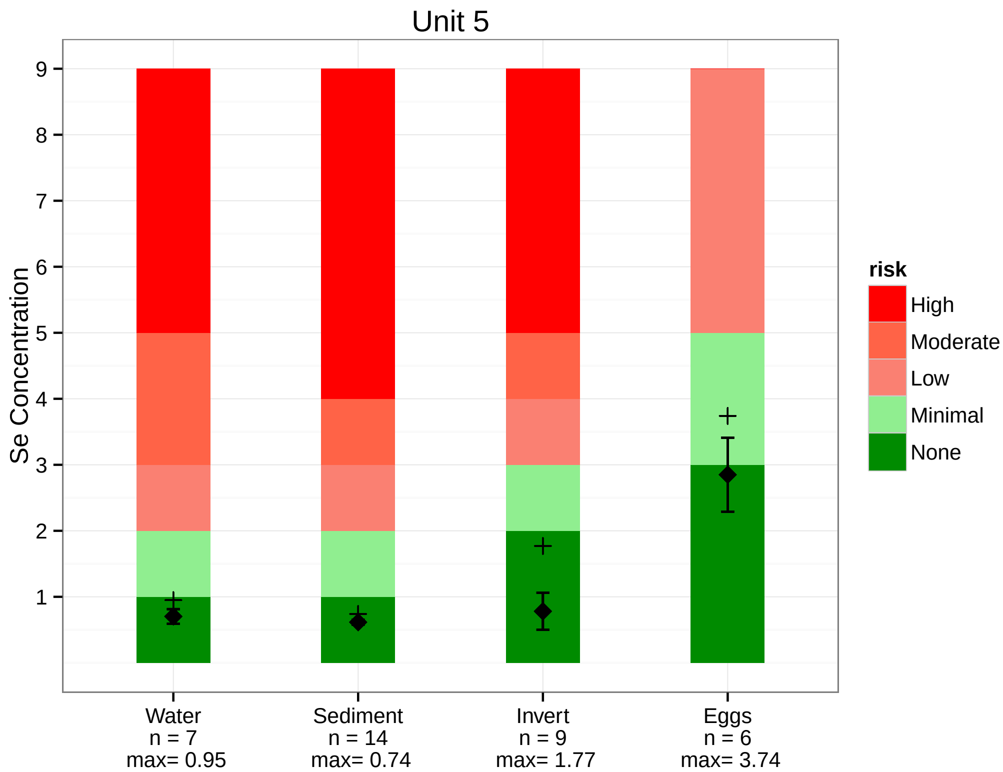

Unit 4A
Se Concentration
9
8
7
6
5
4
3
2
1
Water
n = NA
max= NA
Sediment
n = 13
max= 2
Invert
n = NA
max= NA
Eggs
n = NA
max= NA
risk
High
Moderate
Low
Minimal
None

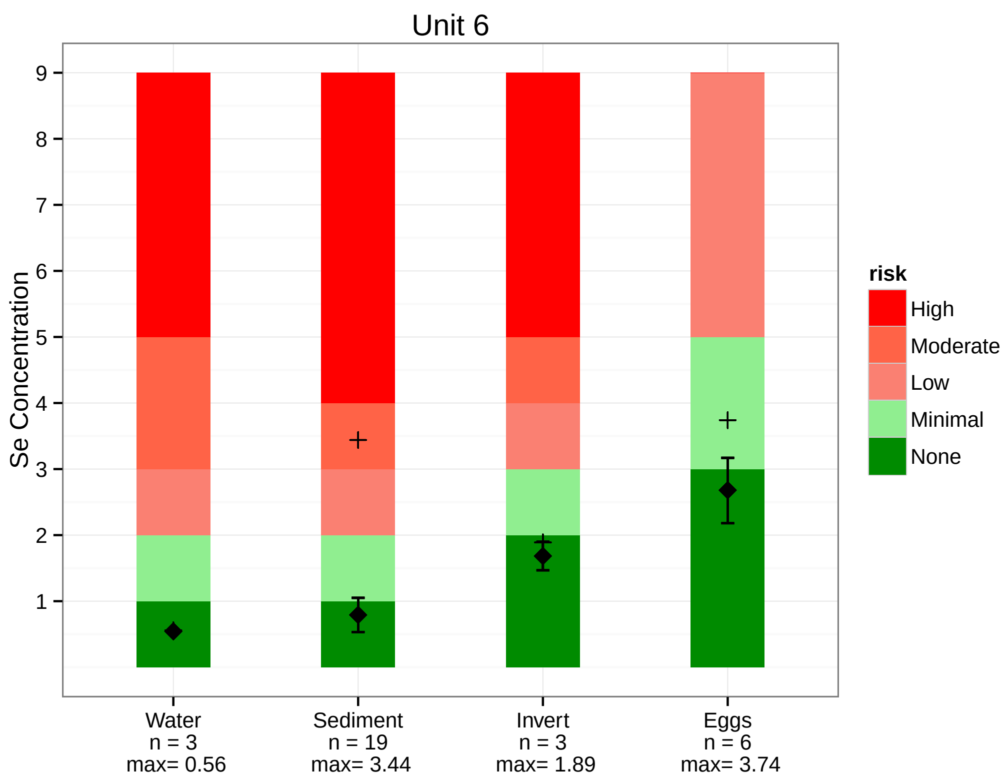


Unit 4B
Se Concentration
9
8
7
6
5
4
3
2
1
Water
n = NA
max= NA
Sediment
n = 16
max= 0.58
Invert
n = NA
max= NA
Eggs
n = NA
max= NA
risk
High
Moderate
Low
Minimal
None



Unit 4C
Se Concentration
9
8
7
6
5
4
3
2
1
Water
n = 6
max= 3.2
Sediment
n = 30
max= 1.08
Invert
n = 6
max= 3.68
Eggs
n = 6
max= 3.74
risk
High
Moderate
Low
Minimal
None


Unit 5
Se Concentration
9
8
7
6
5
4
3
2
1
Water
n = 7
max= 0.95
Sediment
n = 14
max= 0.74
Invert
n = 9
max= 1.77
Eggs
n = 6
max= 3.74
risk
High
Moderate
Low
Minimal
None


Unit 6
Se Concentration
9
8
7
6
5
4
3
2
1
Water
n = 3
max= 0.56
Sediment
n = 19
max= 3.44
Invert
n = 3
max= 1.89
Eggs
n = 6
max= 3.74
risk
High
Moderate
Low
Minimal
None

10 Appendix E: Data Management Procedures and Recommendations

<u>**Evaluation:**</u>

There are four columns for evaluation status, consisting of EvalWater, EvalBugs, EvalSed, and EvalRoot to store the evaluation status for water, macroinvertebrates, sediment, and roots. There is no longer an overall "EvalStatus", since one of the issues we encountered was that something could be evaluated for sediment, but not water (more specific examples given later). For each of the four evaluation statuses, six different evaluation statuses can be input: TS, TN, NT, NN, NE, and NA. Hereafter, "evaluated" refers to whether a point was determined to have water, bugs, sediment, or roots, NOT whether measurements were actually taken. Therefore it is entirely possible for a point to be evaluated, but have no measurements. Separation of evaluated and measured is a necessary distinction in order to correctly quantify information in analysis. The first three evaluation statuses occur when the point is evaluated, and the second three occur when the point is not evaluated.

TS - Target point and sampled.

TN - Target and not sampled. This occurred frequently when points were target for sediment, but not target for anything else. In the field, this resulted in no measurements being taken. Thus even if a point was target for sediment, it may not be sampled. However, this point was still evaluated, and thus gives important information about the extent of the sampling frame area covering the target population .

NT - Not target, not sampled. In order for a point to be deemed not target, it must have been evaluated. In addition, when points are NT, it follows that no measurement for that constituent is recorded.

NN - Not evaluated, but not needed. For not evaluated points, we have many possible reasons as to why they weren't evaluated. One common reason is that the sample was not needed, given that an appropriate amount of samples was already taken and thus more is unnecessary. For future sampling, points could potentially still be evaluated for presence/absence in order to gain more information about the target sampling frame even when they are not needed, resulting instead in either TN or NT samples.

NE - Not evaluated, other. Reasons may vary greatly for this classification. This evaluation status can be further broken down into more specific reasons for not evaluating points, but it is kept simple here. An example occurred in Unit 2 in 2013. Sample points 192 and 209 both indicate as a comment "Not accessible: thick cattails". Clearly sediment and roots were evaluated here: there is no sediment and there are roots. However, due to the inaccessibility of a cattail stand, it is impossible to know whether water and invertebrate samples would have been possible here. Therefore we cannot classify EvalWater or EvalBugs as target or non-target and thus they have not been evaluated.

NA - Not available (truly missing data). This only occurs if a point was truly missed (sample point 33 in 2014), or a recording error is made and the true value is not known.

Example 1) Points that have the comment "Based on veg map, not likely to be sediment" have been evaluated for sediment and roots. In other words, we are confident that there are roots (from cattail/garrison stands) present and therefore no sediment present. However, once again, we have not evaluated the water or invertebrates at this location (they may be present/absent). Therefore, these must be classified as either NE or NN. Specifically, when this happened in 2014, the water and bugs samples were not needed at this point due to already meeting sample size requirements, and thus are classified as NN.

Example 2) Unit 6 presents interesting information. The comments on 58/89 sample points read "dry". These points have been evaluated for water, bugs, sediment, and roots. Clearly there is no water, bugs, or roots, so these are not target (NT) for water, bugs, and roots. Being "dry", there is obviously sediment, so these are target but not sampled (TN) for sediment. Since no measurements were taken, this information is utilized in the frame adjustment only.

Note: EvalSed and EvalRoots are nearly mirror opposites. When one is evaluated, the other has been evaluated as well. If something is target for one, it must then be non-target for the other. This is based on the assumption that we cannot have both roots and enough sediment (2 cm) at a sample point to get accurate measures of both.

Sample Type:

Currently, this category is split up into two columns: **SampleType** and **Sedtype**. In the future, this can be combined into a single column **SampleType**, where a letter indicating each type of sample measured is listed (W,I,S,R).

Example) Selenium measurements are recorded for water, invertebrates, and sediment. The column **SampleType** should consist of: W,I,S. This column allows for data validation between this column and the four evaluation status columns.

Selenium Measures:

Once again, four columns refer to the actual selenium measurements for water, invertebrates, sediment, and roots. Each of the four columns can contain: a numerical measurement, "noData", or NA.

Numerical measure - The actual recorded selenium concentration for either water, invertebrates, sediment, or roots in micrograms/liter. Should correspond with a "TS" in the either EvalWater, EvalBugs, EvalSed, or EvalRoots (depending on column).

noData - This refers to anytime a point is not measured. This occurs when a point is target and not sampled (TN), not target (NT), not needed (NN), or not evaluated (NE). It is important to differentiate this from NA, since this data is not "missing", just not measured.

NA - similar to before, this only occurs when data is truly missing or a mistake in data measurement is made.