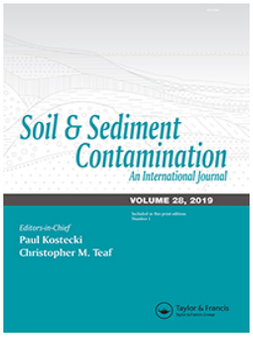

Soil and Sediment Contamination: An International Journal

ISSN: 1532-0383 (Print) 1549-7887 (Online) Journal homepage: https://www.tandfonline.com/loi/bssc20

Incremental sampling methodology for petroleum hydrocarbon contaminated soils: volume estimates and remediation strategies

Kathlyne Hyde, Wai Ma, Terry Obal, Kris Bradshaw, Trevor Carlson, Steven Mamet & Steven D. Siciliano

To cite this article: Kathlyne Hyde, Wai Ma, Terry Obal, Kris Bradshaw, Trevor Carlson, Steven Mamet & Steven D. Siciliano (2019) Incremental sampling methodology for petroleum hydrocarbon contaminated soils: volume estimates and remediation strategies, Soil and Sediment Contamination: An International Journal, 28:1, 51-64, DOI: 10.1080/15320383.2018.1529736

To link to this article: https://doi.org/10.1080/15320383.2018.1529736

SOIL AND SEDIMENT CONTAMINATION
2019, VOL. 28, NO. 1, 51–64
https://doi.org/10.1080/15320383.2018.1529736

Taylor & Francis
Taylor & Francis Group

Incremental sampling methodology for petroleum hydrocarbon contaminated soils: volume estimates and remediation strategies

Kathlyne Hyde[1], Wai Ma[1], Terry Obal[2], Kris Bradshaw[3], Trevor Carlson[3], Steven Mamet[1], and Steven D. Siciliano[1]

[1]Department of Soil Science, University of Saskatchewan, Saskatoon, SK; [2]Department of Scientific Services and Development, Maxxam Analytics, Mississauga, ON; [3]Department of Sustainability, Federated Co-operatives Limited, Saskatoon, SK

ABSTRACT

Current environmental assessments for petroleum hydrocarbon (PHC) contaminated sites are dependent on discrete soil sampling to estimate the degree and extent of contamination, leading to unreliable and non-reproducible results. Incremental sampling methodology (ISM) involves collecting and combining samples within a targeted area and holds promise for being a cost-effective, representative, and reproducible sampling strategy for contaminated site characterization. We hypothesized that traditional Phase II Environmental Site Assessments (ESA) discrete and ISM sampling protocols were not mutually exclusive, and the two approaches can be used to formulate a responsible land management strategy. Results gathered through ISM were compared to those from Phase II ESA for two PHC contaminated sites in Canada. Both methods indicated the sites were impacted with PHC beyond Saskatchewan Tier I guidance, however, the delineation of the PHC plume differed by as much as 75% for the heavier hydrocarbons. The Phase II ESA methods had higher incidences of false positive results and an overestimation of contamination at depth. A laboratory experiment confirmed that ISM does not "dilute" the samples as to cause underestimation, whereby the hydrocarbon concentrations for a single combined sample was equivalent to the mean of 30 discrete samples. Based on our results, sites should undergo risk assessment based on the estimates of the Phase II ESA results using vapor phase logs to estimate contaminant extent. If exposure pathways cannot be eliminated through the risk assessment process, remediation planning based on the ISM results is justified given the demonstrated cost-effectiveness, representativeness, and reproducibility.

KEYWORDS

Soil sampling; Risk assessment; Environmental site assessment; Direct push drilling; Conceptual modeling

Introduction

The collection of a "representative sample" is the hallmark of most scientific studies. Inferences about a population are made based on the characteristics of representative samples that have each element in the same percent composition as in the population

CONTACT Kathlyne Hyde kathlyne.hyde@usask.ca

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/bssc.
Supplementary data for this article can be accessed here.

(Hadley and Bruce, 2014). In the environmental industry, decisions regarding the degree and extent of site contamination are often dependent on the characteristics determined from a handful of samples that are considered representative. However, the small number of samples used in traditional assessments cannot adequately represent the entire site, as the interpretation is limited to the specific amount of soil collected (Brewer *et al.*, 2016a). The costs of overestimating the degree and extent of contamination are measured in time and money required for clean-up and remediation. In contrast, the costs of underestimating the extent of contamination poses potential negative impacts to human and ecological health.

Soil heterogeneity and spatial variability of contaminants have long confounded traditional discrete (or grab) sampling for petroleum hydrocarbon (PHC) contaminated sites (Jenkins *et al.*, 2005). This approach can lead to an inaccurate estimate of the degree and extent of contamination leading to conservative cleanup goals. Discrete sampling is scientifically unrepresentative and nonreproducible given the high heterogeneity in soil composition and contaminant distribution that might exist over short distances (Brewer *et al.*, 2016a; Brewer *et al.*, 2016b; Hadley and Petrisor, 2013). This is especially problematic in western Canada where the highly heterogeneous soils are formed from glacial till and the freeze-thaw conditions make estimating the average concentration of concern for a site difficult and unreliable. However, discrete sampling can be useful for risk identification and management purposes under these circumstances when a PHC contaminated site is highly sampled, analyzed, and delineated or if historical reports have sufficient information to create a more meaningful discrete sampling plan.

Incremental sampling methodology (ISM) is a systematic approach used to obtain more representative data for a site that ultimately leads to improved site management decisions. In conjunction to sample handling processes and subsampling, representative data is collected with detailed planning and site research prior to field sampling to ensure that representativeness requirements are met (Hadley and Bruce, 2014). The sampling design and sample processing methods used in incremental sampling reduces the fundamental error associated with the heterogeneous nature of the soil by increasing the mass of soil used for analysis (Brewer *et al.*, 2016a; Brewer *et al.*, 2016b; Gy, 1998), and reducing analytical costs. ISM is essentially a systematic approach to construct a representative sample of the decision unit (DU) by combining at least 30 increments to create one sample for the DU that provides the average PHC concentrations of the core section sampled.

Investigation areas (IAs) and DUs are set volumes of soil on a site that are incorporated into conceptual site models with a designated purpose for risk assessment or remediation (ITRC (Interstate Technology Regulatory Council), 2012). Typically, PHC sites in Canada contain over 15,000 m^3 of soil to characterize and 3–4 lateral IAs should be implemented. IAs and Single Borehole Decision Units can be used to create targeted bioattenuation zones that allow for targeted remediation efforts and defined areas of concern within the total amount of soil on site (SWRCB, 2012). Additionally, using IAs and DUs eliminates the biased sample collection from "hot spots". Hot spots can be thought of as outlier data portraying PHC concentrations that are much higher than the average concentrations (Brewer *et al.*, 2016a). For example, a 5 g soil sample may hold free product that is not present in the remainder of the core and this would not be entirely useful information when assessing risk. ISM enhances conceptual site modeling by providing structure to remediation sampling and planning and provides mean concentration estimates for large volumes of soil.

A sample collected through ISM must have: (1) a clear environmental objective stated before the initiation of the investigation; (2) adequate spatial coverage of the targeted DU and sampling density considered during the planning stage; and (3) the laboratory aliquots prepared for analysis designed to be representative of the field sample (ITRC, 2012). For PHC sites with limitations of deep subsurface contamination and high costs for drilling boreholes, it is recommended that each borehole serves as a Single Borehole DU with DU layers that will be incrementally sampled (HIDOH, 2016). This approach involves the designation of specific layers or "DU intervals" within a single borehole for the collection of an ISM sample, rather than testing of a small, "discrete" mass of soil from a single point. Although still prone to error due to small-scale heterogeneity, the collection of a sample across a targeted interval rather than a single point provides improved data to assess the presence or absence and the relative magnitude of petroleum contamination at a single borehole location (Brewer, personal communication, 2018). Given site history, IAs (source, plume, and clean delineation) can be implemented to encompass the Single Borehole DUs to further target areas of concern for remedial action plans. Using ISM for contaminated site management can avoid remobilization when traditional discrete sampling does not provide enough information to formulate appropriate decisions.

Most practitioners and regulators have concerns with ISM for environmental assessment regarding the delineation of areas with high concentrations of contamination (hot spots) and regulatory acceptance (Hadley and Petrisor, 2013). While traditional discrete sampling can provide a conservative estimate for risk assessments, there is still uncertainty of whether it is truly overestimating and not underestimating contaminant concentrations. ISM does not provide hot spot spatial location or magnitude of contamination in a DU. However, this hot spot information is not entirely useful as soil heterogeneity and contaminant concentrations will vary even within a 5 cm-wide boring (Brewer, personal communication, 2017). Progressively smaller DU layers can be established while in the field based on odor, staining, and volatile readings to encapsulate the hot spots (ITRC (Interstate Technology Regulatory Council), 2012; HIDOH, 2016). Even so, there are significant time and cost constraints in chasing hot spots with either approach.

Given the strengths and weaknesses of both traditional discrete sampling and ISM, it is prudent to investigate the application of both sampling strategies for estimations of contaminated soil volume and develop management strategies incorporating both approaches. The objectives of this study were to (1) estimate lateral and vertical extent of soil with PHC concentrations above allowable limits for two contaminated sites in Saskatchewan, Canada; (2) quantify and determine the cause for differences in contaminated soil volumes estimated by the two methods; (3) evaluate the precision of traditional discrete sampling methods compared to ISM; and (4) determine how to use the information gathered using both methods to manage the two contaminated sites. In comparison to ISM, we hypothesized that discrete sampling methods will frequently over – or underestimate the mean concentrations of PHC's and the contaminated soil volume and extent.

Materials and methods

Initial sampling activities

Two legacy gasoline and diesel bulk transfer stations located in Saskatoon and Raymore (200 km southeast of Saskatoon), Saskatchewan, Canada, with known spill and leak history were chosen for sampling (Figure 1, co-locator map in Supplemental 1). The Interstate Technology Regulatory Council (ITRC) ISM document was used as a guideline for experimental design and sampling (ITRC (Interstate Technology Regulatory Council), 2012). The site areas were conceptually divided into four investigation areas (IAs) encompassing the source, plume, plume delineation, and clean areas. Each IA contained three Single Borehole DUs in unbiased locations, from which soil cores up to 7.5 m in depth were taken using direct push core drilling with a Geoprobe® 7822DT (Salina, USA). Single Borehole DUs had two co-located boreholes within 0.5 m of each other, one for the Phase II Environmental Site Assessment (ESA) based on discrete sampling methods and one for ISM analysis. Diagrams of conceptual site design and sampling are in Supplemental 1.

Traditional discrete samples were collected on site at depth increments of 0.5 m (for up to 6 or 7.5 m) from each initial borehole by the consultants contracted to perform a Phase II ESA. A single sample was submitted for analysis for each Single Borehole DU. The single or additional bias samples were taken based on visual contamination and odor, and if there were high volatile organic compound (VOC) readings on a photo-ionization detector (PID). Samples for volatiles analysis were collected with a 5 g Terra Core™ Sampler (En Novative Technologies, Dexter, USA) and placed into 40 mL VOC vial pre-charged with HPLC grade methanol. Approximately 200 g of soil was packed into 250 mL jars, from which a 5 g subsample was used for semi-volatiles analysis. From the co-located borehole, the 1.5 m acrylic tube segments were collected by the University of Saskatchewan and sealed with paraffin wax on site. Cores were stored at −20°C prior to further sub-sampling in a laboratory setting.

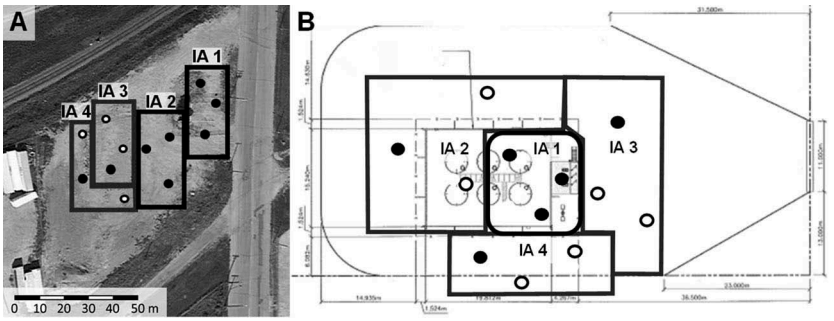

Figure 1. All circles denote Phase II Environmental Site Assessment (ESA) borehole locations and filled circles denote where a co-located Incremental Sampling Methodology (ISM) borehole was drilled. Investigation Areas (IAs) are represented by the rectangles; black rectangles represent plume area of interest and gray rectangles represent delineation areas. A) Aerial view of Saskatoon site and B) Schematic drawing of Raymore site

Stored cores: incremental and discrete sampling

We divided each Single Borehole DU into three DU layers: (1) surface zone at 0–1.5 m, (2) estimated contaminated zone at 1.5–4.5 m, and (3) depth delineation zone at 4.5–6.0 or 7.5 m depending on site. From each DU layer we collected: (1) 30 plug increments to combine for 1 ISM sample, (2) a wedge sample collecting surface soil from the entire length of the core, and (3) a discrete sample from a biased hot spot. The wedge sample was the least reproducible of the three sampling techniques due to core smearing and unreliable soil mass collection. Therefore, for the ISM to Phase II ESA comparisons, we only analyzed the data gathered using the plug sampling method (ISM samples) and the discrete samples taken from cores that ISM samples were collected from. The discrete samples collected were used to illustrate the biased sampling that occurs in the field and how the results drastically differ from the ISM samples collected from the same core.

We thawed the frozen soil cores until they were malleable and amenable to sample collection, but not so much for significant PHC volatile loss from the soil to air. To sample each core in a cost-effective manner, reusable soil corers were constructed out of copper tubing and modeled after the Terra Core™ Sampler (En Novative Technologies, Dexter, USA). To prepare the core for the ISM sample, the top layer was shaved off to expose fresh soil. First, using the plug method, we obtained 30 evenly spaced ~ 2 g cores for each DU layer and combined those into one 250 mL amber Boston bottle of High Performance Liquid Chromatography (HPLC) grade methanol (Fisher Scientific, Fair Lawn, USA) at a ratio of 2:1 methanol:soil for preservation of PHC volatiles – BTEX (benzene, toluene, ethyl-benzene, and *o*-,*m*-,*p*-xylenes) and CCME fraction F1 (C6-C10 hydrocarbons). Immediately following the first plug, a second plug was collected for CCME fractions F2-F4 (C10-C50 hydrocarbons) with a copper corer that held ~ 6.5 g. The second plug was combined in a 2-Dimensional (2-D) Japanese Slabcake (ITRC (Interstate Technology Regulatory Council), 2012) and then sub-sampled using a third copper corer that held ~ 6 g of soil for further homogenizing via chopping and mixing. The corers were washed and rinsed with hexanes, methanol, and acetone between each DU layer.

The collection of ~ 60 g of soil for VOC analysis and ~ 180 g of soil for semi – and non-VOC analysis did not meet the 300 g requirement for a sample of ISM standards (HIDOH, 2016; ITRC 2012). However, the ISM theory was used as 30 increments were collected, combined, and analyzed, in contrast to the three or less, 5 g sample collected and analyzed for Phase II ESA per DU layer.

To collect the discrete sample from the ISM sampled core, we identified PHC "hot-spots" in each core via observable staining or strong odor. We collected a discrete sample from each hotspot for PHC volatiles using a 10 g Terra Core™ sampler. Each sample was deposited into a 40 mL VOC vial, pre-charged with sufficient methanol to ensure a 2:1 methanol to soil ratio, provided by Maxxam Analytics (Mississauga, Canada). A 250 mL jar was packed with soil from the hotspot for CCME F2-F4 hydrocarbon analysis. The core was then inverted for the wedge method that involved using a large laboratory scoopula to shave the soil off along the axis of the core. A portion of the scoopula was deposited in methanol and the other portion was processed with a 2-D Japanese Slabcake for BTEX + F1 and F-F4 hydrocarbon analysis, respectively.

Hydrocarbon analysis

The samples combined in large amber Boston bottles of methanol were extracted at the University of Saskatchewan and then shipped to Maxxam Analytics (Mississauga, Canada) for analysis. We followed the CCME Tier 1 reference method (CCME, 2001) for soil hydrocarbon extraction and analysis for all samples. The soil sample weight for each DU was determined and an internal surrogate spike of 2000 ppm ethyl-benzene D-10 was used for quality control (QC). The samples were shaken vigorously for ~ 10 min and then allowed to settle on ice in a 4°C refrigerator for 24 h. A 2 mL aliquot was subsampled from each chilled sample and shipped to the Maxxam Analytics laboratory (Mississauga, Canada) for analysis of volatile (BTEX + F1). Maxxam Analytics also performed an extraction for F2-F4 hydrocarbons using a 5 g subsample from the jar of soil provided.

ISM vs. discrete methods

Three frozen 1.5 m cores from the Saskatoon site were selected and sampled via ISM and traditional discrete methods. An ISM sample consisting of 30 combined increments was collected from each core and preserved in methanol. An additional 30 discrete samples were collected and individually preserved from the same core for comparison. The same extraction process was used as above, and both the ISM samples and discrete samples were analyzed for F1 + BTEX using the CCME Tier 1 reference method (CCME (Canadian Council of Ministers of the Environment Inc.), 2001) with a Scion™ Gas Chromatograph equipped with a Flame Ionization Detector and a Mass Spectrometer (Bruker, Milton, Canada) at the University of Saskatchewan.

Data analysis

Regulatory guidance

Where available, PHC values were compared to the Saskatchewan Ministry of the Environment Tier 1 guidance values for PHC's in coarse grain soils under residential land use (SME, 2009). The use of residential guidance is for the protection of adjacent residential properties in case chemicals of potential concern migrate off-site.

Plume mapping

Kriging interpolation among contaminated boreholes for the Saskatoon site (n = 11) was used to produce subsurface benzene, F1, and F2 concentration maps using ArcMap as an example of an initial estimate for contaminant extent and volume (ESRI Inc., Redlands, USA). The Phase II ESA traditional discrete data was compared to the ISM sample data from the co-located boreholes. Phase II ESA benzene and F1 plume map concentrations were determined using the discrete concentrations from the Phase II ESA report, and then estimating the vertical extent based on the vapor log reading. The Phase II ESA F2 map concentrations were determined by using the discrete value only at the specific depth it was recorded. The volume of the plume was estimated at depths of 1, 3, and 5 m by multiplying the number of pixels by the pixel area (0.15 m^2) and the thickness of each layer (1.5, 3.0, and 1.5 m). Total plume volume was calculated as the volume summation of the three layers.

False positives and false negatives

The false positive and negative data analysis used only contaminated and co-located boreholes from both sites (n = 11) for the analysis. The false negative rate percentages of the traditional Phase II ESA sampling methods in comparison to the ISM samples and discrete samples taken from the co-located ISM borehole were determined by counting the instances in which the Phase II ESA values were below CCME guidelines and the ISM methods were above guidelines. Likewise, the false positive rate percentages were obtained by counting the instances in which the Phase II ESA value was above guidelines in comparison to the ISM and discrete sample taken from the incrementally sampled core. These definitions explicitly assume that the ISM represents the "true" method. Assuming otherwise would merely reverse false-negative and positive values. Thus, the interpretation would remain the same.

95% Upper confidence limit calculations

The 95% upper confidence limits (95UCL) were calculated using the ITRC UCL Calculator (ITRC (Interstate Technology Regulatory Council), 2012). A Chebychev 95UCL was used due to a large variability in the data noted from previous discrete sampling within an IA in the Phase I ESA mobilization. To assess risk, a 95UCL value was calculated for each DU layer in each Phase II borehole and co-located ISM borehole. Additionally, a 95UCL was calculated for each IA and each DU layer for each sampling method for more detailed data for target remediation plans (Supplemental Material 2). Data used for the calculation with Phase II borings was limited because samples were only taken where contamination was assumed present. The 95UCL calculation method, number of borings used, and soil sample mass for each method is included in Supplemental 2 and 3.

Results

Plume mapping

The kriging interpolation was used as a gross estimation of contaminant extent on site and to compare the traditional discrete sampling methods with ISM sample data collected. Except for benzene, the volume of soil affected by PHC as predicted by ISM at the Saskatoon site was less than that predicted by the traditional Phase II ESA sampling method (Figure 2). The ISM method predicted the benzene-, F1-, and F2-affected soil volume was approximately 10,200, 6,450, and 2,330 m^3, respectively. The Phase II ESA predicted benzene-, F1-, and F2-affected soil volumes were approximately 9,340, 6,930, and 4,090 m^3, respectively. The percent difference between the ISM and Phase II ESA estimates of benzene-, F1-, and F2-affected soil volumes were −8.5, 7.5, and 75.6%, respectively.

False positives and false negatives

The discrete sample taken from the ISM co-located borehole detected PHC contamination when the Phase II ESA method failed to detect contamination, i.e., the Phase II ESA method produced false-negatives (Type II error). The rate of Type II error was 18% for the Phase II ESA when compared to the discrete sample from the ISM core for both benzene

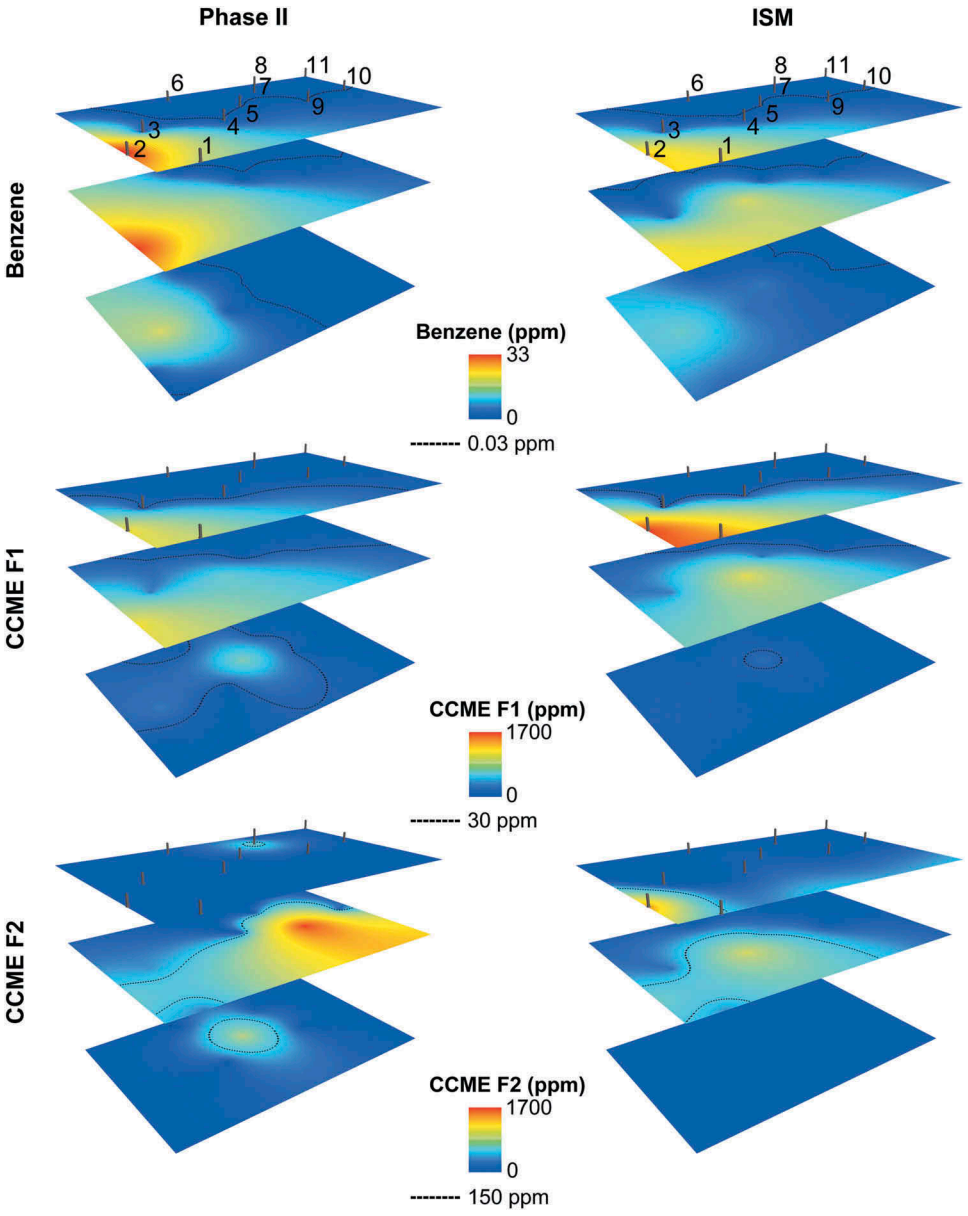

Figure 2. Subsurface plume maps of the Saskatoon site comparing the Phase II Environmental Site Assessment and ISM petroleum hydrocarbon contaminant plume extent and the concentrations for benzene, F1, and F2. Borehole locations are indicated in the top layers by numbered posts (upper panel). Canadian Council Ministry of Environment guideline thresholds for benzene (0.03 ppm), F1 (30 ppm), and F2 (150 ppm) are indicated for each plume with a dashed line

and F2 contamination (Figure 3a). The ISM data agreed with the Phase II ESA approach for both benzene and F2 contamination in soil.

When compared to the ISM samples and discrete samples from the ISM core, the Phase II ESA method detected incidences of contamination when contamination was not present

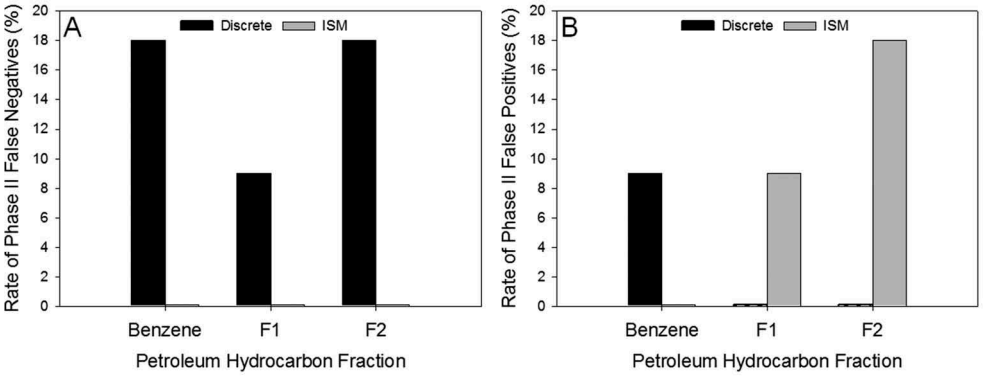

Figure 3. Rates of false negatives (Type II error) and false positives (Type I error) for the Phase II Environmental Site Assessment hydrocarbon results (benzene, F1, and F2) when compared to the discrete and incremental sampling method

(Type I error, i.e., false positive). The Phase II ESA method produced a 9% Type I error rate for benzene when compared to the discrete sampling method (Figure 3b). The ISM sample results indicated the Phase II ESA had Type I error rates of 9, 9, and 18% for benzene, F1, and F2, respectively.

95% upper confidence limit calculations

Assessing the 95UCL of the sites based on DU layer (across the site and not considering IAs), discrete sampling and Phase II ESA estimations of benzene, F1 (excluding BTEX), F2, and F3 concentrations across the site disagreed with ISM estimations (Table 1). Where no data is available with the discrete sampling methods, we assume there is no contamination present (ND in Table 1). Using the approach of analyzing the Chebychev 95UCL per IA and subsequent DU layers, for F1 (excluding BTEX), F2, and F3, at least one DU layer at each site had concentrations greater than the respective guidance values of 30, 150, and 300 mg/kg soil (ppm), respectively. All ISM 95UCL values for F4 were below the guidance value of 2800 mg/kg soil (Supplemental 2).

ISM vs. discrete methods

PHC concentrations measured from the ISM technique were comparable to the mean PHC concentration of 30 discrete samples (Table 2). Apart from where PHC concentrations were low (e.g., Core \$1), PHC concentrations of individual ISM samples were < 10% different than the mean concentrations of 30 discrete samples and the difference was often less than 5%.

Discussion

In applying ISM to two PHC contaminated sites for comparison to traditional discrete sampling, we have proven that ISM defies preconceived notions of over-diluting soil PHC concentrations. Here we have demonstrated that ISM does not: (1) underestimate the plume extent, or (2) underestimate the magnitude of contamination. ISM holds greater

Table 1. Chebychev 95UCL (95% Upper Confidence Limit) data (mg contaminant/kg soil) for each Decision Unit Layer (DUL) at each site

			Contaminant Concentration (mg kg^{-1})		
Site	DUL	Contaminant	ISM	Discrete	Phase II
Saskatoon	1[¶]	Benzene[†]	16	6	ND[‡]
	2	Benzene	15	59	31
	3	Benzene	4	8	ND
	1	F1[†]	1417	542	ND
	2	F1	490	2858	757
	3	F1	44	166	ND
	1	F2[†]	753	1485	ND
	2	F2	425	4591	290
	3	F2	35[$]	370	ND
Raymore	1	Benzene	17	8	0.21
	2	Benzene	5	39	32
	3	Benzene	ND	0.05	ND
	1	F1	2501	926	192
	2	F1	366	2572	2858
	3	F1	ND	ND	ND
	1	F2	2685	2852	6559
	2	F2	179	1393	1554
	3	F2	ND	ND	ND

[†]Values for Saskatchewan Tier 1 guidance values for petroleum hydrocarbons in coarse grain soils under residential land use (Benzene = 0.03 mg kg^{-1}, CCME F1 = 30 mg kg^{-1}, CCME F2 = 150 mg kg^{-1}). The F1 values are F1-BTEX.
[‡]ND = value below instrument detection limit or no data available.
[¶]Amount of soil used for each DU layer for each method in Supplemental 3.
[$]Values below Saskatchewan Tier 1 guidance values.

Table 2. Comparison of 30 discrete samples to 1 ISM sample from three cores

		Contaminant Concentration (mg kg^{-1})	
Core	Method	Benzene[†]	F1[‡]
1	ISM	ND[§]	ND
	Discrete	ND	0.99[¶]
	SE[$]	ND	0.50[¶]
2	ISM	15	45
	Discrete	13	51
	SE	1.2[¶]	3.1[¶]
3	ISM	35	140
	Discrete	38	140
	SE	1.5[¶]	1.3[¶]

[†]Saskatchewan Tier 1 Guideline = 0.03 mg/kg.
[‡]Saskatchewan Tier 1 Guideline = 30 mg/kg.
[§]ND = value below instrument detection limit.
[¶]Values are below Saskatchewan Tier 1 guidance values for petroleum hydrocarbons in coarse grain soils under residential land use.
[$] SE = Standard Error for discrete.

statistical power by overcoming the fundamental and distributional error associated with spatial soil heterogeneity and contaminant distribution (Gy, 1998.; ITRC (Interstate Technology Regulatory Council), 2012; Brewer *et al.*, 2016a; Brewer *et al.*, 2016b). Therefore, inferences made about ISM as the "true" method remain correct by the theory that collecting more soil mass results in a more representative sample.

In comparison with ISM, Phase II ESA discrete sample data underestimated the volume of benzene contamination and overestimated the CCME F1 and F2 volumes for the Saskatoon site (Figure 2). Between the methods, the PHC volume estimates for benzene and F1 compounds were no more than 9% different (Supplemental 4).

However, the volume extent estimations for the Phase II ESA used a single discrete sample and on-site PID readings, whereas the ISM concentrations were analytically measured soil samples. Constructing a plume map using the Phase II ESA data and PID readings may only be suitable for the more volatile hydrocarbons as Phase II ESA overestimated the semi-volatiles plume volume by 75% (e.g., CCME F2, Figure 2 and Supplemental 5). If the objective is to properly delineate the plume, ISM methods are recommended as more reliable gross spatial data will be gathered for the volatile, semi-volatile, and non-volatile PHC's and using PID readings is not necessary for all core sections sampled.

The utility and cost-effectiveness of incremental sampling is best demonstrated where incremental samples identify a greater extent of contamination than discrete samples (Hadley and Petrisor, 2013). The analysis of the 95UCL data for DU layers across the site (not considering Investigation Areas) demonstrated the strength in using ISM by reporting values above CCME guidelines (0.03 mg/kg) where Phase II ESA did not or lacked data. The implication of not finding all the contamination, or not detecting all the specific chemical contaminants, may appear more significant to designing remediation projects. Incremental sampling may be useful when conducted in conjunction with traditional screen techniques to formulate conceptual plans for Single Borehole DUs and DU layers based on the Phase II ESA data that indicates the presence of contamination.

The smaller contaminated soil volumes estimated by the ISM samples was caused by the lower incidences of false positives (as compared to the Phase II ESA; Figure 3b). Improved rates of Type I and Type II errors for ISM over discrete sampling is an important feature of incremental sampling (ITRC (Interstate Technology Regulatory Council), 2012). Detecting contamination when none is present (Type I error) will increase the volume of soil perceived to be "dirty" or even above guidance. Therefore, traditional Phase II ESA methods present a risk of increasing remediation costs. Hence, for remediation purposes, incremental sampling is a better option compared to discrete sampling. Discrete sampling is likely still suitable for site management and risk assessment where worst-case scenario estimates (higher incidences of Type I error) may be perceived as beneficial.

The vertical delineation based on the three DU layers for the two sites highlights the importance of vertical spatial variability on estimated plume size for risk assessment and management or site remediation. The strength in ISM is that there is information for each depth increment (Hewitt *et al.*, 2008), whereas the Phase II ESA method relies on a singular concentration to represent site conditions. It is possible that the Phase II ESA process could be conducted with vertical delineation; however, that is not the current Phase II ESA guidance framework (Canadian Safety Council, 2013). The ISM sampling reported contamination varied amongst the DU layers, whereas the Phase II ESA data were lacking and vertical contamination extent was a matter of personal judgment. Again, the discrepancy in vertical delineation between the two methods underlies how each sampling strategy might be used depending on the objective for the sites.

The bias collection of discrete samples from pockets of obvious heavy contamination in a Phase II ESA investigation will cause the risk posed by the contaminated mass of soil to be overestimated and could lead to unnecessary remedial actions. In contrast, risk should be assessed based on an estimate of the mean or "true" concentration of contaminant for a

targeted exposure area and/or volume of soil rather than small, individual points within the targeted soil (ITRC (Interstate Technology Regulatory Council), 2012; HIDOH, 2016). For example, soil screening levels used to assess direct exposure risk are intended to address long-term and random exposure to contaminants in soil throughout an exposure area over a period of many years. This requires an estimate of the mean contaminant concentration for an area often the size of the site itself. Delineation described through ISM allows practitioners to more appropriately and efficiently assess risk and plan remedial objectives. The main objective is to identify "hot areas" for remediation, not "hot spots".

The data reliability and representativeness of ISM was confirmed through the laboratory experiment. The experiment demonstrated that ISM samples produce similar results to the mean of 30 discrete samples. This analytically confirms that a constructed incremental sample is equivalent to 30 discrete samples. As this was observed for three different cores, ISM is a representative sampling method to determine mean contaminant concentrations from a single combined sample (Brewer *et al.*, 2016a; Brewer *et al.*, 2016b; Jenkins *et al.*, 2005; Ramsey and Hewitt, 2005). Additionally, the standard error reported for the discrete method demonstrated the inherent soil and contamination heterogeneity that exists even within a soil core. Thus, no individual discrete sample could robustly represent the reported mean contamination within a core. Fortunately, ISM is a cost-effective and representative way to estimate the true concentration of contaminants of potential concern for logical targeted areas and volumes of soil.

An integrated strategy for the management and possible remediation of the Saskatoon and Raymore sites can be developed from the ISM and Phase II ESA results. The two sampling approaches confirmed that both sites are impacted by PHC to levels above current Tier 1 guidance values for BTEX, F1, and F2 (potentially F3 from Phase II ESA only). The ISM conceptual design and sampling/processing procedure provided highly informative and reliable results. Initial estimates from Phase II ESA hot spot discrete sampling should be reassessed through the collection of ISM data for more reliability and accuracy. Under ideal circumstances, a full ISM investigation would be carried out with multiple designated IAs and 30 or more borings installed in each IA for the collection of proper ISM samples (ITRC (Interstate Technology Regulatory Council), 2012; HIDOH, 2016). This is not practical or even necessary for many, if not most petroleum-release sites, however, we typically have well-known history of the site in question and the initial information that can be gained from Single Borehole DU investigation methods. This information can be used to carry out initial remedial actions. A full ISM investigation can then be used to confirm the adequacy of remedial actions. For example, through the collection of 30 or more increment samples from the sidewalls and floor of an excavation or from targeted volumes of soil treated using in situ methods. If exposure pathways cannot be eliminated through the risk assessment process (including the evaluation of risk management strategies), remediation planning based on the ISM results is justified given the demonstrated cost-effectiveness and reproducibility.

We recommend that site practitioners use ISM as a remediation planning tool. The current Phase II ESA methods are effective in identifying areas of potential concern, but cannot provide robust estimates of contaminant mass. Following a traditional Phase II ESA with an ISM directed sampling approach will provide statistically robust estimates of contaminant mass and exposure for risk assessment. There are two possible approaches to accomplish this objective. The first, and most straightforward, is to remobilize sampling

crews to impacted sites to implement an ISM approach capitalizing on the information obtained through the initial Phase II ESA to better define and articulate the Investigation Areas, Single Borehole DUs, and DU layers. The second approach is to combine traditional Phase II ESA and ISM approaches for the same mobilization. In this approach, Phase II assessments would be used to identify Investigation Areas, Single Borehole DU's, and DU layers that highlight areas of potential concern – e.g., underground storage tanks and known areas of release. Drill crews would be instructed to take cores that would be sealed on site and stored frozen for later analysis pending Phase II ESA results. Decisions can be made on site if the DU layer depths should be alternatively divided based on obvious cues such as odor, staining, and high PID readings. Such an approach would save on remobilization, and for sites with relatively shallow impact (e.g., < 6 m), the additional drilling costs would be negligible compared to a remobilization. In our experience, at a simple site, one would need three Investigation Areas: source, plume, and clean delineation and three DU layers: vadose zone, capillary fringe, and saturated soil zones. Thus, in total there would need to be at least nine boreholes (3 Investigation Areas × 3 Single Borehole DUs) consisting of 27 units (each borehole × 3 depth increments) taken later for analysis. Furthermore, ISM sampling in a laboratory is relative easy with promise for optimizing the method so that total investigation costs are not greatly increased. For complex sites with large vertical or lateral footprints, we recommend a separate remobilization to ensure the DUs are better defined, allowing ISM to provide improved clarity for remediation planning.

Acknowledgements

This work was supported by a NSERC Industrial Research Chair awarded to SDS. Phase II ESA was performed by Stantec, and the analytical data for both the Phase II ESA and ISM work were provided by Maxxam Analytics. The authors thank an ISM expert, Roger Brewer, for his courteous and valuable input in refining this manuscript. The authors thank Richard Nhan and the SDS laboratory for their efforts in the sampling process.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada [IRCPJ 478012-13].

References

Brewer, R., Peard, J. and Heskett, M. 2016a. A critical review of discrete soil sample data reliability: Part 1–Field study results. *Soil Sediment Cont*. 26, 1–22.

Brewer, R., Peard, J. and Heskett, M. 2016b. A critical review of discrete soil sample data reliability: Part 2 – Implications. *Soil Sediment Cont*. 26, 23–44.

Canadian Safety Council. 2013. Phase II environmental site assessment. CAN/CSA-z769-00 (R2013).

CCME (Canadian Council of Ministers of the Environment Inc.). 2001. Reference method for the Canada-wide standard for petroleum hydrocarbons in soil – Tier 1 method. ISBN 1-896997-01-5, Publication No. 1310.Winnipeg, MB.

Gy, P. 1998. *Sampling for Analytical Purposes. A.G. Royle, Trans*. West Sussex, U.K.:Wiley.

Hadley, P. W. and Bruce, M. L. 2014. On representativeness. *Environ. Forensics*. 15(1), 1–3.

Hadley, P. W. and Petrisor, I. G. 2013. Incremental sampling: Challenges and opportunities for environmental forensics. *Environ. Forensics*. 14, 109–120.

Hewitt, A. D., Ramsey, C. and Bigl, S. 2008. Multi-increment TCE vadose-zone investigation. *Remediat J*. 19, 125–140.

HIDOH, 2016. *Technical Guidance Manual*: Hawai'i Department of Health, Office of Hazard Evaluation and Emergency Response. http://www.hawaiidoh.org/.

ITRC (Interstate Technology Regulatory Council). 2012. *Incremental Sampling Methodology. ISM-1*. Washington, D.C: Interstate Technology & Regulatory Council, Incremental Sampling Methodology Team. www.itrcweb.org

Jenkins, T. F., Hewitt, A. D., Walsh, M. E., Ranney, T. A., Ramsey, C. A., Grant, C. L. and Bjella, K. L. 2005. Representative sampling for energetic compounds at military training ranges. *Environ. Forensics*. 6, 45–55.

Ramsey, C. A. and Hewitt, A. D. 2005. A methodology for assessing sample representativeness. *Environ. Forensics*. 6, 71–75.

SME (Saskatchewan Ministry of Environment). 2009. Risk-based corrective actions for petroleum hydrocarbon impacted sites, EPB $344. www.environment.gov.sk.ca

SWRCB (California State Water Resources Control Board). 2012. Resolution 2012-006: Low Threat Underground Storage Tank Case Closure Policy. https://www.waterboards.ca.gov