To: Jeffrey Kimble, Federal On-Scene Coordinator, U.S. EPA, Region 5
Subject: Enbridge Energy, Limited Partnership (Enbridge) Line 6B, Mile Post (MP) 608, Marshall, MI Pipeline Release
Date: December 5, 2014
Authors: Thomas P. Graan, Ph.D., Weston Solutions, Inc. (WESTON)
Ronald B. Zelt, Professional Hydrologist (Water Quality), U.S. Geological Survey

1.0 Introduction

In March 2013, U.S. Environmental Protection Agency (EPA) released an estimate of the residual volume of Line 6B oil that remained in the 38-mile impacted reach of the Kalamazoo River at the time of sediment sampling in summer 2012. That estimate was based on application of a "state-of-the-science data set" produced using the latest forensic chemistry information and methods available at that time (Graan and Zelt, 2013, p. 10), and was prepared in response to Enbridge Energy's estimate of the residual Line 6B oil volume. The two volume estimates were very different, and Enbridge subsequently retained Geosyntec, Inc., to review both of the estimates. Geosyntec subsequently issued a third estimate of the residual Line 6B oil volume (Geosyntec, 2013). EPA has refined its 2013 estimate because EPA and its contractors continued to develop their forensic chemistry methods research to address some previously unresolved problems with the fingerprinting methods, particularly the performance of the TAS1/T30 biomarkers ratio. Secondly, recent changes in EPA guidance on the recommended handling of censored data have been incorporated into the quantification estimate (Singh and Singh, 2013), and the approach to quantify the vertical extent in the volume quantification model has been refined in view of (1) a sensitivity analysis comparing several methods for quantifying the vertical extent, and (2) new results from July 2014 laboratory analysis of a set of core samples collected in 2012. Third, a refined approach to estimating the uncertainty of the volume estimates was implemented to more appropriately model the range of values that Line 6B oil concentrations can take (i.e., lower limit of uncertainty interval not less than zero) and to avoid the implicit assumption that the concentrations follow a normal (Student's *t*) distribution.

EPA has completed a revision of its report of the Line 6B oil concentrations in the summer 2012 sediment-core samples and, consequently, has refined its estimate of the residual Line 6B oil volume. This technical memorandum addendum presents the 2014 refined estimate of the residual oil volume, discusses the differences in data and methods between the 2013 initial and 2014 refined estimates, and presents some revised conclusions and interpretations of the refined estimates.

2.0 Data and Methods

Several differences between the 2013 initial and 2014 refined estimates of residual Line 6B oil volume relate to key modifications of both the concentration data and data-analysis methods:

1) Although the laboratory analytical determinations of the various hydrocarbon-constituent compounds were not revised between volume estimation studies, there have been (1) additional 2012 core samples analyzed; and (2) modifications in the forensic chemistry method details (described in NewFields, 2014) that affected the reported concentration of Line 6B oil. The changes in the forensic methods affected both the frequency of Line 6B oil detection and the magnitude of its mean concentration, in many cases (see Sec. 2.1 *Changes in Oil Concentration Dataset*).

2) Updated versions of the statistical software and submerged oil volume quantification (SOVQ) spreadsheet calculator were used for the 2014 refined oil-volume estimates. Although the update of ProUCL (ver. 4.1) to version 5.0 did not appear to introduce differences in estimates of the mean concentration computed using the Kaplan-Meier survival-analysis (K-M) methods, there were some small differences in estimates of the standard deviation computed for identical data between versions 4.1 and 5.0.
3) More significantly, the SOVQ spreadsheet calculator methods, and some data records, were revised for the 2014 refined estimates. These changes were made to address some discovered errors in data entry or to improve upon or substitute more appropriate "*ad hoc* decision rules" in view of some modifications to the characteristics of the oil concentrations dataset. In addition, the spreadsheet calculator was revised to handle the refined estimates of uncertainty in the Line 6B oil concentrations and volumes. These changes are discussed in Sec 2.3 *Changes in SOVQ Spreadsheet Tool*.
4) The inclusion in the 2014 refined estimate of a second table of results corresponds to another modification needed, in the view of the authors, to accommodate the increased importance attached to stratum-mean concentrations that were based on fewer than three detectable concentrations of Line 6B oil. Given the inherent unreliability of sample-based estimation of parameters of a population when the sample size is very small, and recognizing that available methods for estimating means from samples containing censored values (non-detections) are not recommended in cases where fewer than 3 to 4 detected values are included in the sample, an alternate approach was warranted. These cases composed 93 (57 percent) of the total 163 discrete vertical increments within the vertical extent of summarized oil volume across all sample strata combined. Details of this selected alternative approach are discussed in Section 2.4 *Analysis of Datasets Including Censored Results*.

2.1 Changes in Oil Concentration Dataset

For the 2014 refinement of the residual oil-volume quantification, the oil-volume equation is unchanged, and input datasets were unchanged for three of the five terms in the equation (eq. 2 in Graan and Zelt, 2013, p. 18). In this section of the addendum, the subject is the oil concentration term in the oil-volume quantification equation.

The oil concentration data for each analyzed sediment core sample resulted from using an improved method for calculating Line 6B oil concentrations, as explained in NewFields (2014). The 2014 dataset used for the 2014 refined oil-volume quantification includes 367 samples collected from 102 cores, after excluding 4 duplicate cores and 22 field sample duplicates. The 2014 data were transmitted from NewFields to EPA in September 2014 (NewFields, 2014).

In addition to changes in the forensic chemistry method, results for 25 recently analyzed 2012 core samples were included in the 2014 refinement that were not part of the EPA (NewFields, 2013) oil concentrations database in 2013. The newly included samples resulted from lab analyses completed in July 2014 to determine concentrations of hydrocarbons and other organic compounds in deeper parts of cores where the deepest sample previously analyzed had been interpreted to contain detectable Line 6B oil. The intention was to fulfill the objective of representing the full, oiled interval of each core, as had been recommended to the EPA FOSC (U.S. EPA SSC, 2012). Moreover, the directive on study design called for analysis of the next-deeper sample of the respective sediment cores involved (U.S. Environmental Protection Agency, 2012b).

Also described by NewFields (2014) are an expanded set of forensic analysis result-level qualifiers. For the purpose of volume quantification, the authors applied the qualifiers as follows.

- Nonqualified (116 results): used Line 6B oil concentration data as reported.
- I-code qualified (2 results): based on GC chromatograph pattern mismatch, oil detected was not Line 6B oil, so treated result as if it were a U-qualified non-detection.

- J-code qualified (23 results): assumed the estimated Line 6B oil concentration represents a detected value as reported. Sample may be discounted if estimated value is well above maximum nonqualified value (but such did not occur in the dataset for this study).
- U-code qualified (186 results): sample is a non-detection and normally reported as “less than” the reported censoring value (the SQL).
- UJ-code qualified (40 results): sample is a non-detection reported as “less than” the reported censoring value (might be lower than the SQL).

Of the 367 samples, 130 (35.4 percent) had detectable concentrations of Line 6B oil. Of the 130 detections in environmental primary samples, 20 samples had detectable Line 6B oil that was less than the sample quantitation limit (SQL) that corresponds to the laboratory preparation (dilutions, etc.) and sample-specific sensitivity that varies with the amount of background hydrocarbons also present in the analyzed sample. There were 23 samples coded as “J-code qualified” detections; although the maximum SQL among these samples was 319 mg/kg, the estimated Line 6B concentration was not more than 81 mg/kg for 20 of these samples. For the other “J-code qualified” detection, the Line 6B concentration was 280 mg/kg, and the vertical extent of this sample is 1 inch. Thus, the effect on oil volume from “J-code qualified” detections was either very limited vertically or associated with relatively low concentrations.

Detection frequency varied from zero (Morrow Lake Light-to-None) to more than 69 percent (Anthropogenic Channel Light-to-None). Other strata with a detection frequency greater than 60 percent included the Heavy-to-Moderate zone of the Depositional Bar stratum and both Cutoff/Oxbow strata (Table 1). In addition, the 2014 oil concentrations dataset has 294 distinct levels used as the SQL, ranging from 13.8 to 1,340 mg/kg, but only 2 distinct SQLs were greater than 513 mg/kg (Figure 1).

Comparison with 2013 dataset. The dataset used for the initial 2013 oil volume estimate consisted of 336 samples collected from 101 cores. There were 254 samples with detected concentrations in the dataset input to the SOVQ calculator. The overall 75.6 percent detection frequency was more than twice that of the detection frequency in the 2014 edition of the concentrations dataset. Two of the Heavy-to-Moderate poling strata (Cutoff/Oxbow and Impoundment geomorphic strata) had a 100 percent detection frequency in the 2013 dataset, and only the Morrow Lake stratum (31.6%) had a detection frequency less than 35 percent. The 2013 concentrations dataset had 76 distinct levels used as limits of detectability (LOD)[1], with LOD ranging from 1 to 1,923 mg/kg, and 24 distinct LOD were greater than 500 mg/kg (with 24 non-detections associated with the latter LODs).

The summary of detection frequency by stratum for the initial 2013 oil-concentrations dataset is given in Table 2, along with concentration percentiles. Percentile statistics and empirical cumulative frequency distribution (ECFD) were calculated using the K-M method on ‘flipped’ data, described by Helsel (2005, p. 63-68). The calculation procedures were those implemented in the S-PLUS “USGS library,” version 4.0 (Lorenz et al. 2011), for S-PLUS software (TIBCO Spotfire S-PLUS, ver. 8.1 for Windows, 2008).

Comparison of the datasets used for the 2013 initial and 2014 refined residual oil-volume estimates was not limited to detection frequencies, but also considered both concentration percentiles and graphical summaries— that is, plots of the ECFD (Figure 2). The selected concentration percentiles (Tables 1 and 2) indicate that the two datasets differ considerably in some characteristics despite having originated from the same set of cores and, with the exception of the 25 newly added samples, essentially the same analytical laboratory determinations.

[1] A discussion of sample quantitation limit (SQL) versus limit of detection (LOD) is presented in NewFields (2014).

- Concentrations in the upper tail of the frequency distribution for the 2014 dataset are about one-half those at corresponding percentiles (90th and 95th) for the 2013 dataset.
- Concentration at the 25th to 75th percentiles of the frequency distribution for the 2014 dataset are less than one-fourth of corresponding percentile concentrations for the 2013 dataset.
- With 95-percent confidence, the datasets differ significantly in their reported percentile concentrations.

These differences are chiefly a result of the modifications to the forensic chemistry approach, discussed in NewFields (2014). The differences between the concentrations and detection frequencies do not result from the processing done using the SOVQ spreadsheet tool; rather, the compared datasets were part of the *inputs* to the spreadsheet tool.

2.2 Changes in ProUCL Software

The 2013 estimates of residual line 6B oil volume used the nonparametric K-M method within ProUCL (ver. 4.1) to estimate a stratum-mean concentration whenever non-detections were present along with three or more detections of Line 6B oil. Some statistical methodology guidance (Helsel, 2005, p. 67) had cautioned that in cases when the smallest SQL among non-detects was less than the smallest detected concentration, there is a small positive bias in the K-M estimated sample mean, but because of the bound at zero concentration, this bias is not large. Guidance released with ProUCL (version 5.0; Singh and Maichle, 2013, p. 36) and another widely used guide (Helsel, 2012) have clarified that the K-M estimated mean in cases with only one SQL present is not equal to a mean obtained when the SQL is substituted for each non-detection, except when all of the detections are greater than the single SQL.

ProUCL (version 5.0) also reports multiple estimates of the upper 95-percent confidence limit (UCL95) of the population mean, including some methods based on K-M estimates that were not present in version 4.1 of ProUCL (Singh and Singh, 2013, p. 114). Although ProUCL reports estimates of K-M based UCLs for datasets containing as few as 3 detected values, the accuracy of these estimates from such small datasets is questionable (Singh and Singh, 2013).

2.3 Changes in SOVQ Spreadsheet Tool

Some of the differences between the 2013 initial and 2014 refined estimates of residual Line 6B oil volume are related to small changes in the SOVQ spreadsheet calculator and more substantial changes to the methods used for managing non-detections and for quantifying the vertical extent term in the quantification equation. As the foregoing sections demonstrate, non-detections compose a substantial fraction of the 2014 refined concentrations dataset, and the handling of these warranted some re-examination. But in the subsection that follows next, explanations are provided for several spreadsheet corrections involving keyed data inputs.

2.3.1 Correction of Erroneous Sample Identifiers

During preparations for the 2014 refined oil-volume estimate, four data-entry errors related to sample identifiers in the 2013-edition SOVQ spreadsheet were discovered. These errors are summarized in Table 3. In two cases, the deepest sample analyzed (core depth interval from 0.9 to 1.5 ft) was overlooked during data entry, affecting the results for cores SEKR3825C701 and SEKR3850C703 in the "ML Fan – Light-None" sample stratum. Although the corrected sample values increased the stratum-mean depth of investigation (2013 method for vertical extent) by 0.15 ft, no Line 6B oil was detected in the affected vertical interval, so the error does not affect evaluations of the depth of nonqualified detection of oil in either core in the 2014 refined estimates. Also, this interval contributed no oil volume to the stratum estimate for either the 2013 or 2014 refined estimate, so this data-entry error was of no consequence.

In another case, the sample identifier (SEKR2175C701S072712D002) was entered with D001 instead of D002, causing no matching concentration to be found, which affected results for a short interval (0.117 ft) of a core in the "Backwater – Light-None" stratum. The omitted sample had detectable Line 6B oil (605 mg/kg). After correcting the error, there was no change in the stratum-mean depth of investigation, but the stratum-mean concentration of Line 6B oil for the affected interval changed from 232 mg/kg in the 2013 estimate to an interval estimate of 150–220 mg/kg in the refined 2014 estimate. Because the vertical interval affected by this data-entry error was small, the estimated Line 6B oil volume for this vertical interval changed only from 1,533 liters (405 gallons) in the 2013 estimate to 575–1,354 liters (152–358 gallons) in the 2014 refined estimate.

In the fourth case, another typographical error caused samples from a different core to be used instead of the two samples from core SEKR3725C709 in sample stratum "Backwater - Heavy-Moderate." Corrected data would have decreased the stratum-mean depth of investigation (DOI) by 0.083 ft, but inclusion of newly analyzed samples actually caused the value to increase by 0.16 ft when evaluated using the 2013 mean DOI method. The affected core actually contained detectable Line 6B oil (652 mg/kg in 2014 dataset, 643 mg/kg in 2013 initial dataset) in a 0.3-ft vertical interval (0.5 to 0.8 ft depth) and, in the 2013 dataset only, also in the 0.417-ft interval lying above it (28 mg/kg); whereas, the erroneously used core's data included detections over a 0.9-ft interval (3 samples, concentration range 32–209 mg/kg) of the 1.3-ft affected length (used for the 2013 estimate). Consequently, the stratum-mean oil concentration for the affected interval would have increased from 322 to 405 mg/kg if the error was corrected, and the estimated Line 6B oil volume for the affected 1.3-ft interval would increase from 5,138 liters (1,357 gal) to 5,492 liters (1,451 gal) if the correction was made. In part, the oil volume was smaller using the erroneous core because most of the interval containing the 209 mg/kg concentration was below the stratum-mean depth of investigation, whereas, for the correct core, the entire investigated interval should have been included. Note that no amended 2013 estimate of oil volume was generated. Using the 2014 concentrations and quantification method changes, a different set of results with lower concentrations are being reported for this core's affected interval (Table 3).

2.3.2 Changes in Quantifying Vertical Extent

The sensitivity of oil volume estimates to the vertical extent factor in the quantification equation varied by stratum (Table 4), largely in relation to the stratum-mean concentration(s) of Line 6B oil in the DVIs near the depth of the maximum vertical extent included in the volume estimate. For the initial 2013 oil volume estimate, there were many cores for which the collected set of sediment samples had partly been analyzed and partly been preserved in an "on hold" status at the laboratory, awaiting further instructions as to their disposition. In its original recommendations to the EPA FOSC, the SSCG (2012; cf. U.S. EPA, 2012a) had suggested characterizing more than the top 2 cm; i.e., collecting at least two additional samples to represent the full oiled interval of each core. Once the first round of samples (including at least 2 samples per core) had been analyzed at the laboratory and forensic methods applied to evaluate the presence of Line 6B oil, additional samples then would be analyzed until a depth was identified in each core at which Line 6B oil no longer was detected. On this basis, the 2013 oil-volume quantification used the depth of investigation as a surrogate for depth of oiled sediment, which was at that time unknown. In July 2014 at EPA's request, a group of 25 "on hold" samples were analyzed to address the missing information about depth of oiled sediment for several cores and to provide analytical results that were subsequently evaluated to determine the concentration of Line 6B oil using the 2014 forensic methods.

During sensitivity analysis conducted in 2014, we realized that if the now-complete dataset was filtered on stratum-mean depth of investigation as determined according to the sample processing protocol directed by EPA (2012b, p. 4; Attachment 1, p. 4-7; Attachment 2, p.2), one could simulate what likely would have been the stopping point for submitting samples and thereby minimize the positive bias in depth of investigation introduced by sometimes long sequences of non-detections (or nonanalyzed intervals) in deeper intervals of some sediment cores. Specifically then, for the upper-bound scenario of the 2014 refined estimates, the vertical extent was quantified using what is herein called the "design depth of investigation" (DDoI), defined as the depth to the base of the next-deepest analyzed sample immediately beneath the deepest sample having a nonqualified detection of Line 6B oil (i.e., none of the U-, J-, or UJ-code-qualified results are nonqualified detections). This DOI is consistent with the EPA directive (2012b, p. 4): "For any sediment core in which the Enbridge Line 6B oil concentration within the initially lowermost sample is above the method detection limit … Enbridge shall submit additional deeper samples from that core for forensic chemical analysis. … Enbridge will continue this procedure for such core until analytical results identify a layer that yields a concentration of Line 6B oil below the method detection limit."

However, the EPA-directed sample processing protocol (U.S. EPA, 2012b, p. 3; cf. U.S. EPA, 2012a) also required that at least the uppermost two samples would be analyzed; that is, the purpose of sampling the top 1-inch interval was to avoid dilution of the freshest sediment deposit, not to define the depth of contamination. So for the oil-volume quantification purpose, the DDoI was evaluated by including at least the two uppermost samples per core. The DDoI was the vertical extent method used for the upper-bound scenario because it yields a depth that typically includes a sampled core interval below the deepest sampled interval containing detectable Line 6B oil, and thus tends to be a conservative estimate for the vertical extent of oiled sediment. The stratum-mean DDoI is the vertical extent factor in the SOVQ spreadsheet calculations that produced the upper-bound scenario of 2014 refined Line 6B oil volume estimates.

Although the actual depth of oiled sediment is yet unknown for many of the collected sediment cores, or only known with relatively coarse resolution, for a lower-bound scenario one might assume that the now-complete set of core sample analytical results was an accurate indication of the vertical distribution of Line 6B oil along each core. Under that assumption then, the vertical extent factor in the quantification equation was estimated simply as the stratum-mean depth to the base of the deepest sample having a nonqualified detection (DoNqD) of Line 6B oil as the forensic result (with nonqualified detection defined identically as for the upper-bound scenario).

Resulting values for the vertical extent factor were compared with "baseline" values for DOI, calculated with the same method used for the 2013 initial volume quantification, but as applied to the 2014 dataset. These comparisons revealed the expected general decrease in the vertical extent factor (Table 4); however, we did not find an equivalent decrease in oil volume, because the Line 6B oil was rarely uniformly distributed in the vertical sequence of sediment samples. For the 2014 lower-bound scenario, stratum-mean vertical extent overall averaged 0.32 ft, or 0.96 ft less than when using the baseline method used for the 2013 initial estimate (Table 4). For the 2014 upper-bound scenario, vertical extent was an average 0.79 ft, or 0.49 ft less than using the 2013 baseline method. This difference in vertical extent means that use of the 2013 method would have resulted in a Line 6B oil volume that was about 28 percent larger overall than what is reported for the 2014 upper-bound scenario, or in absolute volume the difference would be about 24,000 gallons (Table 4). The method-related difference was larger for the lower-bound scenario, where use of the 2013 vertical extent method would have resulted in an oil volume about 89 percent larger than what is reported for the 2014 quantification, or in absolute volume the difference would be about 43,000 gallons (Table 4). The magnitude of these differences indicates that oil volume estimates are quite sensitive to the quantification method used to determine vertical extent.

2.3.2 Changes in Quantifying Uncertainty

The equation for oil-volume quantification includes five factors, but the equation for calculating the composite uncertainty includes terms for each factor plus an additional term included for any pair of factors that are correlated. Inclusion of a term to account for the correlation between oil concentration and sediment bulk density was described in Graan and Zelt (2013, eq. 4, p. 21), and is the first topic of this section.

Given the aforementioned changes in the Line 6B oil concentration dataset, the correlation between oil concentration and sediment bulk density was re-examined for the 2014 refined volume estimate. The 2012 samples analyzed for bulk density were the only ones that could be paired with Line 6B oil concentrations, and 32 such pairs were used. The relation between sediment dry bulk density and Line 6B oil concentration is most linear when a logarithmic scale is used for bulk density as well as oil concentration (Figure 3). Significance of bulk density as a predictor of oil concentration was tested using maximum likelihood methods to estimate the likelihood r^2 (Helsel 2012, p. 225-226; also called the generalized r^2 by Allison 1995). When correlation is the objective, the square root of r^2 is the likelihood *r* correlation coefficient, and the likelihood approach can handle multiply censored data for the dependent variable if the explanatory variable is uncensored (Helsel 2012). Likelihood *r* correlation is implemented in the R language using the "cenreg" function in the 'NADA for R' package (Lee, 2013). Fortunately, there were no censored values in the bulk density data. "The test of the null hypothesis that $r = 0$ is identical to the test for whether the slope equals 0, just as in ordinary linear regression" (Helsel 2012, p. 225).

Bulk density samples generally were collected as the top 0.5 foot of the bed material in 2012. To make the data pairs more consistent with respect to sample depth, the thickness-weighted average Line 6B concentration in the top 0.5 foot of cores collocated with a bulk density sample was calculated for each of 32 core locations paired with a 2012 bulk density sample. The mean concentration result was flagged as a censored value (non-detection) if any of the averaged sample concentrations was qualified as a non-detection.

Results from the "cenreg" function for likelihood *r* correlation indicated that the correlation is marginally significant (likelihood *r* correlation coefficient = –0.302; p-value = 0.081). Since the result indicates marginal significance in the presence of >56% censored values, and we want to be conservative in including the effect of a potentially important relation in our uncertainty estimate, we conclude that the term for this correlation will be included. We used the likelihood *r* correlation coefficient (–0.302) in estimating the uncertainty interval for the lower-bound scenario; however, for the upper-bound scenario where the *ad hoc* decision rules involved replacement of many non-detections by substitution with the censoring level, we estimated the correlation coefficient by substituting the censoring level for the non-detections and calculated Pearson's R coefficient of correlation (R = –0.515). These are the values plotted in Figure 3.

A second change in calculation methods for uncertainty resulted from recognizing that an assumption implicit in 2013 methods had in several cases caused the uncertainty intervals reported *for individual strata* in Graan and Zelt (2013) to show a lower limit of less than zero gallons of Line 6B oil. This was because a symmetrical, Student's *t*-based estimate was used for oil-concentration uncertainty, which involves an implicit assumption that the stratum-mean concentrations follow a frequency distribution that is roughly normal (or *t*-distributed to be more precise). This assumption generally is not appropriate for the 2014 concentration dataset, because after replacement of non-detected results by substitution with either zero or the censoring level (per the *ad hoc* decision rules described in section 2.4.2), the standard deviations reported by either K-M or robust ROS procedures are not accurate. Hence, confidence limits based on those standard deviations also will be inaccurate. Alternatives more appropriate for the type of dataset being analyzed and the often non-normally distributed K-M mean values for oil concentration include parametric approaches (e.g., using the gamma distribution) and non-parametric approaches—such as bootstrap-based estimates to determine the 95-percent confidence interval (i.e., estimation of the 2.5^{th} and 97.5^{th} percentiles of the empirical frequency distribution of the K-M mean).

For the 2014 refined oil-volume estimates, we calculated 95-percent confidence intervals using sample thickness-weighted concentrations, and bootstrap resampling based on K-M means of the weighted concentrations. This weighted-concentrations K-M analysis is similar to, but modified from, the one reported in Helsel (2010). These K-M and bootstrapping statistical analyses were repeated for each stratum and used only the samples that included some part of the depth interval above the lower limit of the stratum vertical-extent factor in the volume quantification (i.e., the DoNqD for lower-bound scenario and the DDoI for the upper-bound scenario). Note that the K-M means of weighted concentrations generated during these uncertainty analyses did <u>not</u> replace the discrete vertical interval (DVI)-specific mean concentrations used for estimating the stratum sum of Line 6B oil volume. Rather, these weighted K-M means were used only to rescale the resulting uncertainty intervals from weighted concentration units to oil volume units, and this was done exclusively at the stratum level, not for the individual DVIs.

2.4 Analysis of Datasets Including Censored Results, or Non-detections

2.4.1 Options for Handling Subsets with Fewer than Three Detections of Line 6B Oil

Given both the generally lower SQLs and the decreased frequency of detections of Line 6B oil in the 2014 refined concentrations dataset, it was reasonable to reassess the situation where the concentration data for an individual discrete vertical increment (DVI) included only one or two detections of Line 6B oil.

With the presence of detectable concentrations of Line 6B oil included in the data for these individual DVI, the stratum-mean concentration for the DVI will be greater than zero, but the presence of non-detections causes much uncertainty about the actual magnitude of the mean. Interval estimates of the stratum mean concentration might be the best one could hope to obtain under these circumstances.

- For the upper-bound of the interval estimate, one workable approach is to substitute the censoring level for each non-detection and calculate the maximum mean concentration possible using the available sample values.
- EPA guidance recommends use of an upper limit of statistical-confidence level (UCL) or other mean-associated statistic to estimate the upper-bound for the mean (Singh and Singh, 2013, p. 26-27). Similarly, the lower-bound of statistical confidence about the mean (LCL) was a desired result from data analysis.
- In cases where at least two detections are present, the KM-based 95-percent UCL reported by ProUCL (ver. 5.0) could potentially provide a second estimate of the upper-bound of the interval estimate. However, the authors noted that some varieties of this UCL method were found to exceed the maximum detected concentration in the summarized data subsets.
- The robust regression-on-order-statistics (ROS) method potentially provides a third estimate of the upper-bound of the interval estimate, but when only two detected values are present, the regression line is fixed by the two sample-concentration data points. While the robust ROS method can estimate the mean and standard deviation using two detections, there are no regression residuals from which to assess the fit for reliability or to estimate the uncertainty of the imputed values used as substitutes for the non-detections.

Ad hoc methods are acknowledged in some EPA guidance as being the best one can do when there are too few data to do anything else (Singh and Singh, 2013, p. 28). For the lower-bound of the interval estimate, at least two options are possible, depending on which of the situations anticipated by the *ad hoc* decision rules was applicable.

- One estimate of the lower-bound of the interval estimate was obtained by substituting zero for each non-detection and calculating the arithmetic mean (i.e., minimum mean concentration possible) using the samples from the respective DVI.
- Given the uniformly small sample sizes for estimation of stratum mean concentrations in this study ($n < 10$), in cases with two detections present the detection frequency always was greater than 20 percent for an individual DVI. Some cases where a single detection was present did involve detection frequencies less than 20 percent. In such cases with a single detection, one approach to estimating the mean is to simply divide the detected concentration by the number of analyzed samples included in the summarized DVI. This is equivalent to a substitution of zero in place of the non-detections, and is herein elsewhere referred to as the minimum mean concentration possible using the samples from the respective DVI.

- As a nonparametric method, K-M does not model variability in the domain less than the lowest SQL, but when multiple SQLs are present, it potentially provides a second estimate of the lower-bound for the mean.

2.4.2 Decision Rules Applied for Managing Non-Detections

The most recent EPA guidance (Singh and Singh, 2013, p. 28) indicates that when the number of detected values is small, it is preferable to use *ad hoc* methods rather than using statistical methods to compute mean concentrations, standard deviations, and other upper limits. By extension to the context of our study, this recommendation applies to lower confidence limits as well. Further, the guidance suggests that for data sets consisting of less than 4 detects (**D**) and for small datasets (e.g., size < 10) with low detection frequency (e.g., < 10%), the project team and the decision makers together should decide on a site-specific basis on how to estimate the average concentration, and it suggests consideration of the median or mode as measures of the central tendency. The following decision rules (**and accompanying codes**) explain the resulting *ad hoc* method that was developed and used to manage non-detections for the 2014 refinement of residual Line 6B oil-volume quantification.

1) For cases where all oil concentrations are detections (stratum rule **1.DA**):
    - Calculate arithmetic mean concentration within the stratum spreadsheet using Excel functions (*average* or equivalent option of the *subtotal* function).

2) Before determining the percentage of censored values or counts of analytical results for comparison with decision rules, remove from consideration any non-detection where the censoring level exceeds the largest detected concentration in the data subset being summarized (code for excluded non-detection is **NDX**). Such a censored value actually contributes no information at all about the sample mean, and software implementing the robust ROS and K-M methods will ignore such values.

    A) For up to 80% censoring:

    - If there are at least 3 detections and multiple SQLs (**2A.3DM**), estimate the mean using K-M method as implemented in ProUCL (ver. 5.0).
    - If there are at least 3 detections and a single SQL (**2A.3DS**), estimate the mean using the robust regression-on-order statistics (ROS) method (used the *cenros* function within the 'NADA for R' package [Lee, 2013]).
    - If there are 1 or 2 detections (**2A.1D** or **2A.2D**), report an interval estimate of the mean. We used the minimum possible mean as the lower-bound and the maximum possible mean as the upper-bound of this interval. The minimum possible mean is the arithmetic mean calculated after substituting a zero value for each non-detection in the DVI subset; the maximum possible mean is the arithmetic mean calculated after substituting the censoring level for each non-detection in the DVI subset.

    B) For > 80% Censoring:

    - Some cases where a single detection was present did involve detection frequencies less than 20 percent. Whether there are detections (**2B.DZ)** or not (**2B.NDZ**), report an interval estimate of the mean. We used the minimum possible mean as the lower-bound and the maximum possible mean as the upper-bound of this interval. The minimum possible mean is the arithmetic mean calculated after substituting a zero value for each non-detection in the DVI subset; the maximum possible mean is the arithmetic mean calculated after substituting the censoring level for each non-detection in the DVI subset.
    - Where there were no detections of Line 6B oil among the samples in a DVI subset, the lower-bound of the interval estimate was zero, per the *ad hoc* method rule (**2B.NDZ**). However, because the zero mean results

from substitution of zero values for analytical non-detections (qualified results) one cannot conclude that the mean Line 6B oil concentration is actually zero, only that the true mean could possibly be as small as zero.

3.0 Results and Discussion

Results of the application of the EPA spreadsheet calculator to estimate the volume of residual Line 6B oil in Summer 2012 are given in Tables 5-A and 5-B. Table 5-A presents the volume estimates for the lower-bound scenario, by stratum, which used the lower-bound of the interval estimate of stratum-mean concentrations whenever an interval estimate was called for (per section 2.4.2 decision rules) and the more conservative method of quantifying the vertical extent of Line 6B oil volume ("maximum depth of nonqualified detection" method, per section 2.3.2). Table 5-B presents results for the upper-bound scenario, by stratum, which used the upper-bound of the interval estimate of stratum-mean concentrations whenever an interval estimate was called for, along with the alternate method of quantifying the vertical extent of Line 6B oil volume ("design depth of investigation" method, per section 2.3.2).

- Total volume in the lower-bound scenario of 2014 refined estimates of Line 6B oil volume is 49,000 gallons.
- Total volume in the upper-bound scenario of 2014 refined estimates of Line 6B oil volume is 86,000 gallons.

3.1 Uncertainty of Estimate of Line 6B Oil Volume

Overall, the range of uncertainty in the estimated volume of Line 6B oil in the refined 2014 estimate varies among strata and also varies between the upper-bound and lower-bound scenarios used for the oil quantification. For the 17 strata in the study design, the ranges of uncertainty are listed in Tables 5-A and 5-B. When expressed as relative uncertainty, i.e., uncertainty given as a percentage of the total volume, the following paragraphs indicate that the range of relative uncertainty was similar between the lower- and upper-bound scenarios of the 2014 refined oil-volume estimate.

For the lower-bound scenario, the uncertainty of the total volume ranges from 19,000 to 101,000 gallons, or from 61% less than to 107% more than the volume estimate. Among the individual strata, the lower-bound scenario's volume estimates have uncertainty limits that range from 0 to 405% of the stratum-total volume, with the average lower confidence limit being 67% less than the total volume and the average upper confidence limit being 142% more than the total.

For the upper-bound scenario, the uncertainty of the total volume ranges from 35,000 to 181,000 gallons, or from 59% less than to 111% more than the volume estimate. Among the individual strata, the upper-bound scenario's volume estimates have uncertainty limits that range from 0 to 455% of the stratum-total volume, with the average lower confidence limit being 62% below the total and the average upper confidence limit being 149% above the total.

The uncertainty limits for the 2014 refined estimate are the sums of the respective stratum-level limits, but the relative (percentage difference from the total) uncertainty values are equal to a volume-weighted average of the relative uncertainty limits for the individual strata. As such, strata where very little Line 6B oil remained in 2012 had an almost insignificant effect on either the overall composite CV or relative uncertainty interval size for the total volume estimate.

3.2 Effects of Changes in Data or Methods on Uncertainty of Oil Volume Estimates

As described in section 2.3.2, there were substantial differences in methods used to estimate uncertainty for the 2014 refined estimates of Line 6B oil volume in comparison with methods used for the 2013 initial estimates. These method changes result in very dissimilar uncertainty limits. In the 2014 refined results, the uncertainty limits are asymmetrical around the volume estimates, whereas for the 2013 initial estimate they were symmetrical.

At least one factor tended to cause smaller uncertainty when expressed in actual oil-volume units than was reported for the 2013 estimates: the oil volumes were smaller overall, and smaller for many individual strata. Because the uncertainty interval is calculated as a function of the combined uncertainty in multiple terms of the oil-volume equation (Graan and Zelt, 2013, p. 18), and the combining of uncertainties is done mathematically using CVs as the standardized "units" (Graan and Zelt, 2013, p. 21), then if the composite CV had been about constant while the overall oil volume decreased, the uncertainty interval in volume units would decrease.

At least one factor tended to cause wider uncertainty intervals in the refined oil-volume estimates than was reported for the 2013 estimates. The choice of a nonparametric method for estimating the stratum-mean oil concentration (K-M method) and percentile-based method for identifying the 95-percent confidence limits around those K-M means (bootstrapping option 'BCa,' which is the adjusted bootstrap percentile method) resulted in asymmetrical confidence intervals that are right-skewed. This was intentional and appropriate for the type and size of sample data at hand.

Finally, there was a third factor that potentially could have tended to cause either smaller or larger uncertainty in the updated oil-volume estimates: as detection limits change, so also does uncertainty in estimates of the statistics of the frequency distribution.

Although the sample-specific quantitation limits were lower in the 2014 concentrations dataset, generally, than in the 2013 dataset, detection frequency also was smaller. Generally, the smaller the detection frequency, the greater the uncertainty of the sample mean concentration. So this third factor likely produced greater uncertainty in the 2014 refined estimate than was indicated for the 2013 initial estimate.

So of the three factors considered here that affect quantification of uncertainty, two tended generally to cause greater relative (percentage) range of the uncertainty interval reported for the 2014 estimate of residual volume of Line 6B oil than was indicated for the 2013 estimate. But the third factor causes the actual range of the uncertainty interval, in gallons, to be narrower for the 2014 estimate than for the 2013 estimate.

3.3 Effects of Changes in Data or Methods on Conclusions and Interpretations

The conclusions reported in Graan and Zelt (2013) were revisited, and differences between them and the conclusions that can be drawn from the refined 2014 estimates of Line 6B oil volume are noted in the following subsections.

3.3.1 Magnitude of 2012 Submerged Line 6B Oil Volume

Clearly, the 2014 refined, lower-bound estimate of oil volume is significantly less than the 2013 initial estimate; whereas, the upper-bound scenario of the 2014 estimate falls within the uncertainty limits of the 2013 initial estimate, and *vice versa*. Encoding and explicitly tracking the method used to estimate the mean concentration for each DVI in the 2014 refined estimates allows more detailed summarization of the results than was done for the 2013 estimate, which enables a clearer understanding of how the estimated total oil volume was distributed by estimation method or detection frequency, as examples. Several factors affecting the differences in oil volume magnitude between 2013 and 2014 estimates relate to changes in data or methods that have been described in this addendum, and these factors affect two of the terms in the volume quantification equation.

First, affecting the oil concentration term are two related factors that together appear to cause the largest decrease in the refined oil volume estimate from the initial 2013 estimate. The magnitude of the Line 6B oil concentration reported for many samples has decreased. The method changes responsible for this decrease, including the abandonment of the biomarkers ratio, TAS1/T30, as a source/quantitation ratio for identification and quantitation of Line 6B oil, are described in NewFields (2014).

The second important factor affecting oil volume is the vertical extent term of the volume quantification equation, which also was affected by a methods change. As discussed in section 2.3.2, after a second round of 2012 core samples was analyzed in July 2014 there was a more nearly complete set of information available from the collected cores. Results from an initial sensitivity test comparing several alternate approaches to quantify vertical extent indicated large sensitivity to replacing the 2013 method, which used a basic evaluation of "depth of investigation" corresponding to the maximum depth of analyzed samples from each sediment core. For the 2014 lower-bound estimate, vertical extent is based on maximum depth of nonqualified detection, which is not actually known for cores with detections in the deepest sample collected from the core, and thus is likely to contain some negative bias. For the 2014 upper-bound estimate, vertical extent is based on design depth of investigation, and because this depth also depends in part on the maximum depth of nonqualified detections in many cases, it also may contain some negative bias, but less pervasively than for the lower-bound scenario.

A third factor, also affecting the oil concentration term, is the much smaller detection frequency in the 2014 dataset (35.4% overall) than in the 2013 initial dataset (75.6% overall). This more than halving of the overall detection frequency multiplied the challenges of estimating stratum-mean concentrations of Line 6B oil, because now there were many more instances with less than three detected concentrations present for a given vertical increment of sediment. Specifically, for 41% (lower-bound scenario) or 63% (upper-bound scenario) of the DVIs contributing nonzero volume to the 2014 refined oil volume estimate, there were fewer than 3 analytical detections of Line 6B oil.

- In the lower-bound scenario, of the 82 DVIs contributing nonzero oil volume[2], 38 (46%) had 4 or more Line 6B oil detections, 14 (17%) had 3 detections, and 30 (37%) had 1 or 2 detections.
- 27% of the lower-bound scenario oil volume estimate is from the 58 DVIs where the stratum-mean concentration was estimated using statistical approaches (either arithmetic mean, K-M mean, or RROS mean).
- In the upper-bound scenario, of the 163 DVIs contributing nonzero oil volume[3], 43 (26%) had 4 or more Line 6B oil detections and 17 (10%) had 3 detections.
- 17.9% of the upper-bound scenario oil volume estimate is from the 70 DVIs where the stratum-mean concentration was estimated using statistical approaches (either arithmetic mean, K-M mean, or RROS mean).

Recall that for cases with less than 3 detections, statistical approaches become unreliable and *ad hoc* approaches become the recommended course. *Ad hoc* decision rules resulted in the substitution of zero (lower-bound scenario) or the substitution of the censoring level (upper-bound scenario) as a sample concentration in many vertical increments of sediment contributing to the oil volume summation for the sampled strata. As the following summary bullets indicate, the large majority of Line 6B oil volume included in the 2014 refined estimate continues to originate from the discrete vertical increments of sampling strata where sample size or detection frequency caused *ad hoc*, rather than statistical, methods to be applicable.

- For the lower-bound scenario, 30 DVIs (34%) had fewer than 3 detections and were instances where substitutions with zero were used for non-detects per the *ad hoc* decision rules.
    - These DVIs contributed 73% of the total lower-bound oil volume.

[2] For the lower-bound scenario there were 88 DVIs within the vertical extent used for volume quantification, i.e., within stratum-mean max.depth of nonqualified detections; but 6 DVIs had zero detections and contributed zero oil volume to the lower-bound oil volume estimate. (Yet in only one case, Morrow Lake stratum, did zero detections correspond to a DVI with more than 2 analytical results.)
[3] For the upper-bound scenario, 163 DVIs were within the vertical extent used for volume quantification, i.e., within stratum-mean design depth of investigation.

- 10.4% of the lower-bound oil volume estimate was contributed by the DVIs with a detection frequency less than 20%; but 67% of that estimate was contributed by the 31 DVIs with a detection frequency less than 40%.
- For the upper-bound scenario, 93 DVIs (57%) had fewer than 3 detections and were instances where substitutions with the censoring limit were used for non-detects per the *ad hoc* decision rules.
    - These DVIs contributed 82% of the total upper-bound oil volume.
- 15.9% of the upper-bound oil volume estimate was contributed by the DVIs with a detection frequency less than 20%; but 57% of that estimate was contributed by the 86 DVIs with a detection frequency less than 40%.

3.3.2 Major Contributing Strata

Based on the initial 2013 estimates, Graan and Zelt (2013, p. 21) concluded that the four strata contributing the most residual Line 6B oil were those that also were the largest in areal extent, and these were all areas where reassessment poling had indicated "Light-None" submerged oil. These four strata occupied 79 percent of the oil-impacted area and contributed 83 percent of the total residual Line 6B oil volume. In comparison, the 2014 refined estimates for these same four most areally extensive strata total 44,000 to 76,000 gallons, or 89–91 percent of the total residual Line 6B oil volume, but no more than 3 of the top 5 most areally extensive strata are now among the top 5 strata ranked by Line 6B oil volume. Instead, the "Channel Deposit – Light-None" stratum contributed the most oil by far: 38,000 gallons (lower-bound scenario; 79 percent of the total) to 62,000 gallons (upper-bound scenario; 72 percent of the total) of residual Line 6B oil.

When standardized by areal extent, the "Depositional Bar – Heavy-Moderate" stratum may have been the most intensive contributor if the upper-bound scenario is indicative (Table 5-B). But if the lower-bound scenario is more indicative, then the four most intensive contributors of Line 6B oil were the Heavy-Moderate areas of the Cutoff/Oxbow and Impoundment geomorphic environments and the more extensive Light-None areas of the Channel Deposit and Depositional Bar environments.

In the 2014 upper-bound scenario, four sample strata contributed more than 1,800 gallons each to the Line 6B oil-volume estimate, accounting for 90 percent of the total volume, and they are from three geomorphic environments, "Channel Deposit," "Depositional Bar" and "Morrow Lake" (Table 5-B). In the 2014 lower-bound scenario, the four strata that contributed the most oil volume (93 percent of the total) were from two of these same geomorphic environments ("Channel Deposit" and "Depositional Bar") plus the "Anthropogenic Channel—Light-None" and "Impoundment—Heavy-Moderate" strata (Table 5-A). In the 2013 estimate, these latter four strata contributed 52 percent of the total oil volume. Focusing on the top 2 strata ranked by oil volume contribution, in the 2013 initial estimate 70 percent of the total oil volume came from the top 2, but for the lower- and upper-bound 2014 scenarios, respectively, 91 and 82 percent of the total volume came from the top 2 strata. Thus, the relative importance of the major contributing strata has risen from what it was in the 2013 estimates. However, most of this proportionate increase is accounted for by one stratum, "Channel Deposit – Light-None," where the relative contribution rose from 43 percent in the 2013 initial estimate to from 72 to 79 percent in the 2014 refined estimates.

3.3.3 Uncertainty of 2012 Submerged Line 6B Oil Volume

The lower- and upper-bound scenarios approach is conservative in recognizing the lower reliability of mean concentrations estimated from small samples (e.g., typically 4 to 6 cores, and up to 8 cores, contributed a sample for a specific vertical increment of sediment) that often contained 50 to 80 percent censoring (non-detections). The interval estimates of stratum-mean concentrations were an improvement in this regard over the methods used for the 2013 initial estimates of stratum-mean concentrations. Though the method may be an improvement, and the 2014 dataset more defensible than ever, the increased frequency of censoring resulted in less than 1,800 gallons (2%) of the Line 6B oil in the upper-bound volume estimate coming from DVIs for which the stratum-mean calculation involved no censored concentration values. The corresponding volume for the lower-bound scenario is <400 gallons (0.8%). However, 12,800 to 13,600 gallons (lower- and upper-bound scenarios, respectively) of additional Line 6B oil came from DVIs with 3 or more detected concentrations where recommended statistical methods allowed estimation of the stratum-mean concentration. For the remaining volume of residual Line 6B oil, the lower- and upper-limits of uncertainty are a more reliable guide to estimate the oil volume as of summer 2012 than are volume estimates based on means calculated using less than 3 detectable concentrations in any discrete vertical interval of sediment.

There is no change in the 2013 general conclusion that uncertainty in the residual Line 6B oil-volume estimates could be reduced by collecting and analyzing additional sediment cores in selected sample strata where both magnitude and uncertainty of the oil-volume estimate are presently large. Following the foregoing analysis of the 2014 oil concentrations dataset, it seems clear that estimates for the "Channel Deposit – Light-None" stratum, in particular, would have benefited most from either a larger sample or a focused follow-up field study, possibly one that used a more finely stratified sampling design.

The second 2013 conclusion pertaining to uncertainty reduction was focused on the deeper intervals of cores collected from the downstream part of Morrow Lake. The subsequent forensic-chemistry research and 2014 oil concentrations dataset show that major progress was made for this concern and there is consistency both vertically and laterally in the updated estimates for the Morrow Lake stratum. No detections of Line 6B oil remain in the refined results for this sample stratum, but oil-volume estimates for this area are qualified, because the presumption of zero Line 6B oil as the stratum-mean concentration is biased low when all that is actually known is that the concentration is less than the quantitation limit.

4.0 References

Allison, P.D., 1995, Survival analysis using the SAS System: A practical guide. Cary, N.Car., SAS Institute, Inc., 292 p.

Graan, T.P., and Zelt, R.B., 2013. U.S. EPA Volume Estimate for Submerged Line 6B Oil in the Kalamazoo River. Marshall, Mich., U.S. EPA, Region V FOSC, Tech. Memorandum, 31 p., appendices. Available online at http://epa.gov/enbridgespill/documents.html. (Attachment 2 to U.S. EPA letter to Enbridge, dated May 8, 2013)

Helsel, D.R., 2005, Non-detects and data analysis: Statistics for censored environmental data. New York, John Wiley, 250 p.

Helsel, D.R., 2010, Summing non-detects: Incorporating low-level contaminants in risk assessment. Integrated Environmental Assessment and Management, v. 6, p. 361-366.

Helsel, D.R. 2012. Statistics for censored environmental data using Minitab and R (2nd ed.). New York, John Wiley, 324 p.

ITRC (Interstate Technology & Regulatory Council). 2012. Groundwater Statistics for Monitoring and Compliance (Draft). Washington, D.C., Interstate Technology & Regulatory Council, Groundwater Statistics for Monitoring and Compliance Team, GSMC-1, 361 p., available online at www.itrcweb.org.

Lee, L., and Helsel, D. R., 2005, Statistical analysis of water-quality data containing multiple detection limits: S-language software for regression on order statistics. Computers & Geosciences, v. 31, p. 1241-1248.

Lee, L., 2013. Package 'NADA for R': Non-detects and data analysis for environmental data, ver. 1.5-6. Comprehensive R Archive Network repository, S-language programs and documentation, 63 p. (http://cran.r-project.org/)

NewFields, 2013. Technical Memorandum – Determination of Line 6B Oil Concentrations in Kalamazoo River Sediments. 8 p., attachments.

NewFields, 2014. Technical Memorandum – Determination of Line 6B Oil Concentration in Kalamazoo River Sediments. 12 p., tables.

Singh, A., and Singh, A.K., 2013, ProUCL version 5.0.00 Technical Guide. Washington, D.C., U.S. Environmental Protection Agency, Office of Research and Development, EPA/600/R-07/041, 312 p.

U.S. Environmental Protection Agency, 2012a, Validation of existing core processing procedure for identification of oil in sediment cores. Tech. memo, 3 p., tables.

U.S. Environmental Protection Agency, 2012b, Directive to complete submerged oil quantification in response to the Removal Administrative Order issued by U.S. EPA on July 27, 2010, pursuant to 311(c) of the Clean Water Act etc.: U.S. EPA, Region 5, letter to Enbridge, Nov. 20, 2012, 6 p., attachments.

U.S. Environmental Protection Agency, 2009, Statistical analysis of groundwater monitoring data at RCRA facilities: Unified guidance. U.S. Environmental Protection Agency, EPA 530/R-09-007, 888 p., available online at http://www.itrcweb.org/gsmc-1/Content/Resources/Unified%20Guidance%202009.pdf.

U.S. Environmental Protection Agency, SSC (Scientific Support Coordinator), 2012, Recommendation to the FOSC: Submerged oil volume quantification sampling design and methods for sediment sampling, processing, and analysis. Tech. memo, 52 p. Transmitted August 8, 2012, with cover from T. Graan, coordinator.

5.0 Tables and Figures

5.1 Tables

Table 1. Percentiles of residual Line 6B oil concentrations and summary by sample stratum of sample counts, detection frequency, sediment core subsamples, summer 2012, Kalamazoo River, Calhoun and Kalamazoo Counties, Michigan: Data from forensic analysis method (NewFields, 2014).

[mg, milligram; kg, kilogram; mg/kg, milligram per kilogram; NA, confidence interval not analyzed; collapsed poling categories: HVY-MOD, heavy to moderate; LGT-NON, light to none. (Method: Kaplan-Meier [mdlnpar] function in USGS Library (ver. 4.0) for S-PLUS)]

	Number of detections (total samples)	Concentration at percentile indicated, in mg/kg (and 95-percent confidence interval of percentile value)					
		10th	25th	50th	75th	90th	95th
Residual Line 6B oil concentration, in mg oil per kg bed sediment (dry weight)	130 (367)	< 0.6 (NA)	5.8 (NA – 5.9)	19.7 (7.3 – 26.1)	95.3 (55.6 – 159)	549 (332 – 652)	940 (NA)

Sample stratum (note poling categories collapsed)	Number of detections	Number of samples	Detection frequency (%)
ANTHROPOGENIC CHANNEL - HVY-MOD	4	11	36.4
ANTHROPOGENIC CHANNEL - LGT-NON	9	13	69.2
BACKWATER - HVY-MOD	8	22	36.4
BACKWATER - LGT-NON	5	24	20.8
CHANNEL DEPOSIT - HVY-MOD	8	19	42.1
CHANNEL DEPOSIT - LGT-NON	11	23	47.8
CUTOFF/OXBOW - HVY-MOD	16	24	66.7
CUTOFF/OXBOW - LGT-NON	11	17	64.7
DELTA - HVY-MOD	11	28	39.3
DELTA - LGT-NON	7	27	25.9
DEPOSITIONAL BAR - HVY-MOD	12	19	63.2
DEPOSITIONAL BAR - LGT-NON	5	23	21.7
IMPOUNDMENT - HVY-MOD	13	33	39.4
IMPOUNDMENT - LGT-NON	4	31	12.9
LAKE - LGT-NON	0	19	0.0
ML FAN - HVY-MOD	1	7	14.3
ML FAN - LGT-NON	5	27	18.5
Total	130	367	35.4

Table 2. Percentiles of residual Line 6B oil concentrations and summary by sample stratum of sample counts, detection frequency, sediment core subsamples, summer 2012, Kalamazoo River, Calhoun and Kalamazoo Counties, Michigan: Data from forensic analysis method (NewFields, 2013).

[mg, milligram; kg, kilogram; mg/kg, milligram per kilogram; NA, confidence interval not analyzed; collapsed poling categories: HVY-MOD, heavy to moderate; LGT-NON, light to none. (Method: Kaplan-Meier [mdlnpar] function in USGS Library (ver. 4.0) for S-PLUS)]

	Number of censored values (total)	Concentration at percentile indicated, in mg/kg (and 95-percent confidence interval of percentile value)						
		5th	10th	25th	50th	75th	90th	95th
Residual Line 6B oil concentration, in mg oil per kg bed sediment (dry weight)	82 (336)	7 (NA)	9 (7 – 12)	28 (20 – 37)	104 (79 – 159)	456 (347 – 544)	1,009 (820 – 1,340)	1,856 (NA)

Sample stratum (collapsed poling categories)	Number of detections	Number of samples	Detection frequency (%)
ANTHROPOGENIC CHANNEL - HVY-MOD	6	10	60.0
ANTHROPOGENIC CHANNEL - LGT-NON	9	13	69.2
BACKWATER - HVY-MOD	11	16	68.8
BACKWATER - LGT-NON	8	19	42.1
CHANNEL DEPOSIT - HVY-MOD	13	19	68.4
CHANNEL DEPOSIT - LGT-NON	19	21	90.5
CUTOFF/OXBOW - HVY-MOD	18	18	100.0
CUTOFF/OXBOW - LGT-NON	12	14	85.7
DELTA - HVY-MOD	21	27	77.8
DELTA - LGT-NON	20	27	74.1
DEPOSITIONAL BAR - HVY-MOD	13	15	86.7
DEPOSITIONAL BAR - LGT-NON	21	23	91.3
IMPOUNDMENT - HVY-MOD	32	32	100.0
IMPOUNDMENT - LGT-NON	29	31	93.5
LAKE - LGT-NON	6	19	31.6
ML FAN - HVY-MOD	4	7	57.1
ML FAN - LGT-NON	12	25	48
Total	254	336	75.6

Table 3. Correction of data-entry errors modified the stratum-mean values of depth of oil-volume investigation and Line 6B oil concentration in the affected cores. (For context, the table also includes entries [shaded] for the two core intervals containing the deepest samples containing nonqualified detections of Line 6B oil in the 2014 concentrations dataset and the associated stratum-mean Line 6B oil concentration.)

Correct core identifier	Core depth investigated (ft)		Stratum-mean depth of oil-volume investigation (ft)		Sample stratum
	2013 version	2014 refined version	2013 version	Corrected value (and 2014 version)	
SEKR2175C701	0.6	0.6	1.07	1.07 (0.58)	Backwater – Light-None
SEKR3725C709	1.3	1.8	1.12	1.28 (0.88)	Backwater – Heavy-Moderate
SEKR3825C701	0.9	1.5	0.89	1.04 (0.40)	ML Fan - Light-None
SEKR3850C703	0.9	1.5	0.89	1.04 (0.40)	ML Fan - Light-None
SEKR1500C701	2.8	2.8	1.30	-- (1.13)	Channel Deposit - Light-None
SEKR1550C701	2.6	2.6	1.86	-- (1.11)	Impoundment - Heavy-Moderate

Correct core identifier	Affected discrete vertical increments, as interval below sediment surface (ft)	Stratum-mean Line 6B oil concentration (mg/kg) for affected discrete vertical increments		Sample stratum
		2013 version	2014 refined version	
SEKR2175C701	0.083 – 0.2	232	150 - 220[1]	Backwater – Light-None
SEKR3725C709	0 – 1.3	322	158 – 194[1]	Backwater – Heavy-Moderate
SEKR3825C701	0.9 – 1.5	No data; and deeper than mean DOI[2]	0 – 110[1,3] (no detections)	ML Fan - Light-None
SEKR3850C703	0.9 – 1.5	No data; and deeper than mean DOI[2]	0 – 110[1,3] (no detections)	ML Fan - Light-None
SEKR1500C701	2.4 – 2.8	--	71.4[3]	Channel Deposit - Light-None
SEKR1550C701	2.1 – 2.6	--	135 – 180[1,3]	Impoundment - Heavy-Moderate

[1] Interval estimate of stream mean concentration, estimated as described in "2.4 Analysis of Datasets Including Censored Results" section of the memorandum.
[2] Mean DOI is the stratum-mean depth of investigation, below which no oil concentrations or volumes were included in the 2013 initial estimates of Line 6B oil volume.
[3] Contributed zero oil volume to the 2014 refined estimate of Line 6B residual oil volume, because vertical interval involved is completely below the vertical extent of oil volume summation analysis.

Table 4. Summary by sampling stratum of sensitivity of stratum-total values of Line 6B oil volume to method used for quantifying vertical extent factor in the oil-volume equation.

	Baseline estimates using depth of investigation as vertical extent			Alternative surrogates for vertical extent		2014 Estimates using alternative surrogates for vertical extent	
Stratum name[*]	Line 6B oil volume in lower-bound scenario (liters)	Line 6B oil volume in upper-bound scenario (liters)	Mean depth of investigation (feet)[†]	Lower-bound mean max. depth of nonqualified detection of Line 6B oil (feet)	Upper-bound mean design depth of investigation (feet)	Line 6B oil volume in lower-bound scenario (liters)	Line 6B oil volume in upper-bound scenario (liters)
Anthropogenic Channel - Hvy-Mod	50	103	1.33	0.82	1.33	50	103
Anthropogenic Channel - Lgt-None	4,563	4,865	0.45	0.25	0.45	2,514	4,865
Backwater - Hvy-Mod	2,764	3,400	1.28	0.42	0.88	934	2,751
Backwater - Lgt-None	3,368	8,047	1.32	0.15	0.58	1,273	4,823
Channel Deposit - Hvy-Mod	3,468	3,868	0.98	0.33	0.73	2,336	3,694
Channel Deposit - Lgt-None	280,372	293,065	1.43	0.75	1.13	144,937	235,245
Cutoff/Oxbow - Hvy-Mod	1,587	1,653	1.13	0.53	0.87	928	1,417
Cutoff/Oxbow - Lgt-None	2,520	2,549	0.78	0.47	0.78	1,891	2,549
Delta - Hvy-Mod	2,393	2,823	1.43	0.38	0.86	1,684	2,364
Delta - Lgt-None	1,938	3,249	1.42	0.63	1.00	1,402	2,316
Depositional Bar - Hvy-Mod	8,072	8,124	1.02	0.45	0.85	659	6,975
Depositional Bar - Lgt-None	30,541	34,820	1.50	0.45	0.88	22,154	32,864
Impoundment - Hvy-Mod	2,723	3,510	1.86	0.64	1.11	2,451	2,908
Impoundment - Lgt-None	538	2,023	1.90	0.17	0.96	115	801
Lake - Lgt-None	0	32,890	1.23	0.	0.62	0	18,420
Ml Fan - Hvy-Mod	258	405	0.63	0.028	0.50	86	377
Ml Fan - Lgt-None	2,677	10,164	1.04	0.071	0.40	835	2,677
Total oil or overall mean depth[‡]	347,832	415,558	1.28	0.32	0.79	184,249	325,150
Total oil, in gallons	91,888	109,779	--	--	--	48,674	85,895

[*] Submerged oil category (part of 'Stratum name'): Hvy-Mod, heavy-to-moderate; Lgt-None, light-to-none.
[†] Method used for 2013 initial oil-volume estimate.
[‡] Overall mean depth is a weighted mean, where areal extents of strata are the weights.

Table 5-A. Summary by sample stratum of 2014 refined estimates of residual Line 6B oil, lower bound.

[H-M, heavy-moderate (indication of submerged oil observed after poling agitation of bed sediment); L-N, light-none (indication of submerged oil after poling agitation); mg/kg, milligrams oil per kilogram sediment (dry weight); --, not estimated]

Stratum name	Number of cores	Frequency of detection of Line 6B oil (percent)	Volume of Line 6B oil (gallons)	Uncertainty of oil volume estimate, lower limit (gallons)	Uncertainty of oil volume estimate, upper limit (gallons)	Volume of Line 6B oil, standardized (gallons/ acre)	Vertical extent, as mean max. depth of nonqualified detection of Line 6B oil (feet)	Mean concentra-tion of Line 6B oil (mg/kg)[1]	Stratum areal extent (acres)
Anthropogenic Channel - H-M	3	36.4	13	3	52	19.8	0.82	63.1	0.7
Anthropogenic Channel - L-N	6	69.2	664	323	1,286	8.6	0.25	90.1	77.3
Backwater - H-M	6	36.4	247	91	780	13.3	0.42	166.	18.5
Backwater - L-N	6	20.8	336	0	675	4.3	0.15	150.	77.7
Channel Deposit - H-M	6	42.1	617	236	1,213	27.3	0.33	284.	22.6
Channel Deposit - L-N	6	47.8	38,288	17,164	70,679	72.8	0.75	337.	525.6
Cutoff/Oxbow - H-M	6	66.7	245	111	551	43.6	0.53	352.	5.6
Cutoff/Oxbow - L-N	6	64.7	500	236	901	26.1	0.47	238.	19.2
Delta - H-M	8	39.3	445	211	669	12.9	0.38	152.	34.6
Delta - L-N	6	25.9	370	143	990	8.5	0.63	60.5	43.8
Depositional Bar - H-M	6	63.2	174	77	331	21.9	0.45	192.	8.0
Depositional Bar - L-N	6	21.7	5,853	0	20,170	52.7	0.45	465.	111.0
Impoundment - H-M	7	39.4	648	306	1,423	39.6	0.64	534.	16.4
Impoundment - L-N	7	12.9	30	0	123	0.7	0.17	35.2	43.7
Lake - L-N	6	0.0	0	0	0	0.0	0.00	--	592.7
ML Fan – H-M	3	14.3	23	0	57	9.3	0.028	1,883.	2.4
ML Fan - L-N	8	18.5	221	0	630	1.2	0.071	97.4	180.2
TOTAL or GRAND MEAN	102	35.4	48,674	18,900	100,531	27.3	0.32[2]	--	1,780

[1] Thickness-weighted mean of stratum-mean concentrations of discrete vertical intervals within vertical extent of oil volume summation analysis (depth given in previous column, "Vertical extent, as mean maximum depth of nonqualified detection of Line 6B oil"). These means were not used in calculating oil-volume estimates and are presented only to inform the reader.
[2] Overall depth is the area-weighted mean of the 17 strata-mean depths.

Table 5-B. Summary by sample stratum of 2014 refined estimates of residual Line 6B oil, upper bound.

[H-M, heavy-moderate (indication of submerged oil observed after poling agitation of bed sediment); L-N, light-none (indication of submerged oil after poling agitation); mg/kg, milligrams oil per kilogram sediment (dry weight); --, not estimated]

Stratum name	Number of cores	Frequency of detection of Line 6B oil (percent)	Volume of Line 6B oil (gallons)	Uncertainty of volume estimate, lower limit (gallons)	Uncertainty of volume estimate, upper limit (gallons)	Volume of Line 6B oil, standardized (gallons/ acre)	Vertical extent, as mean design depth of investigation (feet)	Mean concentra-tion of Line 6B oil (mg/kg)[12]	Stratum areal extent (acres)
Anthropogenic Channel - H-M	3	36.4	27	10	118	40.9	1.33	80.4	0.7
Anthropogenic Channel - L-N	6	69.2	1,285	715	2,405	16.6	0.45	96.9	77.3
Backwater - H-M	6	36.4	727	392	1,615	39.3	0.88	231.	18.5
Backwater - L-N	6	20.8	1,274	0	2,686	16.4	0.58	146.	77.7
Channel Deposit - H-M	6	42.1	976	429	2,161	43.2	0.73	204.	22.6
Channel Deposit - L-N	6	47.8	62,145	29,541	106,657	118.2	1.13	362.	525.6
Cutoff/Oxbow - H-M	6	66.7	374	210	857	66.6	0.87	327.	5.6
Cutoff/Oxbow - L-N	6	64.7	673	406	1,207	35.2	0.78	191.	19.2
Delta - H-M	8	39.3	625	405	914	18.1	0.86	94.9	34.6
Delta - L-N	6	25.9	612	290	1,717	14.0	1.00	63.3	43.8
Depositional Bar - H-M	6	63.2	1,843	435	7,888	231.8	0.85	1,074.	8.0
Depositional Bar - L-N	6	21.7	8,682	1,404	39,542	78.2	0.88	349.	111.0
Impoundment - H-M	7	39.4	768	415	1,825	47.0	1.11	364.	16.4
Impoundment - L-N	7	12.9	212	0	612	4.8	0.96	43.7	43.7
Lake - L-N	6	0.0	4,866	0	9,506	8.2	0.62	136.	592.7
ML Fan - H-M	3	14.3	100	46	153	40.8	0.50	459.	2.4
ML Fan - L-N	8	18.5	707	324	1,389	3.9	0.40	55.3	180.2
TOTAL or GRAND MEAN	102	35.4	85,895	35,023	181,250	48.3	0.79[13]	--	1,780

[12] Thickness-weighted mean of stratum-mean concentrations of discrete vertical intervals within vertical extent of oil volume summation analysis (depth given in previous column, "Vertical extent, as mean design depth of investigation"). These means were not used in calculating oil-volume estimates and are presented only to inform the reader.
[13] Overall depth is the area-weighted mean of the 17 strata-mean depths.

5.2 Figures

Figure 1. Boxplot diagrams of frequency distributions in 2014 dataset of concentration of Line 6B oil (n = 367 samples) in bed-sediment, summer 2012, Kalamazoo River; and sample-specific quantitation limit (SQL) for Line 6B oil (n = 294 levels).
(Sediment mass expressed as dry weight)

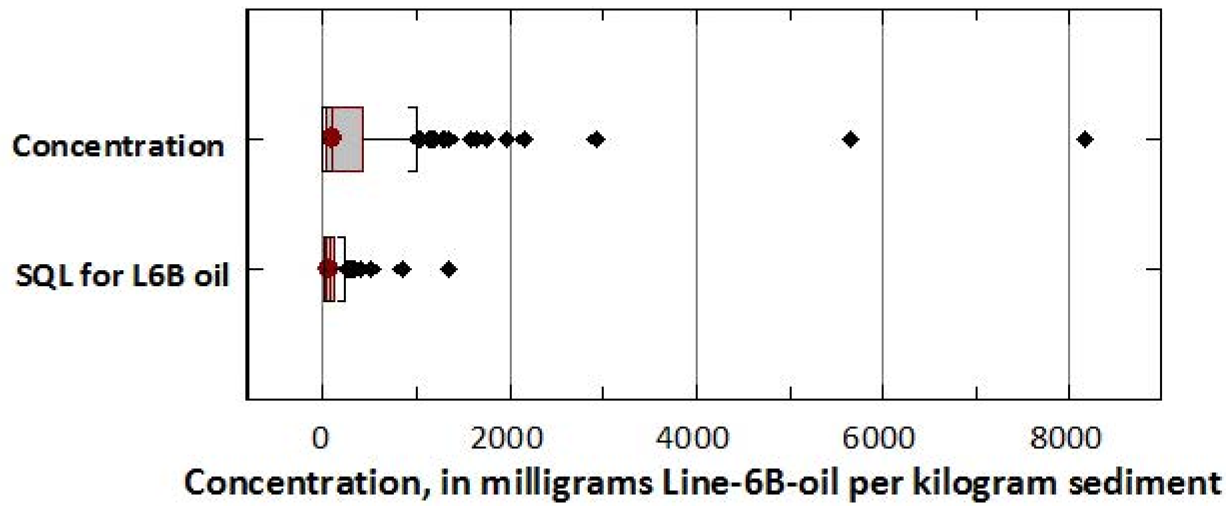

Figure 2. Graphical summaries of Line 6B oil concentrations used for calculation of residual oil volume in Kalamazoo River, summer 2012, shown as empirical cumulative frequency distributions of (A) 2013 concentrations dataset; and (B) 2014 concentrations dataset.

(Concentration units are milligrams oil per kilograms sediment dry weight. Cumulative frequency distributions were estimated using nonparametric Kaplan-Meier method, as implemented in the S-PLUS "USGS library," version 4.0 [Lorenz et al. 2011])

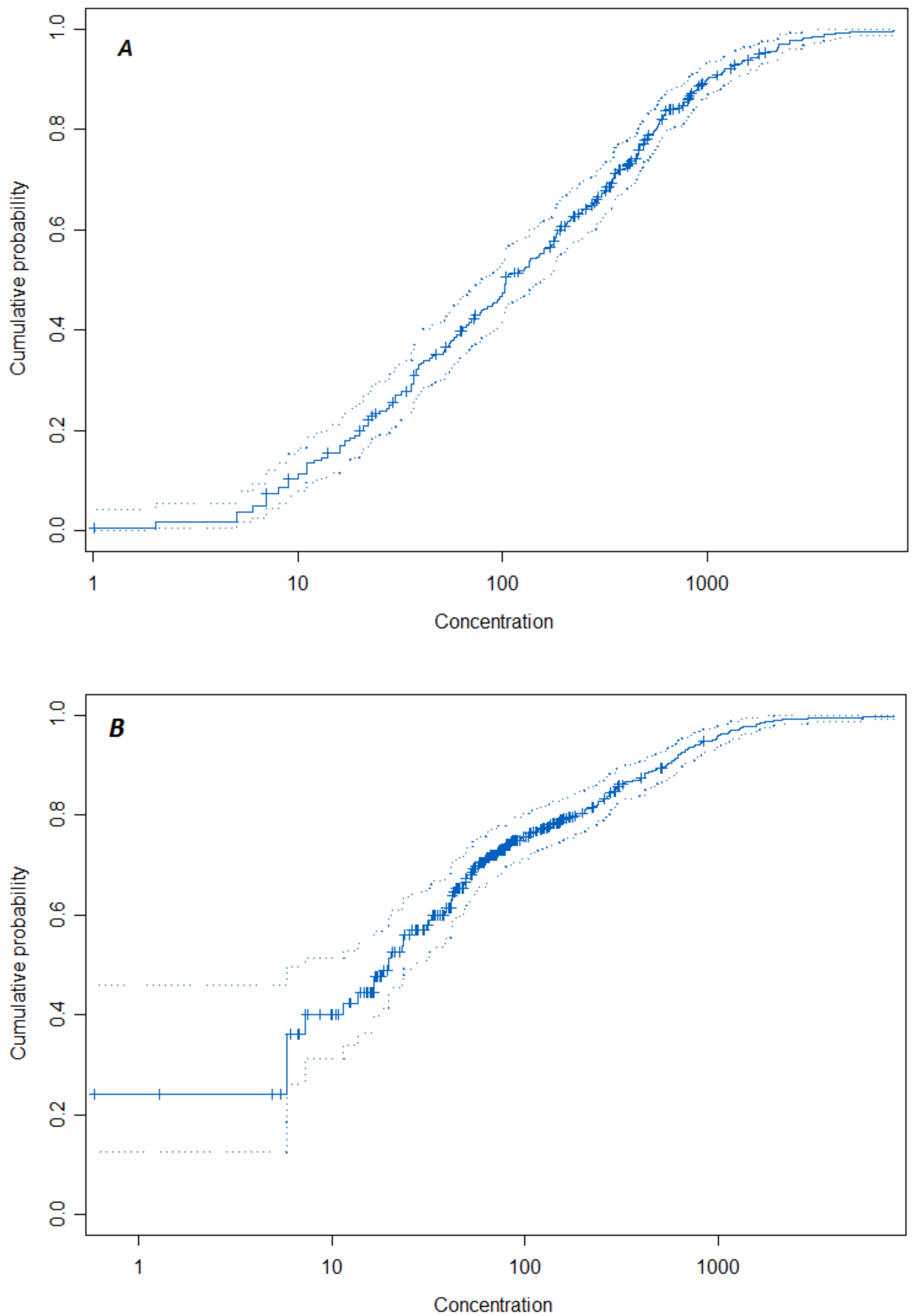

Figure 3. Scatterplot of sediment bulk density and Line 6B oil concentration for top 0.5 ft of cores collocated and sampled for both determinations, summer 2012, Kalamazoo River.