https://doi.org/10.1016/j.scitotenv.2018.03.049


# Three-dimensional spatial variability of arsenic-containing soil from geogenic source in Hong Kong: implications on sampling strategies

Yat Fai Leung[1]*, Wenfei Liu[1], Jiang-Shan Li[1], Lei Wang[1], Daniel C.W. Tsang[1#], Chi Yan Lo[2], Man Tat Leung[2], Chi Sun Poon[1]

[1]Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

[2]Civil Engineering and Development Department, Hong Kong SAR Government

*corresponding author: andy.yf.leung@polyu.edu.hk

[#]corresponding author: dan.tsang@polyu.edu.hk

## Abstract

Soil contamination by trace elements such as arsenic (As) can pose considerable threats to human health, and need to be carefully identified through site investigation before the soil remediation and development works. However, due to the high costs of soil sampling and testing, decisions on risk management or mitigation strategies are often based on limited data at the site, with substantial uncertainty in the spatial distributions of potentially toxic elements. This study incorporates the restricted maximum likelihood method with three-dimensional spatial autocovariance structure, to investigate the spatial variability features of As-containing soils of geogenic origin. A recent case study in Hong Kong is presented, where more than 550 samples were retrieved and tested for distributions of As concentrations. The proposed approach is applied to characterize their spatial correlation patterns, to predict the As concentrations at unsampled locations, and to quantify the uncertainty of such estimates. The validity of the approach is illustrated by utilizing the multi-stage site investigation data, through which the advantages of the approach over traditional geostatistical methods are revealed and discussed. The new approach also quantifies the effectiveness of soil sampling on reduction of uncertainty levels across the site. This can become a useful indicator for risk management or mitigation strategies, as it is often necessary to balance between the available resources for soil sampling at the site and the needs for proper characterization of contaminant distributions.

**Keywords:** site investigation; trace elements; geogenic arsenic; spatial variability; restricted maximum likelihood; soil remediation.

## 1. Introduction

Soil contamination by metals/metalloids poses an increasing threat to human health and environmental quality across the globe (Tóth et al., 2016; Xia et al., 2017; Plessl et al., 2017). Among the potentially toxic elements, arsenic (As) has received considerable attention over the last few decades due to its high toxicity and environmental risk (Bundschuh et al., 2013; Ehlert et al., 2016; Sandhi et al., 2017). Geochemical properties of As are complex because of its various chemical species and amphoteric nature (Yang et al., 2017; Wu et al., 2017), which makes it challenging for proper assessment and remediation (Rahman et al., 2017). Soil can naturally possess high concentrations of As due to weathering of the parent materials, volcanic eruptions, and forest fires (Beiyuan et al., 2017; Li et al., 2017). In recent years, industrialization and urbanization have also transferred As from used products into the environment, resulting in many industrial contaminated sites (Tsang et al., 2014; Gallego et al., 2016; Wcisło et al., 2016). Consequently, many studies have focused on urban contaminated soils with As from anthropogenic sources, such as agricultural activities and industrial and mining processes (Rieuwerts et al., 2014; Cao et al., 2016; Yoon et al., 2016; González-Fernández et al., 2018). Although geogenic As-containing soil/sediment is a common problem worldwide (Fendorf et al., 2010; Yang et al., 2014), including Hong Kong (Li et al., 2017; Cui et al., 2018), there has been limited discussion on the characterization of spatial variations of geogenic As and the corresponding implications on management/remediation of As-containing sites.

To develop cost-effective management and risk mitigation recommendations, the concentration distribution and variability features (or spatial correlation patterns) of the trace

elements should be established foremost through identifying, mapping, and monitoring processes (Bednářová et al., 2016; Pan et al., 2017). It is because the remediation processes, especially for extraction-based approaches, are often expensive and highly dependent on the types of contaminants and estimated amounts of contaminated soils that need to be treated (Bolan et al., 2014; Tsang and Yip, 2014; Wan et al., 2016). In practice, however, it is difficult to accurately predict the concentrations of metals/metalloids due to their complex spatial distribution patterns, including the occasional occurrence of 'hotspots' with high levels of anthropogenic contamination or geogenic formation. Legislations in various countries (e.g., China, the United States, the Netherlands, Australia, and New Zealand) advocate the use of probabilistic sampling schemes (e.g., square grid, simple random, stratified random) for contaminated site assessments (Waterhouse, 1980; Bell et al., 1983; Gilbert, 1987; Horta et al., 2015), which is similarly applied in Hong Kong (HK EPD, 2011). However, these methods focused on detecting high concentration regions of anthropogenic contaminants and quantifying the extent of such hotspots. This is considered appropriate when there is prior knowledge about the contaminants involved, their transport mechanisms, and the human activities causing the contamination.

For trace elements of geogenic nature, the determination of sampling strategies and characterization of their spatial distributions may require different techniques because their existence are not caused by anthropogenic activities (Li et al., 2015; Cui et al., 2018). To this end, various methods such as Geostatistics, multivariate methods, and Geographic Information System (GIS) mapping have been applied to identify and reveal the distributions of these trace elements (Lark, 2000; Lark & Cullis, 2004; Santra et al., 2012; Antunes & Albuquerque, 2013; Hao et al., 2016; Chakraborty et al., 2017; Boente et al., 2017). Geostatistics has been developed for application in various disciplines, and is represented by techniques including various types of kriging (ordinary/disjunctive/indicator kriging),

78 global/local polynomial interpolation (G/LPI), inverse distance weighting (IDW), nearest
79 neighbour interpolation (NNI), radial basis functions (RBF), sequential Gaussian simulation
80 (SGS), etc. Each of them involves different statistical assumptions. Despite the growing
81 literature of these methods, there are major limitations associated with their application for
82 site investigation in an urban setting. For instance, many development sites in densely-
83 populated cities may span across hundreds of metres to tens of kilometres, within which a
84 large number (i.e., hundreds to thousands) of samples may be needed to provide adequate
85 precision for meaningful geostatistical analyses to aid the site development plans. Meanwhile,
86 the concentrations of trace elements may display three-dimensional spatial variations across
87 the subsurface soil domain, which should be properly accounted for in such analyses. Many
88 previous studies (e.g., Santra et al., 2012; Chakraborty et al., 2012; Chakraborty et al., 2017;
89 Zhang & Yang, 2017) discussed the accuracy of various geostatistical approaches through
90 cross validation measures and indicators such as root-mean-square-error (RMSE) and mean
91 percentage error (MPE). However, the uncertainty associated with As distribution across the
92 site is rarely discussed in detail. While quantification of uncertainty is essential from the
93 project management perspective, such estimates are often difficult to verify.

94 This study extends the integrated framework for spatial variability analyses from our
95 recent studies (Liu et al., 2017; Liu & Leung, 2017), incorporating the restricted maximum
96 likelihood (REML) method with a three-dimensional, anisotropic autocorrelation structure,
97 tailored for analysing the concentrations of trace elements in soils. Effectiveness of the
98 approach is illustrated by the implementation on a major development site in Hong Kong,
99 where borehole sampling of As is performed in multiple stages. The current study articulates
100 the spatial extent of the geogenic As, and proposes a rational approach to quantify the
101 associated uncertainty, hence improving the effectiveness of geoenvironmental sampling
102 strategy for site assessment and remediation.

## 2. Methodology

### 2.1. Soil sampling and analysis

A new development site located in the New Territories in Hong Kong is discussed in this study to illustrate the three-dimensional variations of As concentrations, and how the variability can be characterized by the proposed framework in this study. Two stages of geoenvironmental investigation were performed within this development area, with their key information summarized in Table 1. In order to identify potential contamination at the site, the Stage 1 investigation was performed (Fig. 1a) at an early stage of the project, which included drilling of 35 boreholes, with 388 soil samples retrieved from different depths for the testing of As concentrations (HK CEDD, 2015). The locations of boreholes had been strategically selected for broad coverage across the development area (approximately 1,600 m × 2,700 m on plan), considering both site accessibility and the locations of future structures. Within the development area, there was a smaller site (around 100 m × 200 m) of particular concern in the project. The site was where one of the first structures (Building A shown in Fig. 1) would be constructed, and the As concentrations at this location had to be assessed to formulate appropriate mitigation measures. However, during the Stage 1 investigation, no boreholes had been drilled within this site due to accessibility issues at that stage. Predictions were therefore made by the proposed approach (Section 2.2), utilizing all 388 sample values obtained from the Stage 1, and their corresponding spatial information. Shortly before construction of Building A, the Stage 2 investigation was conducted with 12 additional boreholes and 205 samples across the site (Fig. 1b). These additional samples are treated in this study as verification data for independent assessments of the accuracy of the proposed approach.

During the geoenvironmental investigation, the average depth of boreholes is approximately 20 m (with minimum of 2 m and maximum of 52 m), where soil samples were retrieved every 0.5 – 2 m along the depth of borehole. Soil samples obtained at varying depth of the boreholes were sealed to be air-tight at the site, frozen with ice packs in freezing boxes before and during transportation, and stored at 4$^{o}$C in the dark in cold chamber. Selected samples were freeze-dried within two weeks upon sample delivery and stored at -20$^{o}$C in a refrigerator before spectroscopic analysis (Beiyuan et al., 2017; Li et al., 2017). The As concentrations in the digested samples were determined by using an Inductively Coupled Plasma-Atomic Emission Spectrometry (ICP-AES, Perkin Elmer Optima 3300DV), with the limit of detection of 1 mg/kg. In addition, 20% random replicates and spiked samples were included for quality control of sample analysis, and NIST Reference Soil 2710A (containing 1540 mg/kg As) were used for quality assurance. The As recovery rate was 91-93% after digestion and ICP-OES analysis, thus suggesting good recovery and reliability/reproducibility.

During the Stage 1 investigation, the laboratory tests revealed variations of As concentrations ranging from 1 mg/kg to 1,220 mg/kg dry soil in the 388 samples. The enormous variations in concentrations, with the occasional detection of high levels of As, warranted careful assessments of contamination levels during the planning stages of the development. In particular, it was important to make predictions of As levels at the Building A site in order to formulate the mitigation measures before the construction.

## 2.2. Geostatistical characterization method

Based on the locations of the 388 Stage 1 samples and their corresponding As concentrations, the spatial variability features can be established for better prediction at the Building A site using geostatistical approach. In many previous attempts of geostatistical characterization of trace elements (e.g. Zupan et al. 2000; Burgos et al. 2006; Yang et al., 2009, etc.), a

semivariogram ($\gamma$) is developed for the spatial variable, $\boldsymbol{z}$, that represents concentrations of the trace elements at different locations $\boldsymbol{x}$:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \quad (1)$$

where $z(x_i)$ represents the value of $\boldsymbol{z}$ at location $x_i$ ; $N(h)$ is the number of pairs of samples that are separated at a distance $h$ (in any direction). In many of these studies, however, the spatial trend was not explicitly discussed or considered in the analyses. However, the accuracy of the semivariogram analyses can be affected if the 'deterministic' trend is not properly defined based on the site data. In addition, depending on the site history and geochemical properties of the specific trace element, its spatial variations may display features of three-dimensional anisotropy: the concentrations may be relatively uniform in one direction, but show more variations in other directions in the subsurface domain. These features cannot be rigorously considered by Eq. (1).

This study extends the integrated approach from our recent work (Liu et al., 2017; Liu & Leung, 2017) for three-dimensional spatial variability analyses of As concentrations in soils. In these previous studies, the restricted maximum likelihood (REML) is incorporated with rigorous statistical tests to ensure that the assumptions on normality and stationarity are satisfied, while the optimal polynomial order for the trend can be determined, and with potential outliers in the dataset identified in an integrated framework of spatial data analysis. Liu & Leung (2017) also showed that patterns of spatial variations observed in various directions of natural soil/rock properties can be interpreted together with geological settings of the site. In the current study, the framework is implemented with considerations of three-dimensional spatial correlation structure, where the estimates of As concentration profiles can be sequentially updated when additional data become available from the geoenvironmental investigation. The proposed approach will be compared with traditional geostatistical

175 methods in characterizing the spatial variations of As-containing soils. While the analyses
176 enhance our understanding on the distributions of geogenic As in soils, this study further
177 elaborates on their significance to site-specific uncertainty quantification and management
178 strategies. In the current work, the spatial distributions of As, referred to as $\boldsymbol{z}$, consists of a
179 large-scale 'deterministic' trend across the site, and residual effects that represent the local
180 deviations from the trend:

$$\boldsymbol{z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2}$$

181 where $\mathbf{X}$ is a matrix containing the spatial coordinates of sample points in three dimensions. $\boldsymbol{\beta}$
182 is the vector of trend coefficients, and $\mathbf{X}\boldsymbol{\beta}$ combine to produce the trend. $\boldsymbol{\varepsilon}$ represent the
183 residuals, or deviations from the trend, which is often found to be spatially correlated: the
184 values of $\varepsilon_i$ and $\varepsilon_j$ are often similar if the locations $i$ and $j$ are close to each other, while
185 greater variations are observed for larger separation distances. To simulate such effects, the
186 Gaussian (squared exponential) autocorrelation function is adopted:

$$R(h) = \exp\left[-\left(\frac{h_x^2 + h_y^2}{\theta_1^2}\right) - \frac{h_z^2}{\theta_2^2}\right] \tag{3}$$

187 where $h_x$ , $h_y$ and $h_z$ are the separation distances in $x$, $y$ and $z$ directions, respectively. $\theta_1$
188 and $\theta_2$ are range parameters in the lateral and vertical directions, also known as
189 autocovariance distances. The significance of Eq. (3) lies on its flexibility in modelling
190 anisotropic spatial correlations, as different values of $\theta_1$ and $\theta_2$ may arise as a result of the
191 geogenic origin of the trace element in nature. This has not been considered by previous
192 REML analyses (e.g. Lark & Cullis, 2004), and the importance of such features will be
193 illustrated by the case study presented in later sections. Considering all sampling points
194 within the site, $R(h)$ for different $i$ and $j$ locations combine to form the matrix $\mathbf{R}$, which
195 represents the smooth scale spatial variations in As concentrations. Apart from such spatial
196 variations, there is also contribution from the 'white noise' (nugget) effect in the overall

variance of concentrations. This white noise may be due to handling/measurement errors for the specimens, or other random natural variations that do not correlate with separation distances. The variance contributions from the smooth scale variations and white noise effects are represented by $\sigma_e^2$ and $\sigma_n^2$, respectively. The variance of $\boldsymbol{\varepsilon}$, denoted as $\mathbf{V}$ herein, can then be expressed as:

$$\mathbf{V} = \sigma_e^2\mathbf{R} + \sigma_e^2\mathbf{I} = (\sigma_e^2 + \sigma_n^2)[s\mathbf{R} + (1-s)\mathbf{I}] \quad \text{where } 0 \le s = \frac{\sigma_e^2}{(\sigma_e^2 + \sigma_n^2)} \le 1 \tag{4}$$

and $\mathbf{I}$ is the identity matrix. Therefore, the objectives of the analyses are to obtain the trend coefficients, $\boldsymbol{\beta}$, and parameters that characterize three-dimensional spatial variability of $\boldsymbol{\varepsilon}$ (i.e., $s$, $\theta_1$, $\theta_2$) , based on the measured data $\boldsymbol{z}$. The REML method is adopted for this purpose, and the approach mainly involves finding the set of parameters $\boldsymbol{\Theta} = \{s, \theta_1, \theta_2\}$ that maximizes the following log-likelihood function:

$$L(\boldsymbol{\Theta}|\boldsymbol{y}) = -\frac{n-p}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\log|\mathbf{W}| - \frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{Q}\boldsymbol{y} \tag{5}$$

where $\mathbf{W} = \mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{X}$; $\mathbf{Q} = \mathbf{I} - \mathbf{X}\mathbf{W}^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}$; $\boldsymbol{y} = (\mathbf{I} - \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}})\boldsymbol{z}$, and represents the filtered dataset with the trend components filtered out. The maximization can be treated as an optimization problem, i.e. to obtain the set of $\boldsymbol{\Theta}$ parameters that best matches the observed data $\boldsymbol{z}$. This can be achieved by various optimization techniques; the current study adopts the differential evolution algorithm, and details of this heuristic algorithm can be found in Storn & Price (1997). With $\boldsymbol{\Theta}$ determined, the autocorrelation structure of $\boldsymbol{\varepsilon}$ is well defined. The trend coefficients $\boldsymbol{\beta}$ can then be estimated (i.e. $\widehat{\boldsymbol{\beta}}$) using general least squares method, and predictions at unsampled locations, $\hat{\boldsymbol{z}}$, and the corresponding uncertainty in prediction (prediction variance $\boldsymbol{\sigma}_z^2$), can be estimated based on the best linear unbiased prediction (BLUP) technique (Atkinson et al., 2008; Santra et al., 2012):

$$\hat{\boldsymbol{z}} = \mathbf{X_0}\widehat{\boldsymbol{\beta}} + \mathbf{K}^{\mathrm{T}}\mathbf{V}^{-1}\hat{\boldsymbol{\varepsilon}} \tag{6a}$$

$$\boldsymbol{\sigma}_z^2 = \mathrm{diag}[\mathbf{K}_0 - \mathbf{K}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{K} + \mathbf{M}^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{M}] \quad (6b)$$

where $\mathbf{X}_0$ is the matrix containing spatial coordinates of unsampled (prediction) locations; $\mathbf{K}$ is the covariance matrix between observations and predictions, i.e. $\mathbf{K} = \mathrm{cov}[\boldsymbol{z}(\boldsymbol{x}), \boldsymbol{z}(\boldsymbol{x}_0)]$, $\mathbf{K}_0 = \mathrm{cov}[\boldsymbol{z}(\boldsymbol{x}_0), \boldsymbol{z}(\boldsymbol{x}_0)^{\mathrm{T}}]$ and $\mathbf{M} = \mathbf{X}_0^{\mathrm{T}} - \mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{K}$. In general, the level of uncertainty is low near existing samples, and increases with separation distance away from these sample points. In other words, the uncertainty at unsampled areas varies in three dimensions across the site domain, and is dependent on both spatial locations of all existing samples, the trend of the variations, $\boldsymbol{\beta}$, and autocorrelation features represented by $\boldsymbol{\Theta}$.

## 3. Results and Discussion

### 3.1. Spatial variability analysis based on Stage 1 samples

The As concentrations obtained from the 388 samples of Stage 1 investigation are first analyzed using the proposed REML approach. The sample values are log-transformed, as positive-value variables in geochemistry often follow the lognormal distribution. Figure 2 shows the statistical distributions of the As data and the log-transformed values, which further validates the assumption of lognormal distribution. Regarding the deterministic trend function, a quadratic structure for lateral directions and linear trend for vertical direction are adopted to represent the large-scale variations across the site ($\mathbf{X}\boldsymbol{\beta}$). Using the observed data $\boldsymbol{z}$ (vector with 388 components), the $\boldsymbol{\Theta}$ parameters that maximizes the log-likelihood function (Eq. 5) are found to be $s = 0.57$, $\theta_1 = 130.9$ m and $\theta_2 = 4.2$ m. These are the parameters that characterize the spatial variability of As concentrations, and the large difference between $\theta_1$ and $\theta_2$ illustrates the strong anisotropy effects regarding the influence range the correlation: the variations are more abrupt along the vertical direction (small $\theta_2$), compared to a more gradual transition along the lateral directions. Figure 3a shows the corresponding autocorrelation functions ($R$), decoupled in the vertical and lateral directions. As a

comparison with more conventional geostatistical approaches, a separate analysis is conducted using the method of moments (MoM), where the autocovariance is defined by:

$$C(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [z(x_i) - t(x_i)][z(x_i + h) - t(x_i + h)] \quad (7)$$

with $\boldsymbol{t}$ representing the trend component vector (i.e. $\boldsymbol{t} = \mathbf{X}\boldsymbol{\beta}$). The similarity between Eqs. (1) and (7) should be noted, as the MoM approach is closely related to the determination of semivariogram in geostatistics (Reilly & Gelman, 2007; Oliver & Webster, 2014). Together with the assumption of isotropic spatial feature ($\theta_1 = \theta_2$), there is no need to separate $h_x$, $h_y$ and $h_z$, and the corresponding MoM approach is similar to the geostatistical approach adopted by Zupan et al. (2000), Burgos et al. (2006) and Yang et al. (2009). The spatial correlation structure obtained by this MoM analysis is shown in Fig. 3b for comparison. The accuracies of the two approaches will be compared and discussed in more detail in the next section.

The proposed REML approach allows the prediction of As concentrations at unsampled locations ($\hat{\boldsymbol{z}}$), and quantification of the associated uncertainty through the prediction variance ($\boldsymbol{\sigma}_z^2$), across the three-dimensional subsurface domain of the entire development site. For example, the ground surface at the Building A site is at +15 mPD (mPD: metres above Hong Kong Principal Datum, which is 1.230 m below Mean Sea Level), and Fig. 4 shows the predictions of As concentrations and the prediction variance at three different depths (+10 mPD, 0 mPD, -10 mPD).

Based on the raw data of measured As concentrations from Stage 1 investigation, the sample variance is 1.99 (in log-space, since the concentrations are assumed to follow lognormal distribution). Figure 4b shows that the value of prediction variance across the Building A site is relatively uniform, with a magnitude around 1.6, which illustrates two interesting aspects of the spatial correlation analyses. Firstly, the difference between raw data

variance (≈1.99) and prediction variance of residuals (≈1.6) arises from the removal of the deterministic trend ($\mathbf{X}\boldsymbol{\beta}$), or the de-trending process, in the proposed approach. Meanwhile, the almost uniform prediction variance across Building A site means that the uncertainty is not heavily influenced (reduced) by any 'nearby' existing samples. In fact, the closest borehole from Stage 1 investigation is more than 100 m away from the site, while the horizontal range of influence ($\theta_1$) for each sample is only around 131 m according to the analyses. In other words, considering the spatial features of As in this region, the Stage 1 boreholes and samples are not located close enough to the Building A site to aid the assessment of contamination levels there. These spatial correlation features also carry important implications to the predictive capabilities of various approaches, as will be illustrated in the next section.

### 3.2. Predictions and validations with Stage 2 investigation data

While a number of previous studies described the geostatistical characterization of trace elements such as As, there is limited discussion on the accuracy of the approaches, particularly within the context of predictive capabilities verified through subsequent geo-environmental investigation. In this study, the two aspects of predictions, namely the best estimates ($\hat{\boldsymbol{z}}$) and uncertainty quantification ($\boldsymbol{\sigma}_z^2$) by Eqs. (6a) and (6b), are compared with the additional sample data retrieved from Stage 2 investigation at the Building A site. The advantages of the proposed approach over 'conventional' techniques of MoM are also illustrated through the verification process.

Comparisons between predicted and measured As concentrations can be made relatively easily for the Stage 2 sample data. On the contrary, it is more difficult to verify the 'uncertainty' estimates, since the prediction variance varies across the site, but there is only one measured value, and therefore one prediction error, at any location in the subsurface

289 domain. An indirect approach to verify the $\boldsymbol{\sigma}_z^2$ estimates is to normalize the prediction error
290 ($\widetilde{\varepsilon}_i = \widehat{z}_i - z_i$) for each location $i$ by the value of $\sigma_{z,i}$ evaluated at that location. The
291 normalized errors associated with all Stage 2 samples can then be compared with the standard
292 normal distribution. This comparison is based on the proposition that with a proper
293 characterization of uncertainty, the prediction errors across the domain should follow a
294 normal distribution, with the variance denoted by the corresponding $\sigma_{z,i}^2$ at the location.

295 Figure 5 shows the histograms of normalized prediction errors (in log-space) using both
296 the proposed REML approach (Eqs. 2-6), and the conventional MoM approach (Eq. 7). In
297 Fig. 5a, the predictions are made solely based on the Stage 1 data (388 samples), and both
298 approaches generally underestimated the As concentrations at the Building A site, with mean
299 error smaller than zero. As mentioned earlier, considering the horizontal spatial range ($\theta_1$) of
300 131 m, and that the closest borehole is more than 100 m away, it is not surprising that both
301 approaches cannot accurately predict the As concentrations at Building A, as the 'knowledge'
302 of Stage 1 samples do not reach far enough to provide useful information at that site.

303 In the next modelling scenario, data from two of the twelve Stage 2 boreholes is
304 incorporated into the geostatistical analyses, and predictions are then made at the locations of
305 the remaining ten boreholes. Since these boreholes are within the Building A site with plan
306 dimensions in the same order as $\theta_1$, incorporating such additional information is expected to
307 enhance the accuracy of the approach. In each analysis, both the $\hat{\boldsymbol{z}}$ and $\boldsymbol{\sigma}_z^2$ estimates are
308 updated using sample information from the two additional boreholes. Also, for a
309 comprehensive comparison, all possible combinations of locations of the two 'known'
310 boreholes are simulated, and the corresponding normalized prediction errors at the ten
311 'unknown' borehole locations are summarized in the histograms in Fig. 5b. With the
312 knowledge from two additional boreholes, predictions by the proposed REML approach are
313 improved, both in terms of the best estimates (mean error closer to zero) and with error

distribution closely resembling the standard normal distribution. In contrast, results by the MoM are not significantly improved, mainly due to the fact that the three-dimensional nature of spatial variability cannot be properly reflected (Fig. 3) and assimilated into the prediction model. The similar exercise is then repeated for two other scenarios: (1) incorporating sample data from four boreholes and making predictions at the remaining eight borehole locations (Fig. 5c); and (2) incorporating sample data from six boreholes and making predictions at the remaining six locations (Fig. 5d). In both cases, the improvement by the REML approach is more substantial than the conventional MoM approach: the mean prediction error is close to zero using the REML approach, and the distributions of normalized error roughly follow the standard normal distribution. Applying this concept to realistic project situations, it would be desirable to continuously refine the predictions at unsampled locations, based on new sample data as the investigation programme progresses.

As a further comparison between the proposed REML approach and conventional MoM analyses, Fig. 6 shows the results of one simulated scenario incorporating data from two boreholes of Stage 2 investigation, to predict the As concentrations at the other ten borehole locations. In this example, 41% of the measured values fall within ±30% of the REML predictions, which is a significantly higher percentage than using the conventional method. Similarly, Fig. 7 shows the simulated scenario with knowledge of six boreholes and prediction of the other six, in which case almost 60% of the measured values fall within ±30% of the REML predictions. In fact, the substantial variability in the natural occurrence of As makes it unlikely for any modelling approach to precisely pinpoint their concentrations at all locations, and the occasional existence of extreme values does comply with classical statistical theories. Nonetheless, the improved predictive power of the proposed approach, and the capabilities to rationally quantify and represent the uncertainty (Fig. 5), mean that it

can become a useful tool in assessing the risks associated with the occurrence of As, or potentially other contaminants or trace elements, at a development site.

Figures 8 and 9 show the contours of predicted As concentrations, and the associated prediction variance (in log-space) after drilling and sampling at the six and twelve Stage 2 boreholes at the Building A site. It should be noted that the colour scales in Figs. 3, 8 and 9 are consistent with each other, which facilitate comparison between the estimates as more information becomes available. Also, combining the best estimates with prediction variance (after back-transformation to original space), the range of predictions with different confidence levels (e.g., mean ± standard deviation) can be established, which provide an alternative means to demonstrate the risk levels associated with As occurrence across the site.

### 3.3. Effectiveness of sampling in reducing uncertainty of As concentrations

The proposed approach also provides information regarding the effectiveness of additional samples in reducing uncertainty. From a risk management perspective, this would be a useful indicator on the required number of samples to adequately characterize As distributions across the site. Following previous discussions, the uncertainty indicator, i.e. the prediction variance, varies spatially in three dimensions across the entire subsurface domain at all unsampled locations, and can only be shown in a series of contours. In this section, a simpler definition is adopted, and denoted as the prediction variance reduction factor (PVRF). PVRF is calculated by first obtaining the average prediction variance across the site (e.g., Building A), and then evaluating the change of this average variance upon drilling and sampling at a new borehole location. This quantifies the significance of each borehole (and its samples) in reducing the overall uncertainty. To quantify the effectiveness of a particular borehole, $\mathrm{PVRF}_i$ is defined by:

$$\mathrm{PVRF}_i(\%) = \frac{\mathrm{PV}_{i-1} - \mathrm{PV}_i}{\mathrm{PV}_{i-1}} \times 100\% \quad (8)$$

where $\mathrm{PV}_i$ is the average prediction variance across the site after drilling borehole $i$. Therefore, PVRF may be interpreted as the percentage reduction in overall uncertainty associated with each successive borehole. A low level of $\mathrm{PVRF}_i$ refers to a low impact from borehole $i$ in further reducing uncertainty in As variations. By evaluating the value of PVRF associated with each borehole, its effectiveness on the overall uncertainty can be quantified. It is then possible to make informed decisions regarding the appropriate number of boreholes and samples, to balance the need to characterize site uncertainty and the time and resources associated with drilling, sampling and testing at each additional borehole location.

Figure 10 shows the PVRF values obtained for the twelve boreholes at the Building A site. The data point represents the average value obtained by simulation of 1000 realizations, with the error bars illustrating the standard deviations. These realizations are necessary since the calculation of PVRF is a sequential process, and depends on the order of drilling/sampling at the twelve locations. For example, sampling at the first borehole will reduce the overall uncertainty by about 16% on average, while the subsequent boreholes will become less effective as more and more information becomes available within Building A site. In particular, after drilling and sampling at the sixth borehole location, the uncertainty reduction by each subsequent borehole will be generally less than 5%. Figures 5d and 7 also show that with the sample data from six boreholes, the As distributions and uncertainty can be predicted with reasonable accuracy using the proposed approach. Such information can become useful when site-specific geoenvironmental sampling strategies are devised.

## 4. Conclusions

This study outlines the geostatistical approach using REML, which has been extended for consideration of three-dimensional spatial autocovariance structure. The proposed approach

386 is applied to investigate the spatial distributions of geogenic As concentrations at a
387 development site in Hong Kong, which involved 388 samples and 205 samples retrieved
388 during two separate stages of geoenvironmental investigation. Data from the multi-stage
389 investigation are utilized in the verification of the proposed approach, both regarding the best
390 estimates of As levels and the associated uncertainty at previously unsampled areas. The
391 proposed approach is shown to produce more accurate predictions than conventional
392 geostatistical approaches, even with a relatively small dataset in a large development area.
393 The importance of site-specific characterization of spatial variability is also highlighted, as
394 the accuracy of predictions depend heavily on correlation parameters, geometry of the site
395 and locations of existing sample information. The proposed approach also addresses another
396 key consideration in development sites, which is the quantification of uncertainty at
397 unsampled locations. The estimated prediction variance can be condensed into a simple
398 indicator defined as PVRF in this study, which quantifies the effectiveness of
399 geoenvironmental sampling in reducing the uncertainty levels of As concentrations across the
400 site. This helps to supplement the conventional approach that rely on qualitative expert
401 opinions, as it provides quantitative indication to support the decision-making process
402 associated with the necessity and strategies of future sampling, considering the tolerable risk
403 levels and financial setup of the project.
404
405 **Acknowledgement**
406 The study is financially supported by the Hong Kong Research Grants Council (25201214
407 and 15222115) and the Civil Engineering and Development Department of the Hong Kong
408 SAR Government.

**References**

Antunes I.M. & Albuquerque M.T. (2013). Using indicator kriging for the evaluation of arsenic potential contamination in an abandoned mining area (Portugal). Science of the Total Environment, 442, 545-552.

Atkinson, P. M., Pardo-Iguzquiza, E., & Chica-Olmo, M. (2008). Downscaling cokriging for super-resolution mapping of continua in remotely sensed images. IEEE Transactions on Geoscience and Remote Sensing, 46, 573-580.

Bednářová, Z., Kalina, J., Hájek, O., Sáňka, M., & Komprdová, K. (2016). Spatial distribution and risk assessment of metals in agricultural soils. Geoderma, 284, 113-121.

Beiyuan, J., Li, J. S., Tsang, D. C. W., Wang, L., Poon, C. S., Li, X. D., & Fendorf, S. (2017). Fate of arsenic before and after chemical-enhanced washing of an arsenic-containing soil in Hong Kong. Science of the Total Environment, 599, 679-688.

Bell, R.M., Gildon, A. & Parry, G.D.R. (1983). Sampling strategy and data interpretation for site investigation of contaminated land. In Reclamation of former iron and steel works sites, Durham and Cumbria County Councils, Durham.

Boente, C., Matanzas, N., García-González, N., Rodríguez-Valdés, E. & Gallego, J.R. (2017). Trace elements of concern affecting urban agriculture in industrialized areas: A multivariate approach. Chemosphere 183, 546-556.

Bolan, N., Kunhikrishnan, A., Thangarajan, R., Kumpiene, J., Park, J., Makino, T., Kirkham, M. B., Scheckel, K. (2014). Remediation of heavy metal(loid)s contaminated soils–to mobilize or to immobilize? Journal of Hazardous Materials, 266, 141-166.

Bundschuh, J., Bhattacharya, P., Nath, B., Naidu, P., Ng, J., Guilherme L.R., Ma L.Q., Kim K.W. & Jean J.S. (2013). Arsenic ecotoxicology: The interface between geosphere, hydrosphere and biosphere. Journal of Hazardous Materials, 262, 883-886.

Burgos, P., Madejón, E., Pérez-de-Mora, A., & Cabrera, F. (2006). Spatial variability of the chemical characteristics of a trace-element-contaminated soil before and after remediation. Geoderma, 130, 157-175.

Cao, M., Ye, Y., Chen, J., & Lu, X. (2016). Remediation of arsenic contaminated soil by coupling oxalate washing with subsequent ZVI/Air treatment. Chemosphere, 144, 1313-1318.

Chakraborty, S., Man, T., Paulette, L., Deb, S., Li, B., Weindorf, D. C., & Frazier, M. (2017). Rapid assessment of smelter/mining soil contamination via portable X-ray fluorescence spectrometry and indicator kriging. Geoderma, 306, 108-119.

Chakraborty, S., Weindorf, D. C., Zhu, Y., Li, B., Morgan, C.L.S., Ge, Y. & Galbraith, J. (2012). Assessing spatial variability of soil petroleum contamination using visible near-infrared diffuse reflectance spectroscopy. Journal of Environmental Monitoring, 14, 2886-2892.

Cui, J. L., Zhao, Y. P., Li, J. S., Beiyuan, J. Z., Tsang, D. C. W., Poon, C. S., & Li, X. D. (2018). Speciation, mobilization, and bioaccessibility of arsenic in geogenic soil profile from Hong Kong. Environmental Pollution, 232, 375-384.

Ehlert, K., Mikutta, C., & Kretzschmar, R. (2016). Effects of manganese oxide on arsenic reduction and leaching from contaminated floodplain Soil. Environmental Science & Technology, 50, 9251-9261.

Fendorf, S., Michael, H.A. & van Geen, A. (2010). Spatial and temporal variations of groundwater arsenic in South and Southeast Asia. Science, 328,1123–1127.

Gallego J.R., Rodriguez-Valdés, E., Esquinas, N., Fernández-Braña, A. & Afif, E. (2016). Insights into a 20-ha multi-contaminated brownfield megasite: An environmental forensics approach. Science of the Total Environment, 563-564, 683-692.

Gilbert, R.O. (1987). Statistical methods for environmental pollution monitoring. 320 pp. New York: Van Nostrand Reinhold.

González-Fernández, B., Rodríguez-Valdés, E., Boente, C., Menéndez-Casares, E., Fernández-Braña, A. & Gallego, J.R. (2018). Long-term ongoing impact of arsenic contamination on the environmental compartments of a former mining-metallurgy area. Science of the Total Environment, 610-611, 820-830.

Hao, L., Tian, M., Zhao, X., Zhao, Y., Lu, J., & Bai, R. (2016). Spatial distribution and sources of trace elements in surface soils, Changchun, China: Insights from stochastic models and geostatistical analyses. Geoderma, 273, 54-63.

HK CEDD (2015). Collaboration research study on sustainable land decontamination methods for application in land development projects. Final Report produced by the research team from the Hong Kong Polytechnic University.

HK EPD (2011). Practice guide for investigation and remediation of contaminated land, Environmental Protection Department, Hong Kong SAR Government.

Horta, A., Malone, B., Stockmann, U., Minasny, B., Bishop, T. F. A., McBratney, A. B., & Pozza, L. (2015). Potential of integrated field spectroscopy and spatial analysis for enhanced assessment of soil contamination: a prospective review. Geoderma, 241, 180-209.

Lark, R.M. (2000). Estimating variograms of soil properties by the method-of-moments and maximum likelihood. European Journal of Soil Science, 51, 717-728.

Lark, R.M. & Cullis, B.R. (2004). Model-based analysis using REML for inference from systematically sampled data on soil. European Journal of Soil Science, 55, 799-813.

Li, J. S., Beiyuan, J., Tsang, D. C. W., Wang, L., Poon, C. S., Li, X. D., & Fendorf, S. (2017). Arsenic-containing soil from geogenic source in Hong Kong: Leaching characteristics and stabilization/solidification. Chemosphere, 182, 31-39.

Li, Y., Duan, Z., Liu, G., Kalla, P., Scheidt, D., & Cai, Y. (2015). Evaluation of the possible sources and controlling factors of toxic metals/metalloids in the Florida Everglades and their potential risk of exposure. Environmental Science & Technology, 49, 9714-9723.

Liu, W. F., Leung, Y. F., & Lo, M. K. (2017). Integrated framework for characterization of spatial variability of geological profiles. Canadian Geotechnical Journal, 54, 47-58.

Liu, W. F., Leung, Y. F., (2017). Characterising three-dimensional anisotropic spatial correlation of soil properties through in situ test results. Géotechnique, doi:10.1680/jgeot.16.P.33.

Oliver, M. A., & Webster, R. (2014). A tutorial guide to geostatistics: Computing and modelling variograms and kriging. Catena, 113, 56-69.

Pan, H., Lu, X., & Lei, K. (2017). A comprehensive analysis of heavy metals in urban road dust of Xi'an, China: Contamination, source apportionment and spatial distribution. Science of The Total Environment, 609, 1361-1369.

Plessl, C., Jandrisits, P., Krachler, R., Keppler, B. K., & Jirsa, F. (2017). Heavy metals in the mallard Anas platyrhynchos from eastern Austria. Science of the Total Environment, 580, 670-676.

Rahman, M. S., Clark, M. W., Yee, L. H., Comarmond, M. J., Payne, T. E., Kappen, P., & Mokhber-Shahin, L. (2017). Arsenic solid-phase speciation and reversible binding in long-term contaminated soils. Chemosphere, 168, 1324-1336.

Reilly, C., & Gelman, A. (2007). Weighted classical variogram estimation for data with clustering. Technometrics, 49, 184-194.

Rieuwerts J.S., Mighanetara, K., Braungardt, C.B., Rollinson, G.K., Pirrie, D. & Azizi, F. (2014). Geochemistry and mineralogy of arsenic in mine wastes and stream sediments in a historic metal mining area in the UK. Science of the Total Environment, 472, 226-234.

Santra, P., Das, B. S., & Chakravarty, D. (2012). Spatial prediction of soil properties in a watershed scale through maximum likelihood approach. Environmental Earth Sciences, 65, 2051-2061.

Sandhi, A., Greger, M., Landberg, T., Jacks, G., & Bhattacharya, P. (2017). Arsenic concentrations in local aromatic and high-yielding hybrid rice cultivars and the potential health risk: a study in an arsenic hotspot. Environmental Monitoring and Assessment, 189, 184.

Storn, R., & Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization, 11, 341-359.

Tóth, G., Hermann, T., Szatmári, G., & Pásztor, L. (2016). Maps of heavy metals in the soils of the European Union and proposed priority areas for detailed assessment. Science of the Total Environment, 565, 1054-1062.

Tsang, D.C.W., Yip, A.C.K. (2014). Comparing chemical-enhanced washing and waste-based stabilisation approach for soil remediation. Journal of Soils and Sediments, 14, 936-947.

Tsang, D.C.W., Yip, A.C.K., Olds, W.E., Weber, P.A. (2014). Arsenic and copper stabilisation in a contaminated soil by coal fly ash and green waste compost. Environmental Science and Pollution Research, 21, 10194-10204.

Wan, X., Lei, M., & Chen, T. (2016). Cost–benefit calculation of phytoremediation technology for heavy-metal-contaminated soil. Science of the Total Environment, 563, 796-802.

Waterhouse, P.S. (1980). The sampling and analysis of polluted sites. In Reclamation of contaminated land, Society of Chemical Industry, London.

Wcisło, E., Bronder, J., Bubak, A., Rodríguez-Valdés, E. & Gallego, J.L.R. (2016). Human health risk assessment in restoring safe and productive use of abandoned contaminated sites. Environment International, 94, 436-448.

Wu, C., Huang, L., Xue, S. G., Pan, W. S., Zou, Q., Hartley, W., & Wong, M. H. (2017). Oxic and anoxic conditions affect arsenic (As) accumulation and arsenite transporter expression in rice. Chemosphere, 168, 969-975.

Xia, X., Yang, Z., Li, G., Yu, T., Hou, Q., & Mutelo, A. M. (2017). Practicability of monitoring soil Cd, Hg, and Pb pollution based on a geochemical survey in China. Chemosphere, 172, 217-224.

Yang N., Winkel L.H. & Johannesson K.H. (2014). Predicting geogenic arsenic contamination in shallow groundwater of south Louisiana, United States. Environmental Science and Technology, 48, 5660-6.

Yang, Z., Wang, Y., Shen, Z., Niu, J., & Tang, Z. (2009). Distribution and speciation of heavy metals in sediments from the mainstream, tributaries, and lakes of the Yangtze River catchment of Wuhan, China. Journal of Hazardous Materials, 166, 1186-1194.

Yang, K., Kim, B. C., Nam, K., & Choi, Y. (2017). The effect of arsenic chemical form and mixing regime on arsenic mass transfer from soil to magnetite. Environmental Science and Pollution Research, 24, 8479-8488.

Yoon, Y., Kim, S., Chae, Y., Jeong, S. W., & An, Y. J. (2016). Evaluation of bioavailable arsenic and remediation performance using a whole-cell bioreporter. Science of the Total Environment, 547, 125-131.

Zhang, B. & Yang Y. (2017). Spatiotemporal modeling and prediction of soil heavy metals based on spatiotemporal cokriging. Scientific Reports, 7, 16750.

Zupan, M., Einax, J. W., Kraft, J., Lobnik, F., & Hudnik, V. (2000). Chemometric characterization of soil and plant pollution: Part 1: Multivariate data analysis and geostatistical determination of relationship and spatial structure of inorganic contaminants in soil. Environmental Science and Pollution Research, 7, 89-96.

**Table**

Table 1 Site area and sampling information for Stages 1 and 2 investigations

| | Stage 1 (Fig. 1a) | Stage 2 (Fig. 1b) |
|---|---|---|
| Target area | Entire development area (1,600 m × 2,700 m) | Building A site (100 m × 200 m) |
| No. of boreholes | 35 | 12 |
| No. of samples | 388 | 205 |
| Average spacing of boreholes | 870 m | 57 m |
| Vertical sampling interval | 0.5 – 2 m | 0.5 – 2 m |

**Figure**

**Click here to download Figure: STOTEN-D-18-00068_Figues_R1.docx**

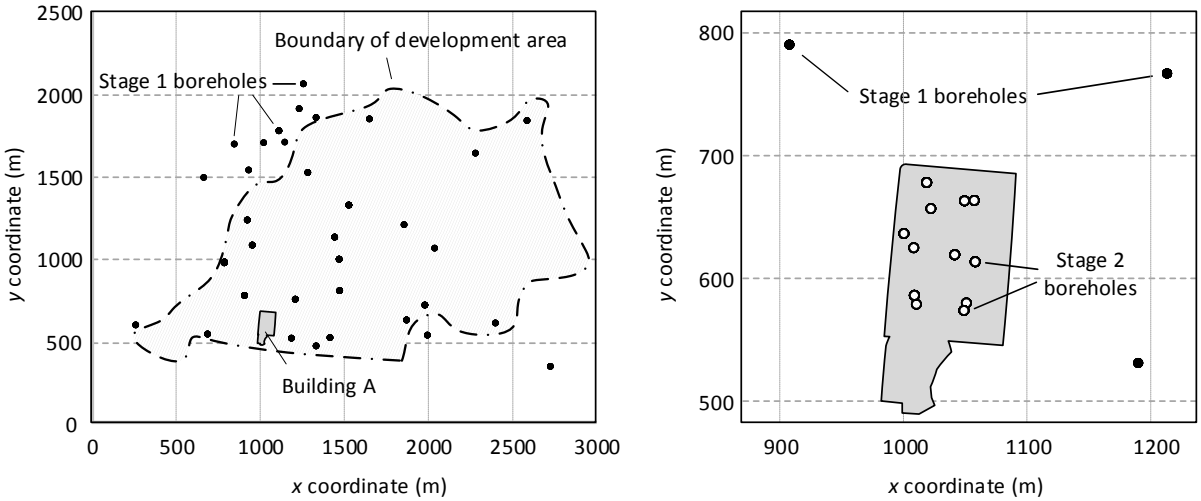


Figure 1 Site area and sampling locations for Stages 1 and 2 investigations

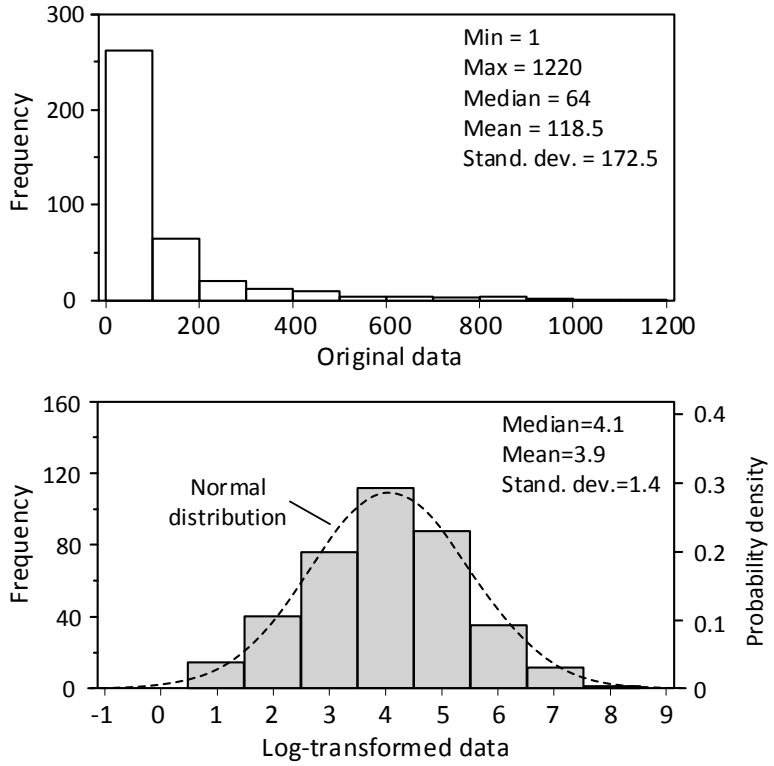


Figure 2 Statistical distribution of As concentrations from Stage 1 investigation

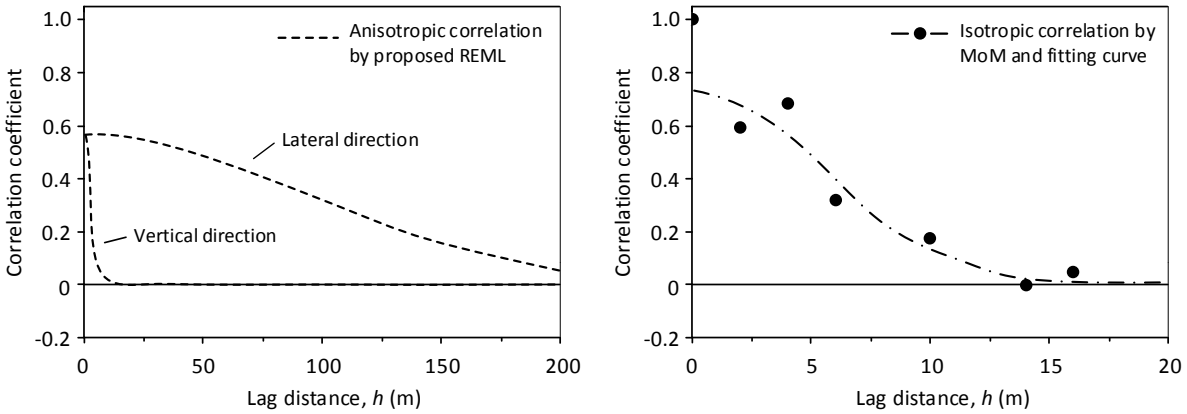


Figure 3 Autocorrelation structure estimated by proposed approach and conventional MoM

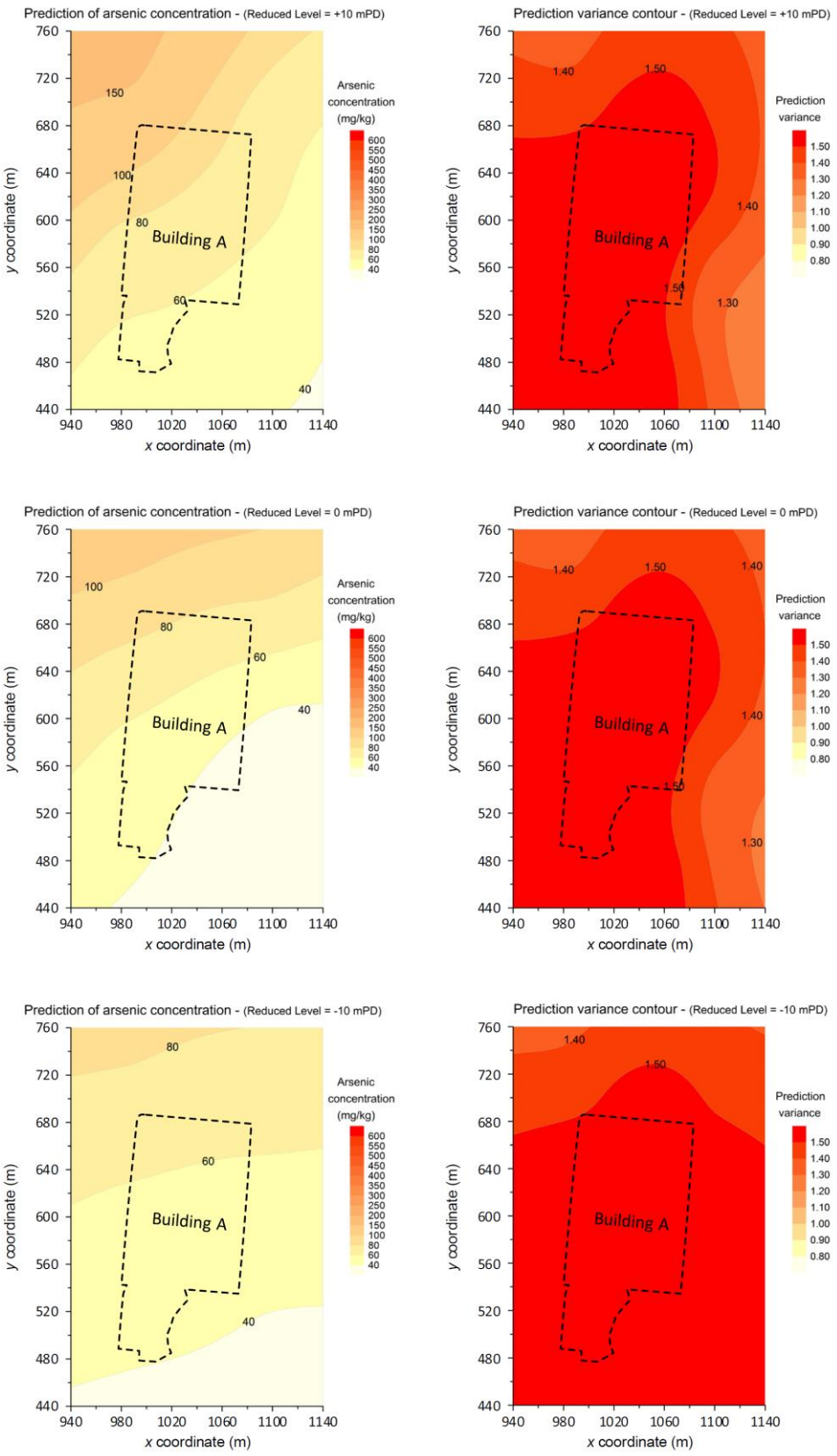


Figure 4 Predicted As concentrations (left) and prediction variance (right) using Stage 1 samples

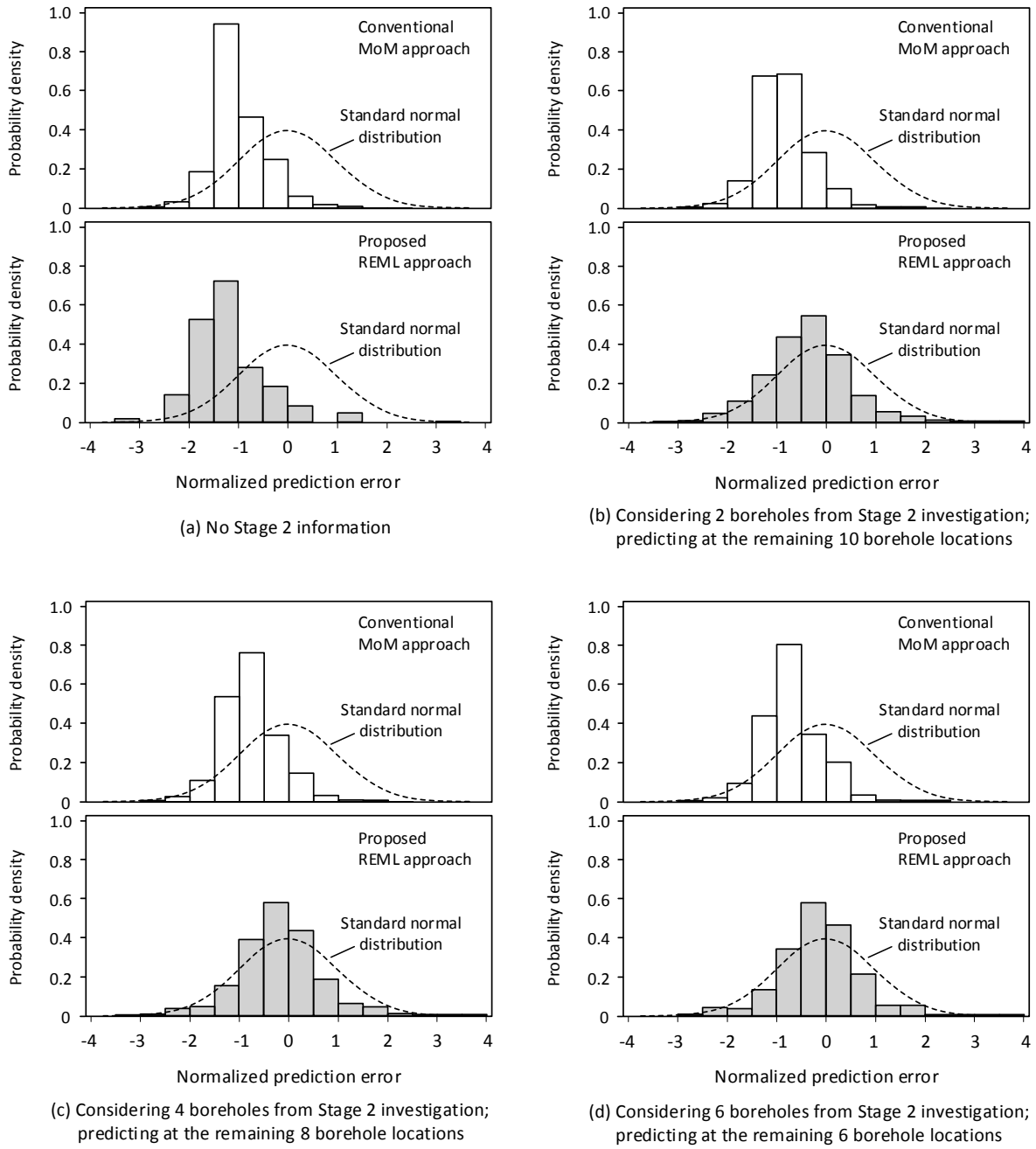


Figure 5 Histograms of normalized prediction errors: (a) with no stage 2 sample data; (b) incorporating two Stage 2 borehole data; (c) incorporating four Stage 2 borehole data; (d) incorporating six Stage 2 borehole data

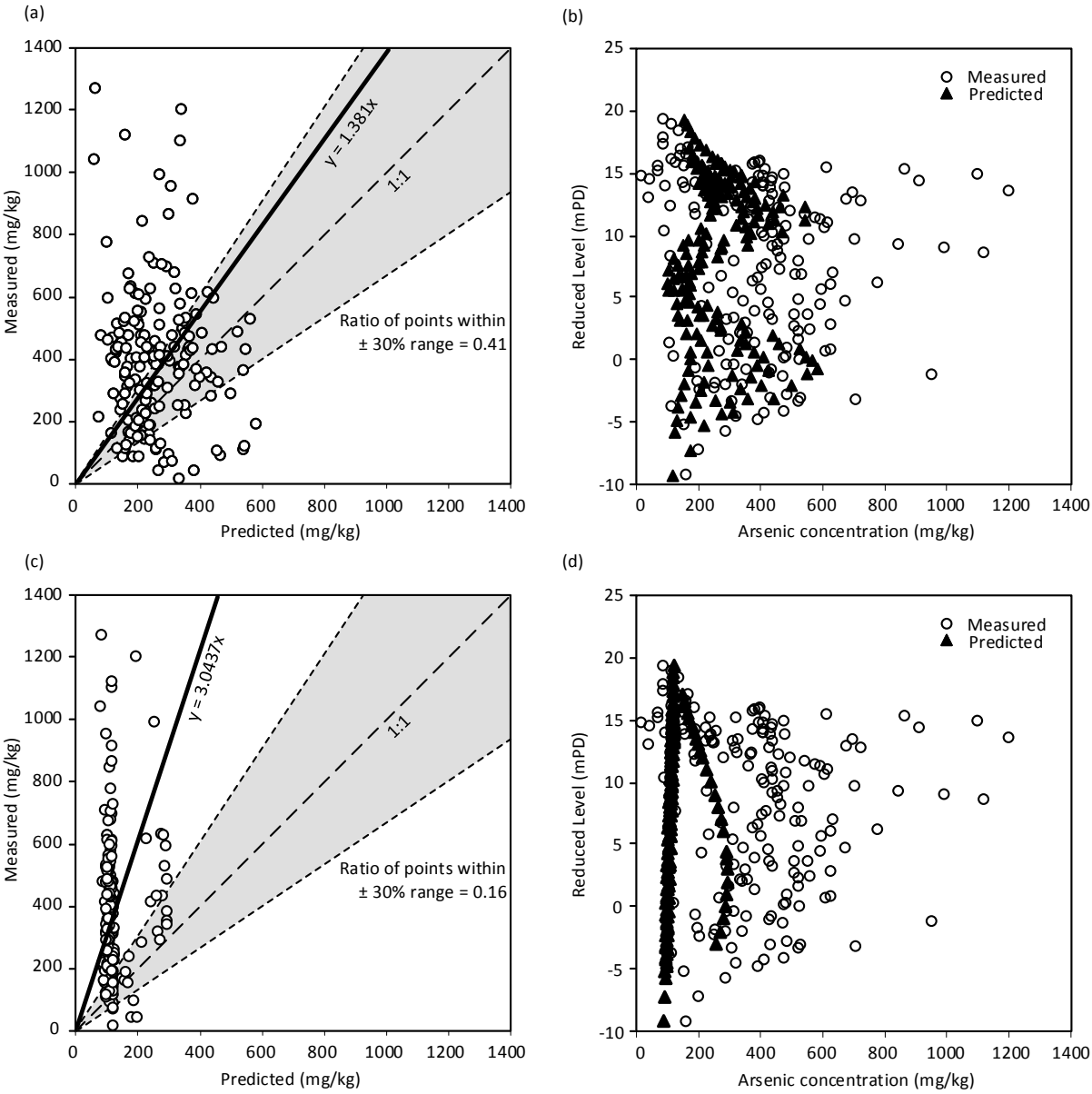


Figure 6 Comparisons between measured As concentration with (a, b) REML and (c, d) MoM predictions incorporating two Stage 2 borehole data

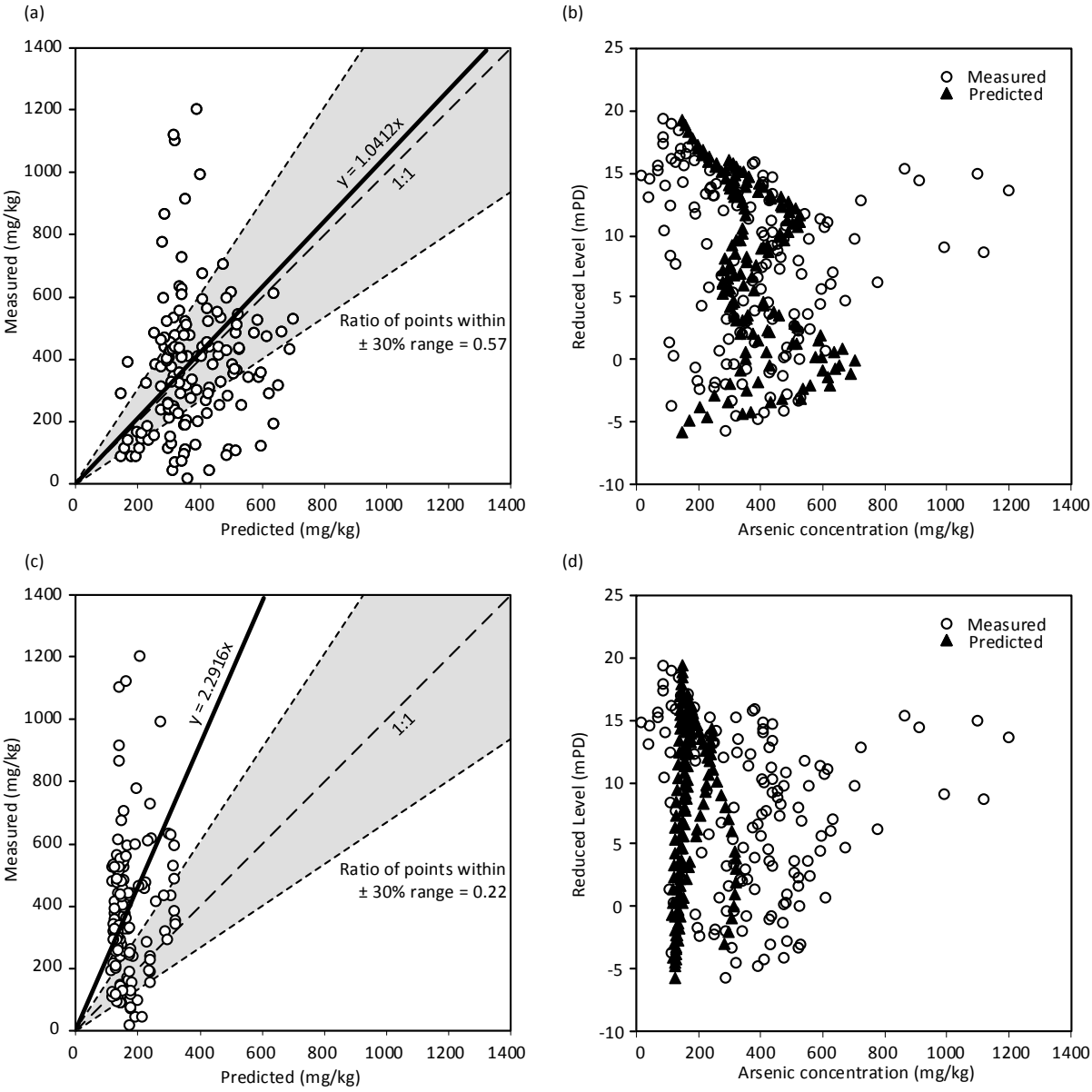


Figure 7 Comparisons between measured As concentration with (a, b) REML and (c, d) MoM predictions, incorporating six Stage 2 borehole data

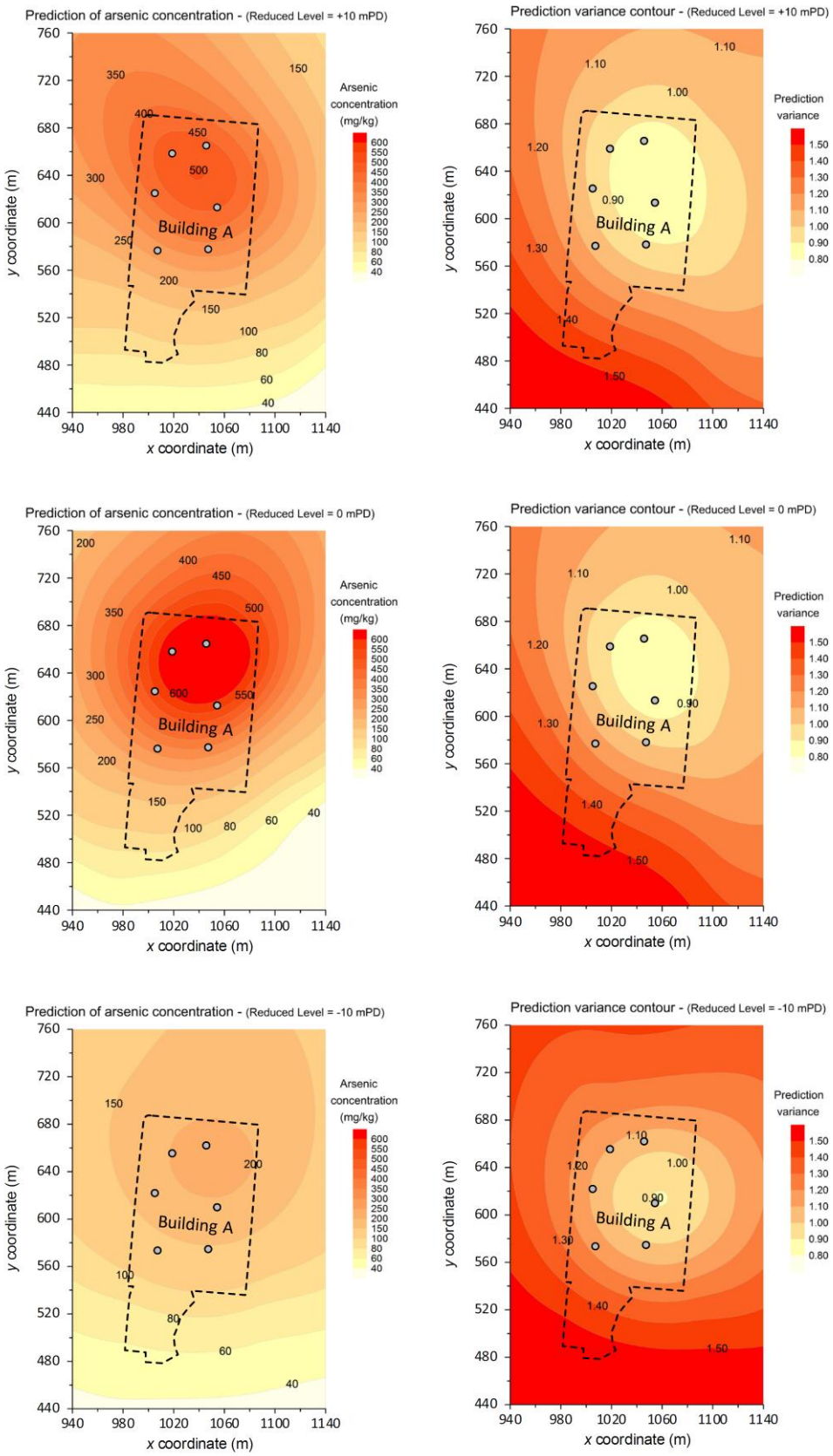


Figure 8 Predicted As concentrations (left) and prediction variance (right) after drilling and sampling at six borehole locations (grey circles) in Stage 2

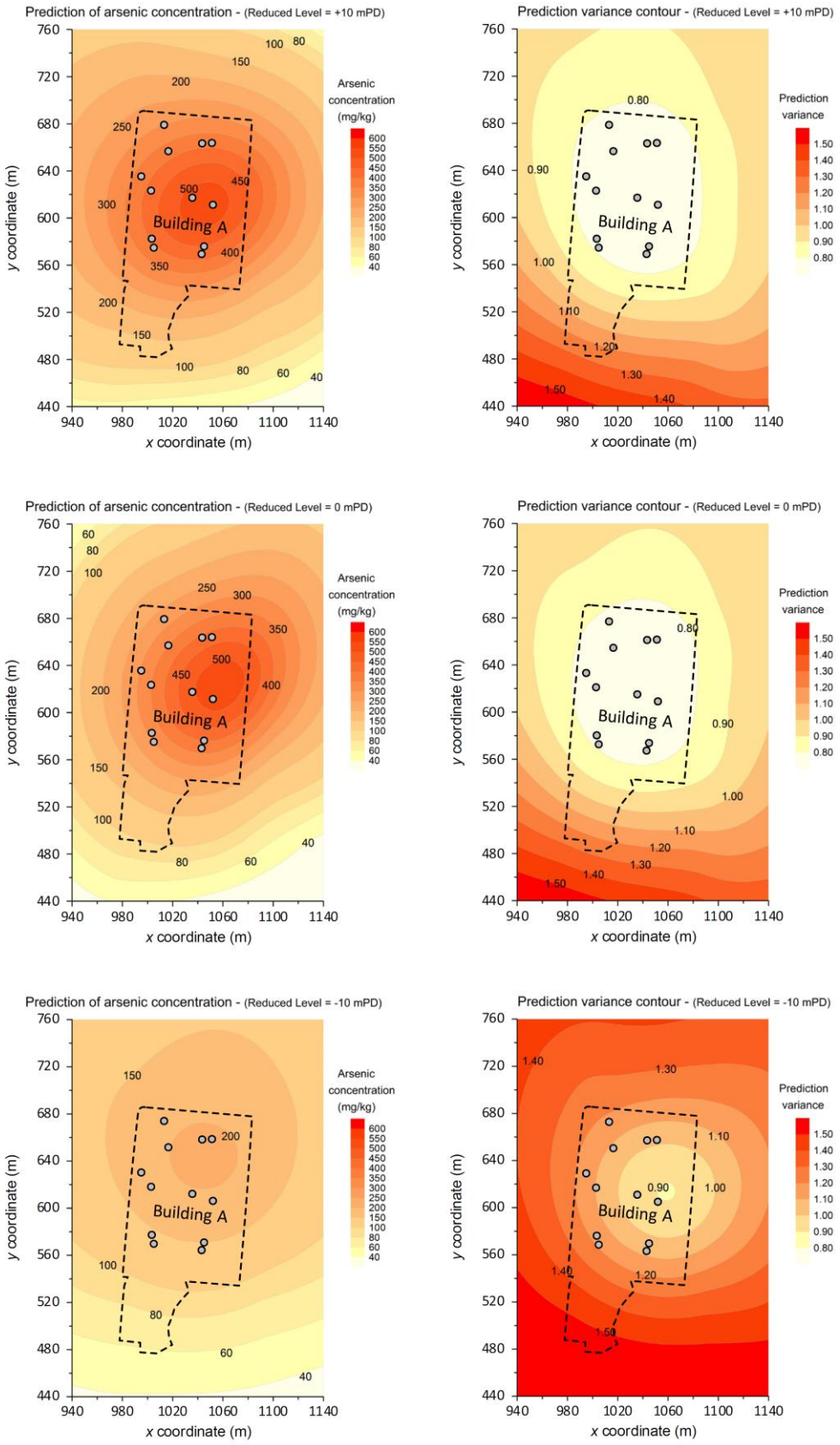


Figure 9 Predicted As concentrations (left) and prediction variance (right) after drilling and sampling at twelve borehole locations (grey circles) in Stage 2

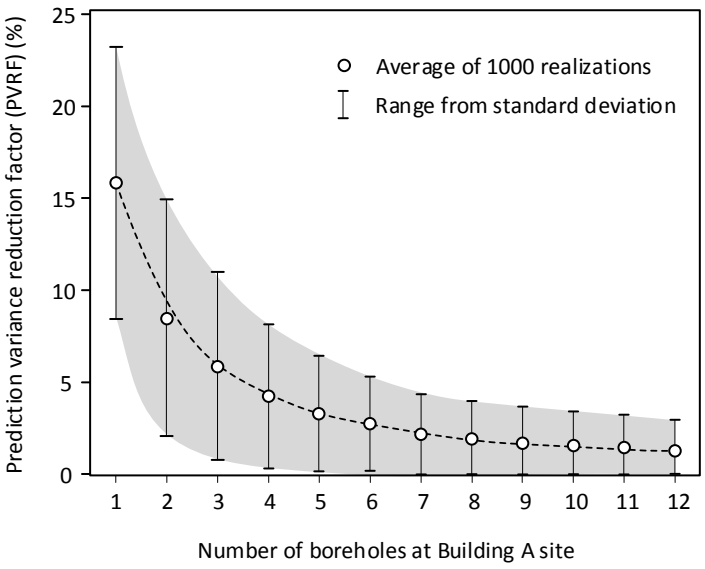


Figure 10 Changes of PVRF with sampling at successive borehole locations