NRC Research Press

ARTICLE

# Quantifying hydraulic conductivity spatial variability for cement-based solidification/stabilization (S/S) remediation project: case study

Gordon A. Fenton, Craig B. Lake, and Rukhsana Liza



**Abstract:** This paper presents statistical analyses of hydraulic conductivity data collected from an existing cement-based solidification/stabilization (S/S) system. The goal is to characterize the spatial variability of hydraulic conductivity and to examine sampling recommendations for the quality control (QC) program of that system to achieve target decision error probabilities regarding the acceptance or rejection of the system with respect to hydraulic conductivity. Over 2000 QC hydraulic conductivity samples, taken over an area of 300 000 m², are used as a basis for these analyses. The hydraulic conductivity spatial variability is described by a marginal lognormal distribution with correlation function parameterized by directional correlation lengths, which are estimated by best fitting an exponentially decaying correlation model to sample correlation functions. The spatial variability associated with hydraulic conductivity of the studied S/S system is then utilized to assess sampling requirements for the QC program of that system. Considering the "worst case" correlation length and the hydraulic conductivity mean and variance, hypothesis test error probabilities are used to provide recommendations for conservative sampling requirements. It is believed that the analysis of this large construction project represents a unique opportunity to review the current practice of S/S field sampling requirements.

*Key words:* case study, sampling requirements, contaminated soil, remediation, hypothesis test errors.

**Résumé :** Cet article présente des analyses statistiques de données de conductivité hydraulique collectées à partir d'un système existant de solidification/stabilisation (S/S) à base de ciment. L'objectif est de caractériser la variabilité spatiale de la conductivité hydraulique et d'examiner les recommandations d'échantillonnage pour le programme de contrôle de la qualité (CQ) de ce système afin d'atteindre les probabilités d'erreur de décision cibles concernant l'acceptation ou le rejet du système en ce qui concerne la conductivité hydraulique. Plus de 2000 échantillons de conductivité hydraulique de CQ, prélevés sur une surface de 300 000 m², servent de base à ces analyses. La variabilité spatiale de la conductivité hydraulique est décrite par une distribution log-normale marginale avec une fonction de corrélation paramétrée par des longueurs de corrélation directionnelle, qui sont estimées en ajustant au mieux un modèle de corrélation à décroissance exponentielle aux fonctions de corrélation de l'échantillon. La variabilité spatiale associée à la conductivité hydraulique du système de S/S étudié est ensuite utilisée pour évaluer les besoins d'échantillonnage pour le programme de CQ de ce système. En considérant la longueur de corrélation du « pire cas » et la moyenne et la variance de la conductivité hydraulique, les probabilités d'erreur du test d'hypothèse sont utilisées pour fournir des recommandations pour des exigences d'échantillonnage prudentes. On pense que l'analyse de ce grand projet de construction représente une occasion unique de revoir la pratique actuelle des exigences d'échantillonnage sur le terrain en matière de S/S. [Traduit par la Rédaction]

*Mots-clés :* étude de cas, exigences en matière d'échantillonnage, sol contaminé, assainissement, erreurs de test d'hypothèse.

## Introduction

Contaminated sites produced from previous industrial legacies often contain a mixture of contaminants with varying differing physical and chemical characteristics, which makes selection of a remediation technology difficult. Solidification/stabilization (S/S) is a common risk management technique used to manage environmental risk for these sites. Between the fiscal years of 1982–2011, S/S was the most commonly used ex situ treatment technology and the second most common in situ treatment technology for United States Superfund sites (USEPA 2013). This soil remediation treatment technology is accomplished by adding a binder(s) such as Portland cement, fly ash, lime, etc. to the contaminated soil. Portland cement is the most common binder used in practice (ITRC 2011); referred to as "cement-based S/S" in this paper. Reviews of the science behind cement-based S/S treatment technology have been provided by authors such as Conner (1990), Al-Tabbaa and Perera (2002), and Hills et al. (2015) and will not be repeated here. In essence, the goal of cement-based S/S treatment of contaminated soil is to both physically (i.e., solidify) and chemically (i.e., stabilize) modify the soil, soil moisture, and (or) contaminants such that the risk of contaminant migration from the resulting treated soil is below an acceptable risk level.

As discussed by the Interstate Technology and Regulatory Council (ITRC 2011), the performance objectives for a S/S remediation project usually involve strength, leachability, and hydraulic con-

Received 30 January 2020. Accepted 13 April 2020.

**G.A. Fenton.** Department of Engineering Mathematics, Dalhousie University, Halifax, NS, Canada.
**C.B. Lake* and R. Liza.** Department of Civil and Resource Engineering, Dalhousie University, Halifax, NS, Canada.
**Corresponding author:** Gordon A. Fenton (email: gordon.fenton@dal.ca).

*Craig B. Lake currently serves as an Editor-in-Chief; peer review and editorial decisions regarding this manuscript were handled by Daichao Sheng.





Published at www.nrcresearchpress.com/cgj on 17 April 2020.

ductivity. Strength objectives relate to the end-use goals for the site while leachability relates to minimizing the flux of contaminants from the treated soil. Hydraulic conductivity objectives, the subject of this paper, relate to reducing the porosity of the contaminated soil to limit mass transfer of the contaminants from the contaminated soil. The reduction of hydraulic conductivity after treatment also assists in reducing the flow of groundwater and (or) surface water through the treated contaminated soil by making the hydraulic conductivity of the treated soil lower than the surrounding soil (i.e., to reduce mixing of clean and contaminated water (ITRC 2011)). For monolithic cement-based treated soils, the resulting hydraulic conductivity after treatment plays a major role in limiting the contaminant migration from the site. Hence, it is imperative to ensure that hydraulic conductivity performance objectives are being achieved during construction. Meeting performance objectives for hydraulic conductivity in practice are achieved by developing sound quality control (QC) and quality assurance (QA) programs. The focus of this paper relates to QC programs for S/S projects.

During construction, QC programs are important to regulators and owners to ensure that the contractor is meeting the performance objectives stated at the beginning of the project. Without these QC programs, there would be doubts as to the success of the completed project. For cement-based S/S projects, hydraulic conductivity QC in the field is accomplished by taking samples of the treated soil and performing laboratory hydraulic conductivity testing (ITRC 2011). The treatment is deemed a success if the resulting conductivity is below the specified performance objective (i.e., a prescribed hydraulic conductivity must not be exceeded). In practice, this hydraulic conductivity objective could be based on past projects or contaminant migration assessments (ITRC 2011), but the typical specification for hydraulic conductivity ranges from $10^{-8}$ to $10^{-9}$ m/s, with $10^{-8}$ m/s being the most common. The number of samples required to be taken, and subsequently tested for hydraulic conductivity during the QC process, are established prior to the initiation of a project. The predominant method for deciding on the number of samples is based on the sample density method (USACE 2000), which requires a given number of samples per unit volume of treated soil. For in situ cement-based S/S construction purposes, the contaminated soil area to be treated is often subdivided into a number of grid cells (ITRC 2011), as is done for this case study, and hydraulic conductivity tests are performed within each cell and each cell average reported for the entire collection of construction cells within the grid. Under the current sample density approach, the risk of making an error in judgement based on the number of samples selected cannot be assessed, as the variability of the treated material does not play a role in the selection of the number of samples.

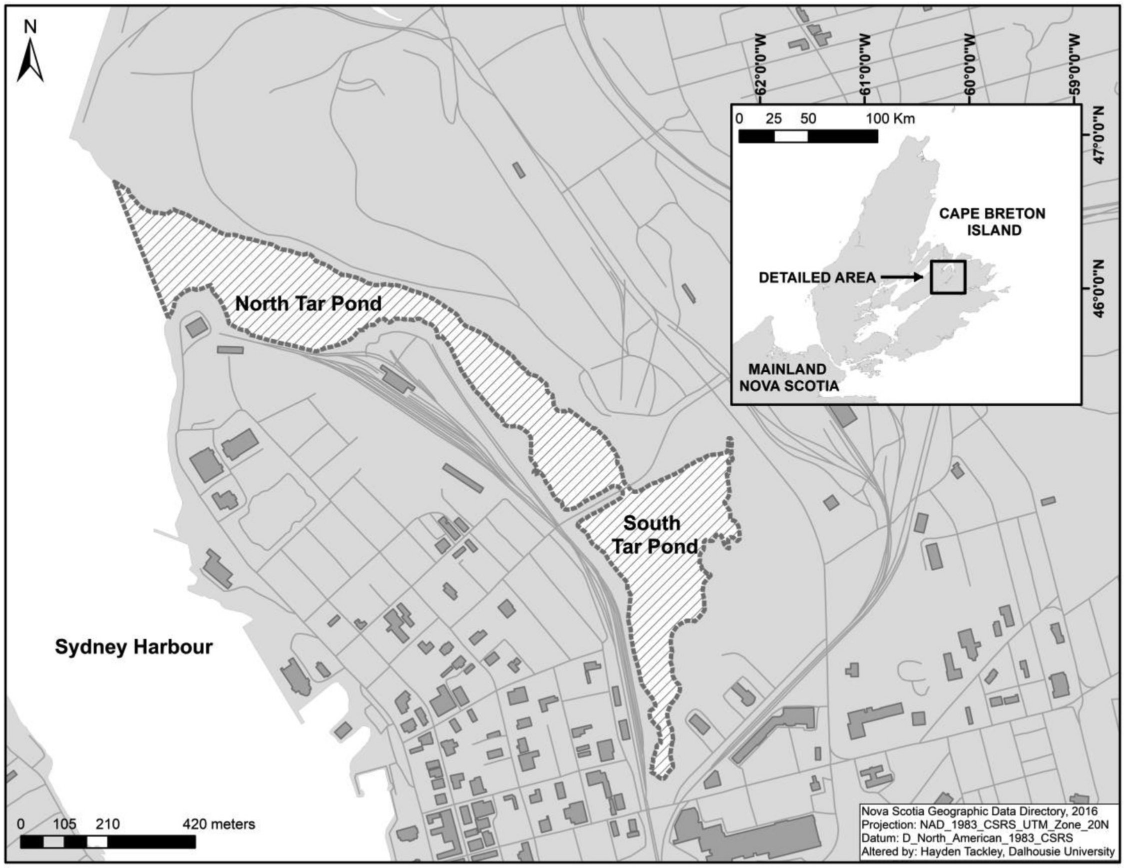
As discussed by Fenton et al. (2015), the question often raised with sample density approaches for QC programs is: "Are enough samples being taken to ensure performance?" To answer this question, statistical methods can be used to analyze the probability of excessive hydraulic flow through systems and (or) the risk associated with QC of such systems (e.g., Fenton et al. 2015; Liza et al. 2017). Given that hydraulic conductivity is spatially variable both for natural soil (Byers and Stephens 1983; Freeze and Cherry 1979) and compacted soil liners (Rogowski et al. 1985; Benson 1993), the hydraulic conductivity of the cement-based S/S material can be modeled as a random field, having a marginal "point" distribution and an associated spatial correlation function (Vanmarcke 1984). In this study, the "cell" values are local averages of "point" values within the "cell". Details about the relationship between point and local averages can be found in Fenton and Griffiths (2008), for example. However, these details are not particularly important, because, as will be seen, the correlation length (discussed in the following paragraphs) is significantly larger than the cell size so that the cell hydraulic conductivity is not significantly different than the point distribution. It is assumed in this paper that they are the same.

The correlation function is parameterized in this study by a correlation length, $\theta_{\ln K}$, which is a measure of the degree of spatial "persistence" between hydraulic conductivity values. Past statistical studies indicate that the distribution of hydraulic conductivity at a point is reasonably lognormal for both natural and compacted soils (Freeze 1975; Krapac et al. 1989; Johnson et al. 1990; Benson 1993). Benson (1991) found the correlation length to be 1–3 m for compacted soil liners. Unfortunately, even though the data exist for past cement-based S/S projects, no similar spatial statistical analyses can be found in the literature for cement-based S/S materials. Different methods are available in the literature to estimate the correlation length. In probably what is the most common method, the correlation length is estimated by best fitting the theoretical correlation model to the sample correlation function (Degroot and Baecher 1993; Fenton 1999; Jaksa et al. 1999; Fenton and Griffiths 2008; Wackernagel 2003; Zhang et al. 2008; Lloret-Cabot et al. 2013, Dasaka and Zhang 2012). Vanmarcke (1984) proposed a method based on the variance reduction function that was used by Wickremesinghe and Campanella (1993) and Lloret-Cabot et al. (2012 and 2013) to estimate correlation lengths in practice. Jaksa et al. (1993) matched theoretical to sample semivariograms to estimate the correlation length of clay using cone penetration test (CPT) data. Phoon and Fenton (2004) used the bootstrap approach to estimate sample correlation functions. Lloret-Cabot et al. (2014) describe how to match a theoretical correlation function to a sample correlation function.

As discussed previously, during QC programs of cement-based S/S systems, a set of field hydraulic conductivity data are obtained by collecting and testing the samples from a collection of construction cells. Hereafter a construction cell is referred to as a "cell". The hydraulic conductivity test dataset can be used to estimate the effective hydraulic conductivity, $k_{\mathrm{eff}}$, of a given cell. The effective hydraulic conductivity, $k_{\mathrm{eff}}$, is defined to be the single hydraulic conductivity value that gives the same total flow through the cell as does the actual spatially variable hydraulic conductivity field. The estimated effective hydraulic conductivity is used to make a decision about the acceptance or rejection of the cell. The cell is considered to be acceptable if the actual effective hydraulic conductivity is less than or equal to the prescribed hydraulic conductivity (limiting hydraulic conductivity specified by the regulatory agency), $k_{\mathrm{crit}}$. Otherwise the cell is deemed unacceptable and must be repaired (i.e., remixed with cement) or replaced. As the actual effective hydraulic conductivity is unknown, and only estimated by the sample, the decision about the acceptability of the cell may result in an error that is best investigated using concepts of hypothesis testing.

As discussed by Fenton et al. (2015) and Liza et al. (2017), the decision-making process with regards to hydraulic conductivity acceptability can be carried out by a simple hypothesis test. If the null hypothesis is that the cell is unacceptable with respect to hydraulic conductivity, two types of errors in the decision may result: a type I error where it is concluded that the cell has adequate hydraulic conductivity when it actually does not and a type II error where one fails to conclude that the cell has adequate hydraulic conductivity, when it actually does. From a safety point of view, it is more important to avoid the type I error. Details regarding the hypothesis test and the type I and II errors considered in this paper can be found in Fenton et al. (2015). These error probabilities are not considered in the current practice for QC sampling of cement-based S/S, which as mentioned previously, is typically based on the sample density method (USACE 2000). The sample density method requires a certain number of samples per unit volume and this number does not vary with the statistics of the sampled field. Given that the uncertainty of the S/S system depends on its variability, a fixed sampling density implies higher uncertainty in the acceptance decision for systems with higher

**Fig. 1.** Map of study site prior to cleanup. Map created in ARCGIS using data from the Nova Scotia Geographic Data Directory (2016) based on the North American Datum of 1983 Canadian Spatial Reference System (NAD83(CSRS) / UTM zone 20N).






variability. To avoid this, the number of samples should increase as the system variability increases — and, conversely, decrease for relatively uniform S/S systems.

Many cement-based S/S treatment projects acquire hydraulic conductivity data during QC programs, but due to the sensitive nature of these projects, few, if any, allow data mining to proceed to be used as a case history. The purpose of this paper is to present the results of a hydraulic conductivity QC program for the largest known cement-based S/S project in Canada to date. What is unique about this dataset is that it is spatially distributed, which allows the methods proposed by Fenton et al. (2015) and Liza et al. (2017) to be assessed and applied from a practical standpoint. As with Fenton et al. (2015) and Liza et al. (2017), the planar extent of this case study is much larger than the contaminated depth and hence the field can be assumed to be two-dimensional (2D) with the random field representing depth averages.

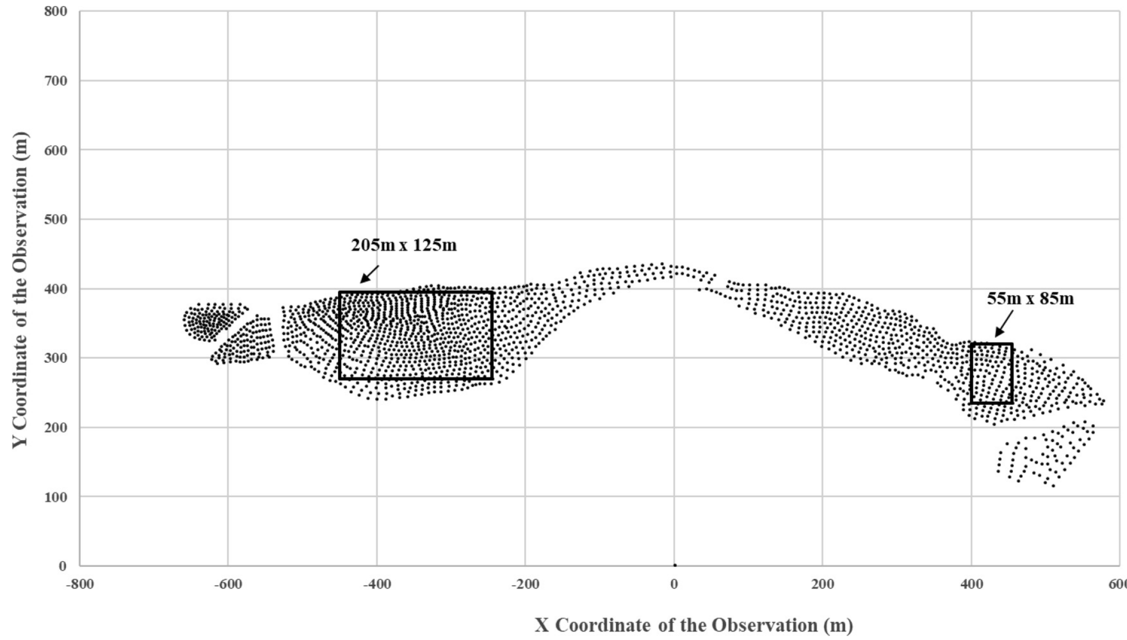
The case study also allows, for the first time, estimates of mean, variance, and correlation length to be reported in some detail for a large cement-based S/S remediation project. No previous study that attempts to find the distribution and correlation length describing the spatial variability of hydraulic conductivity for a cement-based S/S project can be found in the literature. The spatial variability associated with hydraulic conductivity of this cement-based S/S project is then utilized to assess sampling requirements for the QC program of this case study.

## Background: Sydney Tar Ponds project

The Sydney Tar Ponds, located in Sydney, Nova Scotia, Canada, represented one of Canada's largest industrially contaminated sites up until 2012 (STPA 2019). Since the 1800s, coal mining operations existed near Sydney and surrounding areas. In 1901, the first steel plant began production on the site and, in later years, coke production from coal was also established on the site. According to the Sydney Tar Ponds Agency (STPA 2019), it is estimated that Sydney produced half of the steel made in Canada in 1912. Over the subsequent years, coke and steel making activities declined, causing economic hardships for the local steelmaking industry. A century later, the steel making industry was gone, but the legacy of contamination from a century of coke production and steel making remained.

A map of the site prior to the start of cleanup (circa 2000) is shown in Fig. 1. The site, having total area of approximately 1 000 000 $m^2$, is situated in the urban centre of Sydney. Although this project presented many challenges, one of the more interesting challenges involved treatment and containment of approximately 550 000 $m^3$ of contaminated sediments over a surface area of 310 000 $m^2$ that had accumulated in what was referred to as the North and South Tar Ponds (see Fig. 1). The primary contaminants included heavy metals, volatile organic compounds, polycyclic aromatic hydrocarbons, and polychlorinated biphenyls. These contaminated sediments were identified (AMEC 2005) as appearing in layers ranging in thickness up to 4.9 m and covered by seawater. In situ cement-based S/S treatment was selected as the treatment and containment approach for these contaminated sediments. The treatment process consisted of dewatering the surface water over the sediments (see AECOM 2008) and using in situ mixing of a cement binder to achieve performance standards related to unconfined compressive strength, hydraulic conductivity, and contaminant leaching. Pertinent to this paper is the hydraulic conductivity specification ($k_{crit}$) used for the performance-based specification of the project, which was $1 \times 10^{-8}$ m/s.

**Fig. 2.** Sampling locations of hydraulic conductivity dataset shown as small black squares.






Cement binder addition was accomplished by subdividing the area of the site into over 2000 construction cells (see Fig. 2). Figure 2 shows the subdivision of the various construction cells. To keep the volume of each cell relatively constant (i.e., the contaminated depth was different for different cells), the cells have different surface areas. During the QC program, multiple samples were collected from each cell and tested for hydraulic conductivity following ASTM (2010) standard D5084. A given cell was approved individually if it was determined that the average of hydraulic conductivity measurements over that cell was at or below the $k_{crit}$ regulatory value (i.e., $1 \times 10^{-8}$ m/s). The centres of each of these cells were considered to be the sampling locations of the cell average hydraulic conductivity values in this study. As discussed by Fenton et al. (2015), over the entire site, 2086 hydraulic conductivity samples were taken and the estimate of the mean hydraulic conductivity, $\mu_{K'}$ (where $K'$ is the random hydraulic conductivity normalized by a regulatory value of $1 \times 10^{-8}$ m/s), was 0.47 with a coefficient of variation, $\nu_{K'}$, of 1.7. The uppercase $K$ is used in this paper to denote random quantities. The resulting fitted distribution is shown in Fig. 3. It should be noted that the 2086 sample results used for this paper represent the majority of samples taken for the project, but some samples, particularly from the south pond, were not made available to the study due to property privacy concerns.

## Statistical analyses

The statistical analyses performed in this study were based on the available 2086 hydraulic conductivity test values, $K$. It was assumed that the mean, variance, and marginal distribution of $K$ were stationary (spatially constant). In the following, the hydraulic conductivity values, $K$, were converted to a normalized form, $K'$, by dividing by the regulatory hydraulic conductivity, $1 \times 10^{-8}$ m/s (i.e., $K' = K/k_{crit}$). The hydraulic conductivity is assumed to be lognormally distributed, so that $K'$ is also lognormally distributed with parameters $\mu_{\ln K'} = \mu_{\ln K} - \ln k_{crit}$ and $\sigma_{\ln K'} = \sigma_{\ln K}$. To test this assumption, a frequency–density plot of the normalized hydraulic conductivity, $K'$, with a superimposed fitted lognormal probability density function, $f_{K'}(k')$, having parameters $\mu_{K'} = 0.47$ and $\nu_{K'} = 1.7$ ($\mu_{\ln K'} = -1.31$ and $\sigma_{\ln K'} = 1.02$) is shown in Fig. 3. Note that in

**Fig. 3.** Frequency–density plot of hydraulic conductivity, with fitted lognormal distribution.

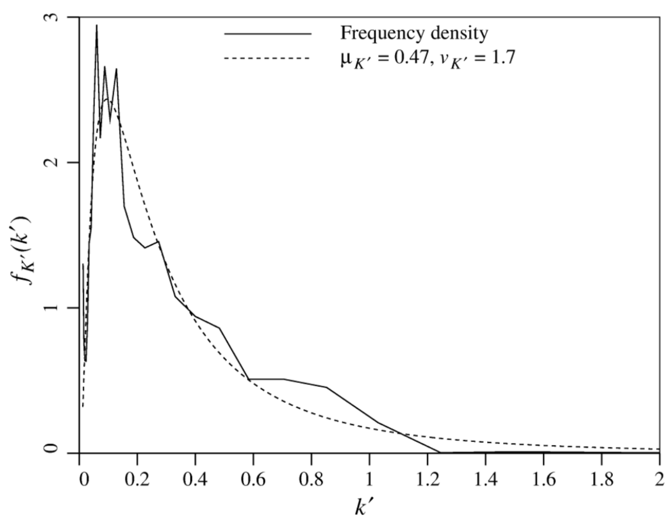



Fig. 3, $k'$ is the deterministic counterpart of the random $K'$, as well as the argument variable of the function $f_{K'}(k')$. There is evidently considerable sampling error in Fig. 3, despite the fact that the sample size, $n = 2086$, is relatively large. A Chi-square goodness-of-fit test has a $p$-value that is effectively 0, implying that the null hypothesis that the data follow a lognormal distribution is rejected. However, this result is not particularly surprising when $n$ is so large — goodness-of-fit tests do not necessarily indicate if a distribution is reasonable, just whether or not the data are precisely lognormally distributed. As seen in Fig. 3, the data are clearly not precisely lognormally distributed, but neither do they precisely follow any other common distribution. Nevertheless, the lognormal distribution can be seen to be a reasonable fit to the frequency–density plot. In addition, the lognormal distribution was also used by Liza et al. (2017) to develop analytical solutions to the probability of type I and II errors, which will be used here to estimate the sample size required for the QC program of this case study.

## Correlation length

Having established a reasonable marginal lognormal distribution for hydraulic conductivity at each point in the field, the next step is to estimate the correlation structure of the hydraulic conductivity random field model since this accounts for the spatial dependence between conductivity values. As in Fenton et al. (2015) and Liza et al. (2017), the correlation structure is assumed to be Markovian

$$\rho(\tau) = \exp\left(-\frac{2|\tau|}{\theta}\right) \tag{1}$$

where $\tau$ is the distance between two points in the field. Equation 1 has a single parameter, $\theta$, the correlation length, which must be estimated from the sample.

The method of matching the theoretical correlation function, eq. 1, to the sample correlation function was used in this study (e.g., Lloret-Cabot et al. 2014) as it is commonly used due to its simplicity and reasonable accuracy. To do this, a linear regression was performed to obtain the value of $\theta$ that led to the minimum sum of squared errors between the theoretical and sample functions. To obtain the sample correlation function, the method of moments was employed, and three different correlation lengths were estimated — the correlation length in the $x$ direction, $\theta_x$, the correlation length in the $y$ direction, $\theta_y$, and the correlation length obtained by assuming isotropy, $\theta_i$. The sample correlation functions for the three cases are obtained from equispaced data on a $\Delta_x \times \Delta_y$ grid of size $n_x \times n_y$ as follows:

$$\hat{\rho}_x(m\Delta x) = \frac{1}{\hat{\sigma}^2_{K'}[n_y(n_x - m) - 1]} \sum_{i=1}^{n_x-m} \sum_{j=1}^{n_y} X_{i,j}X_{i+m,j} \qquad m = 0, 1, \ldots, n_x - 1 \tag{2}$$

$$\hat{\rho}_y(m\Delta y) = \frac{1}{\hat{\sigma}^2_{K'}[n_x(n_y - m) - 1]} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y-m} X_{i,j}X_{i,j+m} \qquad m = 0, 1, \ldots, n_y - 1 \tag{3}$$

$$\hat{\rho}_i(m\Delta x) = \frac{1}{\hat{\sigma}^2_{K'}[(n_x - m)n_y + n_x(n_y - m) - 1]} \times \left( \sum_{i=1}^{n_x-m} \sum_{j=1}^{n_y} X_{i,j}X_{i+m,j} + \sum_{i=1}^{n_x} \sum_{j=1}^{n_y-m} X_{i,j}X_{i,j+m} \right) \qquad m = 0, 1, \ldots, \min(n_x, n_y) \tag{4}$$

where $\Delta x$ and $\Delta y$ are the spacings of the samples in the $x$ and $y$ directions, respectively; $n_x$ and $n_y$ are the number of samples in the $x$ and $y$ directions, respectively; and $X_{i,j}$ is the deviation in normalized hydraulic conductivity about its sample mean ($= K'_{i,j} - \hat{\mu}_{K'}$), $K'_{i,j}$ is the normalized conductivity value at spatial coordinate $((i-1)\Delta x, (j-1)\Delta y)$. Note that eq. 4 can only be properly used if $\Delta x = \Delta y$, which will be assumed in this study, as discussed shortly. The sample mean and variance, used in eqs. 2–4, are obtained using the classical unbiased estimators

$$\hat{\mu}_{K'} = \frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} K'_{i,j} \tag{5}$$

$$\hat{\sigma}^2_{K'} = \frac{1}{(n_x n_y - 1)} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \left(K'_{i,j} - \hat{\mu}_{K'}\right)^2 \tag{6}$$

The regression used to estimate the correlation length follows the procedure given by Fenton et al. (2018, eqs. 30 to 35).

It can be seen in Fig. 2 that the hydraulic conductivity sampling locations used in this study are not equispaced. Because the estimators used in eqs. 2–4 require equispaced data, the scattered hydraulic conductivity dataset was transformed into a $\Delta x = \Delta y =$ 5 m spaced dataset, using an interpolation method called Delaunay triangulation (Delaunay 1934) as provided by the MATLAB class "TriscatteredInterp". The basic idea is to use the data locations to define a set of triangles with vertices at the data locations for randomly located data. Each triangle then defines a plane that can be used to interpolate between the data points at the desired grid points. The MATLAB code required to perform the grid interpolation is presented in Appendix A.

An assumption made in eqs. 2–4 is that the grid is rectangular, of size $n_x \times n_y$. To account for the nonrectangular shape of the actual S/S site shown in Fig. 2, the correlation function estimation was restricted to the two rectangular areas shown in the figure using dashed lines. The reason for selecting two different sizes, as shown on Fig. 2, was to examine whether size and (or) location had a significant effect on the correlation length estimate. These areas were relatively densely sampled and are of sizes 205 m × 125 m on the left and 55 m × 85 m on the right. The two areas were then interpolated on a 5 m × 5 m grid to produce the directional and isotropic sample correlation functions that were then matched via regression to produce estimates of the directional and isotropic correlation lengths. The 5 m × 5 m size was chosen to represent the actual cell size as closely as possible. As shown next, the size and location of these two rectangular areas had negligible effect on the calculated correlation length.

For the 205 m × 125 m rectangle, estimated $x$ and $y$ directional and isotropic correlation lengths are 16.0, 10.2, and 12.3 m, respectively. For the 55 m × 85 m rectangle, estimated $x$ and $y$ directional and isotropic correlation lengths are 14.6, 8.9, and 11.1 m, respectively. The isotropic correlation lengths are between the $x$ and $y$ direction correlation lengths, as expected, because they are obtained by averaging over all data pairs in both directions. Figures 4 and 5 show directional and isotropic estimated correlation functions, for the 205 m × 125 m and 55 m × 85 m subsites, respectively, along with the fitted isotropic correlation function. The estimated functions become increasingly erratic at higher separation distances (higher values of $\tau$) because there are fewer sample pairs to average as the separation distance increases. In Figs. 4 and 5, $\hat{\theta}$ is the estimate of the true correlation length. This estimate is performed by fitting the Markov correlation function (eq. 1) to the sample correlation by least squares regression. Because the fitted correlation function is exponentially decaying with only a single parameter, it weights more heavily the behavior of the sample correlation function at small $\tau$. In other words, the erratic behavior of the sample correlation function at large $\tau$ is largely ignored by the regression, which is appropriate given that the correlation estimates at large separation distances become increasingly uncertain.

As the estimated directional correlation lengths are not substantially different for practical purposes, this paper will assume an isotropic correlation length of 12 m to estimate the type I and type II error probabilities in the next section. The error probabilities will be assessed using the analytical solution presented by Liza et al. (2017), which assumes an isotropic correlation structure. If an anisotropic correlation structure is actually desired, the error probabilities would need to be estimated using random field simulations (see, e.g., Fenton et al. 2015).

**Fig. 4.** Directional and isotropic correlation functions at different lags, estimated using 205 m × 125 m subsite having 5 m spaced hydraulic conductivity values.

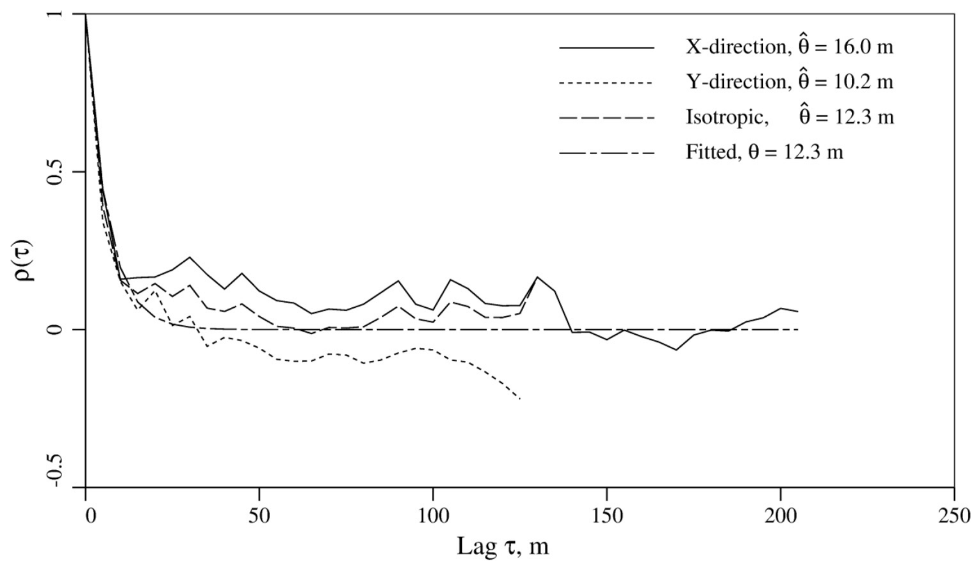
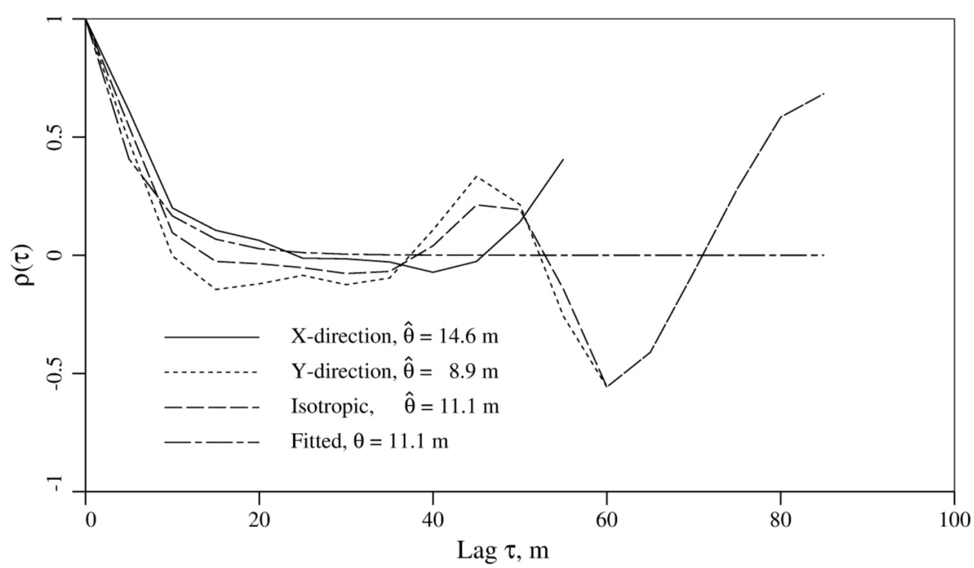



**Fig. 5.** Directional and isotropic correlation functions at different lags, estimated using 55 m × 85 m subsite having 5 m spaced hydraulic conductivity values.




## Error probabilities: evaluating sample density approach

As mentioned previously, the current sampling regulation for cement-based S/S is based on the sample density method, which specifies the same required number of samples for equivalent sized sites even if the sites have different levels of spatial variability. Given that differing variability results in differing reliability of a QC sampling program, Fenton et al. (2015) and Liza et al. (2017) proposed a method to determine the sampling requirement for the QC of cement-based S/S based on the decision error probabilities (i.e., types I and II discussed in Fenton et al. 2015). In this section, the sampling requirement for the QC of this studied cement-based S/S site is determined based on Liza et al.'s (2017) analytical error probability approach.

This paper will concentrate on a subsite of size 55 m × 85 m, from which a series of $n = n_x \times n_y$ virtual samples (i.e., samples extracted from the random field simulation) will be assumed to be gathered. For simplicity, $n_x = n_y$, and the number of virtual samples taken from the subsite will range from 1 to 900 equally spaced samples throughout the subsite. Virtual sampling will be performed by discretizing the 55 m × 85 m subsite into 2048 × 2048 potential virtual sampling subareas, each of size 55/2048 m × 85/2048 m. The virtual sampling will be carried out at equispaced intervals in the $x$ and $y$ directions. For example, if $n_x$ samples are taken in the $x$ direction, then they are taken from the subarea numbered $i[m_x/(n_x+1)]$ where $m_x = 2048$ is the number of subareas in the $x$ direction and $i = 1, 2, \ldots, n_x$. Sampling is performed similarly in the $y$ direction.

The geometric average, $k_G$, of these sample values is an estimate of the "actual" effective conductivity of the cell, $k_{\mathrm{eff}}$, which is assumed to be a geometric average of conductivities over the subsite. Given that, as shown previously, the hydraulic conductivity is at least approximately lognormally distributed, both $k_G$ and $k_{\mathrm{eff}}$ will also be lognormally distributed (see Fenton and Griffiths 2008) and their means and variances are derived by Liza et al. (2017) in their Appendix A.

The probabilities of type I ($p_1$) and type II ($p_2$) errors are computed using the following equations, as derived by Owen (1956) and used by Liza et al. (2017).

If $hw > 0$ or if $hw = 0$ with $h$ or $w \geq 0$ (note that $h$, $w$, $a_h$, and $a_w$ are defined subsequently in eqs. 9 and 10):

$$(7a)\qquad p_1 = \frac{1}{2}\Phi(h) - \frac{1}{2}\Phi(w) + T(h, a_h) + T(w, a_w)$$

$$(7b)\qquad p_2 = \frac{1}{2}\Phi(w) - \frac{1}{2}\Phi(h) + T(h, a_h) + T(w, a_w)$$

If $hw < 0$ or if $hw = 0$ with $h$ or $w < 0$:

$$(8a)\qquad p_1 = \frac{1}{2}\Phi(h) - \frac{1}{2}\Phi(w) + T(h, a_h) + T(w, a_w) + \frac{1}{2}$$

$$(8b)\qquad p_2 = \frac{1}{2}\Phi(w) - \frac{1}{2}\Phi(h) + T(h, a_h) + T(w, a_w) + \frac{1}{2}$$

where

$$(9a)\qquad h = \frac{\ln k_{\mathrm{crit}} - \mu_{\ln k_{\mathrm{G}}}}{\sigma_{\ln k_{\mathrm{G}}}}$$

$$(9b)\qquad w = \frac{\ln k_{\mathrm{crit}} - \mu_{\ln k_{\mathrm{eff}}}}{\sigma_{\ln k_{\mathrm{eff}}}}$$

$$(10a)\qquad a_h = \frac{w}{h\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}}$$

$$(10b)\qquad a_w = \frac{h}{w\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}}$$

$$(11a)\qquad T(h, a_h) = \frac{1}{2\pi}\int_0^{a_h} \frac{\exp\left[-\frac{1}{2}h^2(1+u^2)\right]}{1+u^2}du \qquad \text{when } a_h \leq 1$$

$$(11b)\qquad T(h, a_h) = \frac{1}{2}\Phi(h) + \frac{1}{2}\Phi(ha_h) - \Phi(h)\Phi(ha_h) - T\left(ha_h, \frac{1}{a_h}\right) \qquad \text{when } a_h > 1$$

$$(11c)\qquad T(w, a_w) = \frac{1}{2\pi}\int_0^{a_w} \frac{\exp\left[-\frac{1}{2}w^2(1+v^2)\right]}{1+v^2}dv \qquad \text{when } a_w \leq 1$$

$$(11d)\qquad T(w, a_w) = \frac{1}{2}\Phi(w) + \frac{1}{2}\Phi(wa_w) - \Phi(w)\Phi(wa_w) - T\left(wa_w, \frac{1}{a_w}\right) \qquad \text{when } a_w > 1$$

and where $\Phi$ is the standard normal cumulative distribution function; $\mu_{\ln k_{\mathrm{G}}}$ and $\sigma_{\ln k_{\mathrm{G}}}$ are the mean and standard deviation of the log-sample geometric average, $\ln k_{\mathrm{G}}$, respectively; $\mu_{\ln k_{\mathrm{eff}}}$ and $\sigma_{\ln k_{\mathrm{eff}}}$ are the mean and standard deviation of the log-effective hydraulic conductivity, $\ln k_{\mathrm{eff}}$, respectively; $\rho$ is the correlation coefficient between $\ln k_{\mathrm{eff}}$ and $\ln k_{\mathrm{G}}$; $u = (r - \mu_{\ln k_{\mathrm{eff}}})/\sigma_{\ln k_{\mathrm{eff}}}$; and $v = (s - \mu_{\ln k_{\mathrm{G}}})/\sigma_{\ln k_{\mathrm{G}}}$.

The parameters of the random hydraulic conductivity field (i.e., normalized mean, $\mu_{K'} = 0.47$, coefficient of variation, $\nu_{K'} = 1.7$, and correlation length, $\theta_{\ln K} = 12$ m) estimated from the Sydney Tar Ponds site cement-based S/S system (Fenton et al. 2015) are used to compute the type I and type II error probabilities for number of samples of 1, 4, 9, and 16, taken from the 55 m × 85 m subsite of the existing project S/S site. The results are presented in Table 1.

**Table 1.** Probabilities of type I ($p_1$) and type II ($p_2$) errors for $\mu_{K'} = 0.47$, $\nu_{K'} = 1.7$, $\theta_{\ln K} = 12$ m, and different $n$ over 55 m × 85 m subsite.

| $n$ over 55 m × 85 m subsite | $p_1$ | $p_2$ |
|---|---|---|
| 1 | <0.0001 | 0.1003 |
| 4 | <0.0001 | 0.0064 |
| 9 | <0.0001 | 0.0002 |
| 16 | <0.0001 | <0.0001 |

Table 1 indicates that when the effective hydraulic conductivity of the cell is known ahead of time to be about half the critical value, i.e., $\mu_{K'} = 0.47$, both type I and type II error probabilities are extremely small when the number of samples is 16 or greater. According to the current sampling requirements of 1 sample/500 m³ specified by the USACE (2000) for the QC program of cement-based S/S, this 55 m × 85 m subsite would require 36 samples (= (1/500) × 55 × 85 × 3.9, using an average thickness of 3.9 m). Thus, the current QC sampling regulation of cement-based S/S seems to be conservative for this particular S/S system (see Table 1) if the value of $\mu_{K'}$ is known to be significantly smaller than 1.0.

However, the type I and type II error probabilities presented in Table 1 cannot be obtained prior to the QC program, given that the hydraulic conductivity mean, coefficient of variation, and correlation length would be unknown. To assess QC sampling requirements, the "worst case" hydraulic conductivity mean, coefficient of variation, and correlation length need to be used in the determination of error probabilities. The "worst case" parameters are those that result in the highest probabilities of type I and II errors. According to the results presented by Fenton et al. (2015), the "worst case" normalized mean and coefficient of variation of hydraulic conductivity are approximately 1.5 and 1.0, respectively. A mean of 1.0 will also be investigated. It will be assumed that most sites of this nature will have a similar correlation length, and so the correlation length of 12 m will be retained for this investigation. The type I and type II error probabilities are computed for the number of samples ranging from 1 to 900 and the results are presented in Tables 2 and 3.

The results presented in Table 2 indicate that for $\theta_{\ln K} = 12$ m, $\mu_{K'} = 1.0$, and $\nu_{K'} = 1.0$, at least 25 samples are required to achieve a target 5% probability for both type I and II errors. Notice that for $n = 25$, the probability of a type I error (failing to detect an unacceptable cell, which is more important to avoid) is only 0.14%.

The results presented in Table 3 indicate that for $\theta_{\ln K} = 12$ m, $\mu_{K'} = 1.5$, and $\nu_{K'} = 1.0$, at least 400 samples are required to achieve a target 5% probability for both type I and II errors.

As this is a "worst case", $n = 400$ samples taken over a 55 m × 85 m subsite of the entire S/S site should be conservative in targeting a 5% probability for both type I and II errors. If the "worst case" exists, the United States Army Corps of Engineers, USACE (2000), sampling requirement over 55 m × 85 m subsite of 36 would exceed a 5% probability for both type I and type II errors: $p_1 = 13\%$ and $p_2 = 8\%$. A probability of failing to detect an unacceptable effective hydraulic conductivity of 13% is probably too high, hence the USACE requires too few samples in this hypothetical scenario.

These results raise a number of questions: (*i*) What error probabilities are acceptable? and (*ii*) Should the worst case be adopted in developing sampling requirements? The determination of acceptable error probabilities should depend on the type of error (lower for the type I error, as it is more important to avoid) and can only really be properly estimated through a full-fledged risk as-

**Table 2.** Probabilities of type I ($p_1$) and type II ($p_2$) errors for $\mu_{K'} = 1.0$, $\nu_{K'} = 1.0$, $\theta_{\ln K} = 12$ m, and different $n$ over the 55 m × 85 m subsite.

| $n$ over 55 m × 85 m subsite | Error probability $p_1$ | $p_2$ |
|---|---|---|
| 1 | 0.0020 | 0.3357 |
| 4 | 0.0017 | 0.2058 |
| 9 | 0.0016 | 0.1224 |
| 16 | 0.0015 | 0.0752 |
| 25 | 0.0014 | 0.0488 |
| 36 | 0.0013 | 0.0339 |
| 49 | 0.0010 | 0.0244 |
| 64 | 0.0009 | 0.0188 |
| 81 | 0.0009 | 0.0151 |
| 100 | 0.0009 | 0.0122 |
| 225 | 0.0007 | 0.0059 |
| 400 | 0.0005 | 0.0039 |
| 625 | 0.0004 | 0.0031 |
| 900 | 0.0004 | 0.0022 |

**Table 3.** Probabilities of type I ($p_1$) and type II ($p_2$) errors for $\mu_{K'} = 1.5$, $\nu_{K'} = 1.0$, $\theta_{\ln K} = 12$ m, and different $n$ over the 55 m × 85 m subsite.

| $n$ over 55 m × 85 m subsite | Error probability $p_1$ | $p_2$ |
|---|---|---|
| 1 | 0.3165 | 0.1737 |
| 4 | 0.2449 | 0.1287 |
| 9 | 0.2073 | 0.1136 |
| 16 | 0.1769 | 0.1009 |
| 25 | 0.1527 | 0.0904 |
| 36 | 0.1336 | 0.0817 |
| 49 | 0.1173 | 0.0736 |
| 64 | 0.1051 | 0.0675 |
| 81 | 0.0951 | 0.0622 |
| 100 | 0.0862 | 0.0575 |
| 225 | 0.0587 | 0.0418 |
| 400 | 0.0447 | 0.0326 |
| 625 | 0.0376 | 0.0276 |
| 900 | 0.0298 | 0.0228 |

sessment, taking into account remediation costs and costs of failure. The main problem would be the estimation of cost of failure, as it involves potential environmental and health impacts over time, which could be difficult to quantify. For this paper, a target maximum error probability of 5% has been assumed, which is believed to be a reasonable value.

Regarding whether the worst case should be assumed, it will certainly be generally true that the actual effective hydraulic conductivity of a cell will be unknown prior to sampling. The use of a worst case is thus conservative. As it is unlikely that the effective hydraulic conductivity will actually be at the worst case of $\mu_{K'}$ = 1.5, the actual error probabilities will be less than 5%. Notice, by comparing Table 3 with Table 2, that the probability of a type I error falls rapidly as $\mu_{K'}$ falls below the worst case.

It is important to realize that the actual project had a normalized mean of 0.47 with a coefficient of variation of 1.7 (estimated from a very large sample). In this case, if a limited number of samples had been taken, the probability of type I and II errors would still have been small. See Table 1, where for the actual site only four samples are required to reduce the error probabilities to less than 5%.

When a project is started, the true mean is of course unknown. One practical approach is to target the field mean to be less than

**Table 4.** Probabilities of type I ($p_1$) and type II ($p_2$) errors for varying $\mu_{K'}$, $\nu_{K'} = 1.0$, $\theta_{\ln K} = 12$ m, and $n = 16$.

| $\mu_{K'}$ | $p_1$ | $p_2$ |
|---|---|---|
| 1.00 | 0.0015 | 0.075 |
| 0.95 | 4.5E–04 | 0.051 |
| 0.90 | 1.2E–04 | 0.032 |
| 0.85 | 1.7E–05 | 0.019 |
| 0.80 | 7.3E–06 | 0.0099 |
| 0.75 | 4.4E–06 | 0.0048 |
| 0.70 | 2.0E–06 | 0.0020 |
| 0.65 | 2.0E–07 | 7.4E–04 |
| 0.60 | 1.3E–08 | 2.3E–04 |
| 0.55 | 5.0E–10 | 5.7E–05 |
| 0.50 | 1.0E–11 | 1.1E–05 |
| 0.47 | 6.6E–13 | 3.3E–06 |

the critical mean value. Table 4 shows how the error probabilities change as the target (assumed true) mean drops below the critical value for a fixed number of samples using the analytical formulations developed by Liza et al. (2017). The USACE (2000) requirement for this particular 55 m × 85 m subsite is approximately 16 samples. Table 4 is thus developed assuming 16 samples.

The practical motivation for Table 4 is that it can be used to decide on what the construction field target should be. As shown in Table 4, the error probabilities fall below 5% when $\mu_{K'} \leq 0.90$. This does not mean that the critical mean value has been moved, just that the construction target can be conservatively aimed to achieve desired error probabilities.

## Summary and conclusions

In this paper, a set of hydraulic conductivity data with corresponding locations obtained from an existing cement-based S/S system is statistically analyzed to assess its spatial variability. The spatial variability of hydraulic conductivity is described by a random field with a distribution and correlation length. To make use of the classical estimators for the correlation structure (which are based on equispaced data), an irregularly scattered hydraulic conductivity dataset is interpolated onto a 2D 5 m grid using the linear interpolation method available in MATLAB under the class "TriscatteredInterp". Two subsites, having interpolated 5 m spaced hydraulic conductivity values, of sizes 205 m × 125 m and 55 m × 85 m are used to estimate directional and isotropic correlation lengths. To assess QC sampling requirements, the spatial variability of hydraulic conductivity of the system is then used to compute the error probabilities (i.e., type I and type II) for different numbers of samples taken from a 55 m × 85 m subsite of the entire cement-based S/S site. The type I and type II error probabilities are also computed for the "worst case" conditions of hydraulic conductivity mean and coefficient of variation using the average correlation length estimated for the site, and varying number of samples to provide recommendations for conservative QC sampling requirements over 55 m × 85 m subsite.

The following conclusions can be drawn from this study:

- A lognormal distribution with mean $\mu_{K'}$ = 0.47 and coefficient of variation $\nu_{K'}$ = 1.7 ($\mu_{\ln K'}$ = –1.31 and $\sigma_{\ln K'}$ = 1.02) was found to be a reasonable fit to the site's hydraulic conductivity data.
- The $x$ and $y$ directional and isotropic correlation lengths are estimated to be 16.0, 10.2, and 12.3 m, respectively, over the 205 m × 125 m subsite, while the $x$ and $y$ directional and isotropic correlation lengths are estimated to be 14.6, 8.9, and 11.1 m, respectively, over the 55 m × 85 m subsite. An average isotropic correlation length for the site is found to be 11.7 m, or approximately 12 m.
- For the spatial variability of hydraulic conductivity of the S/S system, i.e., for $\mu_{K'}$ = 0.47, $\nu_{K'}$ = 1.7, and $\theta_{\ln K}$ = 12 m, the type I

error probabilities for any of the number of samples of 1, 4, 9, and 16 are found to be less than 0.0001; whereas, the type II error probabilities for the number of samples of 1, 4, 9, and 16 are found to be 0.1003, 0.0064, 0.0002, and <0.0001, respectively. The USACE (2000) sampling recommendation for the QC program of cement-based S/S, which would specify 36 samples for this subsite volume, is conservative for this particular case. However, this is somewhat misleading because the knowledge about the value of $\mu_{K'}$ is actually based on 2086 samples (over the entire site). If only 36 samples had been taken, and the sample mean had been $\mu_{K'} = 0.47$, there would be considerably less confidence in this estimate, and consequently a considerably higher probability that the actual value of $\mu_{K'}$ exceeds $k_{\text{crit}}$.

- The probabilities of type I and type II errors computed for various numbers of samples and the "worst case" conditions of hydraulic conductivity mean (i.e., 1.0 and 1.5 times the regulatory value), coefficient of variation (i.e., 1.0), and correlation length (i.e., 12 m) suggests that 400 samples should be taken over a 55 m × 85 m site (3.8 m in thickness) to achieve a maximum 5% probability for both type I and type II errors. Although this worst case scenario was not actually present for this case study, the USACE (2000) sampling recommendation for the QC program of cement-based S/S would be unconservative for the worst-case conditions.
- For practical purposes, when undertaking an S/S project such as this, to limit type I and II errors if the USACE (2000) sample size approach is used, the conservative approach may be to target a field value of mean hydraulic conductivity less than what the "regulatory" mean hydraulic conductivity is required. It is shown in this work that by targeting a mean hydraulic conductivity in the field to even 90% of what a "regulatory" value will be, will reduce type I and II errors to acceptable values. In practice, this would involve adjusting the mix design to reduce the hydraulic conductivity.

## Acknowledgements

The authors gratefully acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada and its STEWARD program. The authors also gratefully acknowledge Belinda Campbell from Public Service and Procurement Canada for providing access to the dataset. The academic work performed in this paper is not a reflection of the opinions/views of Public Service and Procurement Canada in any way and is the sole work of the authors. The authors also thank Kirklyn Davidson for assistance in preparing Fig. 1 and Alana Devanney from Public Service and Procurement Canada for assistance in processing the original dataset.

## References

AECOM. 2008. Remediation of the Tar Ponds and Coke Ovens sites design and construction oversight services: design element TP6 – Part B, Solidification/stabilization and channel construction: 100% Design Report. Available from http://www.tarpondscleanup.ca/ [accessed 26 July 2019].

Al-Tabbaa, A., and Perera, A.S.R. 2002. UK stabilisation/solidification treatment and remediation. STARNET State of Practice Report. Cambridge University Engineering Department, Cambridge, U.K.

AMEC. 2005. Remediation of Sydney Tar Ponds and Coke Ovens sites, Environmental Impact Statement, Sydney, Nova Scotia. Available from http://www.tarpondscleanup.ca/ [accessed 26 July 2019].

ASTM. 2010. Standard test methods for measurement of hydraulic conductivity of saturated porous materials using a flexible wall permeameter. ASTM standard D5084-10. ASTM, West Conshohocken, Pa. doi:10.1520/D5084-10.

Benson, C.H. 1991. Predicting excursions beyond regulatory thresholds of hydraulic conductivity from quality control measurements. *In* Proceedings of the 1st Canadian Conference on Environmental Geotechnics, Canadian Geotechnical Society, Montreal, Quebec, pp. 447–454.

Benson, C.H. 1993. Probability distributions for hydraulic conductivity of compacted soil liners. Journal of Geotechnical Engineering, **119**(3): 471–486. doi:10.1061/(ASCE)0733-9410(1993)119:3(471).

Byers, E., and Stephens, D. 1983. Statistical and stochastic analysis of hydraulic conductivity in fluvial sand. Soil Science Society of America Journal, **47**(6): 1072–1081. doi:10.2136/sssaj1983.03615995004700060003x.

Conner, J.R. 1990. Chemical fixation and solidification of hazardous wastes. Van Nostrand Reinhold, New York.

Dasaka, S.M., and Zhang, L.M. 2012. Spatial variability of in situ weathered soil. Géotechnique, **62**(5): 375–384. doi:10.1680/geot.8.P.151.3786.

DeGroot, D.J., and Baecher, G.B. 1993. Estimating autocovariance of in-situ soil properties. Journal of Geotechnical Engineering, **119**(1): 147–166. doi:10.1061/(ASCE)0733-9410(1993)119:1(147).

Delaunay, B. 1934. Sur la sphère vide, Bulletin de l'Acacdémie des Sciences de l'URSS. Class des sciences mathématiques et naturelles, **6**: 793–800.

Fenton, G.A. 1999. Estimation for stochastic soil models. Journal of Geotechnical and Geoenvironmental Engineering, **125**(6): 470–485. doi:10.1061/(ASCE)1090-0241(1999)125:6(470).

Fenton, G.A., and Griffiths, D.V. 2008. Risk assessment in geotechnical engineering. John Wiley and Sons, New Jersey, USA.

Fenton, G.A., Liza, R., Lake, C.B., Menzies, W.T., and Griffiths, D.V. 2015. Statistical sample size for quality control programs of cement-based "solidification/stabilization". Canadian Geotechnical Journal, **52**(10): 1620–1628. doi:10.1139/cgj-2013-0478.

Fenton, G.A., Naghibi, F., and Hicks, M.A. 2018. Effect of sampling plan and trend removal on residual uncertainty. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, **12**(4): 253–264. doi:10.1080/17499518.2018.1455106.

Freeze, R.A. 1975. A stochastic-conceptual analysis of one-dimensional groundwater flow in non-uniform homogeneous media. Water Resources Research, **11**(5): 725–741. doi:10.1029/WR011i005p00725.

Freeze, R.A., and Cherry, J. 1979. Groundwater. Prentice Hall, Inc., New Jersey, USA.

Hills, C.D., Bates, E., and Gunning, P. 2015. Stabilisation and solidification of contaminated soil and waste. Science and Practice. [ISBN-10: 0993272940.]

ITRC. 2011. Development of performance specifications for solidification/stabilization. S/S-1. Interstate Technology and Regulatory Council, Solidification/Stabilization Team, Washington, D.C. Available from www.itrcweb.org.

Jaksa, M.B., Kaggwa, W.S., and Brooker, P.I. 1993. Geostatistical modelling of the undrained shear strength of stiff, overconsolidated, clay. *In* Proceedings of the Conference on Probabilistic Methods in Geotechnical Engineering, Canberra, Australia, pp. 185–194.

Jaksa, M.B., Kaggwa, W.S., and Brooker, P.I. 1999. Experimental evaluation of the scale of fluctuation of a stiff clay. *In* Proceedings of the 8th International Conference on the Application of Statistics and Probability. Vol. 1. Sydney, Australia, pp. 415–422.

Johnson, G., Grumbley, W.S., and Boutwell, G.P. 1990. Field verification of clay liner hydraulic conductivity. *In* Waste containment systems: Construction, regulation, and performance. pp. 226–245.

Krapac, I.G., Panno, S.V., Rehfeldt, K.R., Herzog, B.L., Hensel, B.R., and Cartweight, K. 1989. Hydraulic properties of an experimental soil liner: preliminary results. *In* Proceedings of the 12th Annual Madison Waste Conference, University of Wisconsin Extension, pp. 395–411.

Liza, R., Fenton, G.A., Lake, C.B., and Griffiths, D.V. 2017. An analytical approach to assess quality control sample sizes of cement-based "solidification/stabilization". Canadian Geotechnical Journal, **54**(3): 419–427. doi:10.1139/cgj-2016-0218.

Lloret-Cabot, M., Hicks, M.A., and Wong, S.Y. 2012. Soil characterisation of an artificial island accounting for heterogeneity. *In* Proceedings of the American Society of Civil Engineering Geocongress 2012, San Francisco, Calif., pp. 2816–2825. doi:10.1061/9780784412121.288.

Lloret-Cabot, M., Hicks, M.A., and Nuttall, J.D. 2013. Investigating the scales of fluctuation of an artificial sand island. *In* Proceedings of the International Conference on Geotechnical Installations in Geotechnical Engineering 2013. Rotterdam, pp. 192–197.

Lloret-Cabot, M., Fenton, G.A., and Hicks, M.A. 2014. On the estimation of scale of fluctuation in geostatistics. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, **8**(2): 129–140. doi:10.1080/17499518.2013.871189.

Nova Scotia Geographic Data Directory. 2016. Available from https://nsgi.novascotia.ca/gdd/

Owen, D.B. 1956. Tables for computing bivariate normal probabilities. The Annals of Mathematical Statistics, **27**(4): 1075–1090. doi:10.1214/aoms/1177728074.

Phoon, K.K., and Fenton, G.A. 2004. Estimating sample autocorrelation functions using bootstrap. *In* Proceedings of the 9th Specialty Conference on Probabilistic Mechanics and Structural Reliability, American Society of Civil Engineering, Reston, Va.

Rogowski, A., Weinrich, B., and Simmons, D. 1985. Permeability assessment in a compacted clay liner. *In* Proceedings of the Eighth Annual Madison Waste Conference, University of Wisconsin, Madison, pp. 315–337.

STPA. 2019. Sydney Tar Ponds Agency website. Available from http://www.tarpondscleanup.ca/ [accessed 26 July 2019].

USACE. 2000. Solidification/Stabilization (S/S) of contaminated material. Section 02160 A, Washington, D.C.

USEPA. 2013. Superfund Remedy Report. 14th ed. EPA 542-R-13-016. Office of Solid Waste and Emergency Response, United States Environmental Protection Agency.

Vanmarcke, E.H. 1984. Random fields: analysis and synthesis. The MIT Press, Cambridge, Massachusetts, USA.

Wackernagel, H. 2003. Multivariate geostatistics: An introduction with applications. Springer, Germany.
Wickremesinghe, D., and Campanella, R.G. 1993. Scale of fluctuation as a descriptor of soil variability. *In* Probabilistic methods in geotechnical engineering. *Edited by* K.S. Li and S.C.R. Lo. Balkema, Rotterdam, pp. 233–239.
Zhang, L.L., Zhang, L.M., and Tang, W.H. 2008. Similarity of soil variability in centrifuge models. Canadian Geotechnical Journal, **45**(8): 1118–1129. doi:10.1139/T08-066.

## List of symbols

$a_h$ parameter defined by eq. 10*a*
$a_w$ parameter defined by eq. 10*b*
$f_{K'}$ probability density function of $K'$
$h$ standardization of $k_{\mathrm{crit}}$ with respect to $k_G$
$i, j, m$ integer counters
$K$ random hydraulic conductivity field
$K_G$ sample geometric average
$K'$ random hydraulic conductivity field normalized by $k_{\mathrm{crit}}$
$K'_{i,j}$ normalized hydraulic conductivity at spatial location $((i-1)\Delta x, (j-1)\Delta y)$
$k'$ realization of $K'$
$k_{\mathrm{crit}}$ regulatory hydraulic conductivity
$k_{\mathrm{eff}}$ effective hydraulic conductivity
$n$ number of samples
$n_x, n_y$ number of samples in $x$ and $y$ directions, respectively
$p_1$ probability of type I error
$p_2$ probability of type II error
$r$ dummy variable of integration
$s$ dummy variable of integration
$T$ bivariate normal probability approximation functions given by eq. 11
$u$ variable of integration $(= (r - \mu_{\ln k_{\mathrm{eff}}})/\sigma_{\ln k_{\mathrm{eff}}})$
$v$ variable of integration $(= (s - \mu_{\ln k_G})/\sigma_{\ln k_G})$
$v_{K'}$ coefficient of variation of $K'$
$w$ standardization of $k_{\mathrm{crit}}$ with respect to $k_{\mathrm{eff}}$
$X_{i,j}$ deviation in normalized hydraulic conductivity about its sample mean $(= K'_{i,j} - \mu_{K'})$
$x, y$ coordinate directions
$\Delta x, \Delta y$ sample spacing in $x$ and $y$ directions, respectively
$\theta$ random field correlation length
$\hat{\theta}$ estimate of random field correlation length
$\theta_{\mathrm{i}}$ random field isotropic correlation length
$\theta_{\ln K}$ correlation length of $\ln K$ random field
$\theta_x$ random field correlation length in $x$ direction
$\theta_y$ random field correlation length in $y$ direction
$\mu_{K'}$ mean hydraulic conductivity
$\mu_{\ln K}$ mean of log-hydraulic conductivity field, $\ln K$
$\mu_{\ln K'}$ mean of tlognormalized hydraulic conductivity field, $\ln K'$
$\mu_{\ln k_{\mathrm{eff}}}$ mean of log-effective hydraulic conductivity, $\ln k_{\mathrm{eff}}$
$\mu_{\ln k_G}$ mean of og-sample geometric average, $\ln k_G$
$\rho$ correlation coefficient between $\ln k_{\mathrm{eff}}$ and $\ln k_G$
$\rho(\tau)$ correlation function, giving correlation coefficient between log-hydraulic conductivity points separated by distance $\tau$
$\hat{\rho}_{\mathrm{i}}$ sample correlation function in all directions (isotropic)
$\hat{\rho}_x$ sample correlation function in $x$ direction
$\hat{\rho}_y$ sample correlation function in $y$ direction
$\sigma_{\ln K'}$ standard deviation of lognormalized hydraulic conductivity field, $\ln K'$
$\sigma_{\ln k_{\mathrm{eff}}}$ standard deviation of log-sample geometric average, $\ln k_G$
$\sigma_{\ln k_G}$ standard deviation of log-effective hydraulic conductivity, $\ln k_{\mathrm{eff}}$
$\tau$ separation distance (lag) between two hydraulic conductivity values–observations
$\Phi$ standard normal cumulative distribution function

## Appendix A. MATLAB code that generates interpolated hydraulic conductivity values on 2D, 5 m grid

A spreadsheet containing three columns of information: $x$, $y$ and $z$, where $(x, y)$ is the position of each observation in metres and $z$ is the corresponding normalized hydraulic conductivity value, is used to obtain the interpolated values on a 5 m grid in 2D space. Using the scattered dataset, the "TriscatteredInterp" first creates a function, which fits a convex hull. A grid is then created with $x$ positions ranging from −660 to 580 and $y$ positions ranging from 115 to 440, with 5 m spacing in both $x$ and $y$ directions. The function is then evaluated at each query location (i.e., at each grid point). The MATLAB code that generated the interpolated 5 m spaced hydraulic conductivity values is as given below:

```
n = 2086;
x = xlsread('correlation_length.xls','c4:c2089');
y = xlsread('correlation_length.xls','d4:d2089');
z = xlsread('correlation_length.xls','e4:e2089');
F = TriScatteredInterp(x,y,z);
min_x = −660;
delta_x = 5;
max_x = 580;
grid_x = min_x:delta_x:max_x;
min_y = 115;
delta_y = 5;
max_y = 440;
grid_y = min_y:delta_y:max_y;
[qx,qy] = meshgrid(grid_x,grid_y);
qz = F(qx,qy)
```