Stoch Environ Res Risk Assess (2013) 27:1735–1740
DOI 10.1007/s00477-013-0710-5

ORIGINAL PAPER

A field sampling strategy for semivariogram inference of fractures in rock outcrops

Ryan M. Pollyea · Jerry P. Fairley ·
Robert K. Podgorney · Travis L. McLing

Abstract The stochastic continuum (SC) representation is one common approach for simulating the effects of fracture heterogeneity in groundwater flow and transport models. These SC reservoir models are generally developed using geostatistical methods (e.g., kriging or sequential simulation) that rely on the model semivariogram to describe the spatial variability of each continuum. Although a number of strategies for sampling spatial distributions have been published in the literature, little attention has been paid to the optimization of sampling in resource- or access-limited environments. Here we present a strategy for estimating the minimum sample spacing needed to define the spatial distribution of fractures on a vertical outcrop of basalt, located in the Box Canyon, east Snake River Plain, Idaho. We used fracture maps of similar basalts from the published literature to test experimentally the effects of different sample spacings on the resulting semivariogram model. Our final field sampling strategy was based on the lowest sample density that reproduced the semivariogram of the exhaustively sampled fracture map. Application of the derived sampling strategy to an outcrop in our field area gave excellent results, and illustrates the utility of this type of sample optimization. The method will work for developing a sampling plan for any intensive property, provided prior information for a similar domain is available; for example, fracture maps or ortho-rectified photographs from analogous rock types could be used to plan for sampling of a fractured rock outcrop.

Keywords Field sampling · Fractured rock · Stochastic continuum · Box Canyon · Idaho

1 Introduction

Accounting for the influence of fracture heterogeneity in numerical flow and transport models remains a challenge (Berkowitz 2002), yet despite the promulgation of increasingly complex methods, stochastic continuum (SC) representations continue to provide robust support for simulating regional scale fracture heterogeneity (Neuman 2005). The SC model invokes the equivalent porous media assumption (Long et al. 1982), which, in a fractured rock reservoir model, is represented by spatially distributed subdomains (continua) with unique hydrogeologic properties, e.g., grid cells defined by fracture or matrix controlled parameters (Finsterle 2000).

SC reservoir models are commonly developed with geostatistical methods (e.g., kriging or sequential simulation), which are appealing in their flexible accommodation of source data; for example, the spatial variability of a chosen continuum is described by the model semivariogram and may be inferred from analog outcrop data, discrete borehole data, or, preferably, some combination of both (Deutsch 2002, pp. 9–14). Despite this flexibility,

R. M. Pollyea (✉)
Department of Geology and Environmental Geosciences,
Northern Illinois University, DeKalb, IL 60115, USA
e-mail: rpollyea@niu.edu

J. P. Fairley
Department of Geological Sciences, University of Idaho,
Moscow, ID 83844, USA

R. K. Podgorney
Energy Resource Recovery and Sustainability, Idaho National
Laboratory, Idaho Falls, ID 83415, USA

T. L. McLing
Center for Advanced Energy Studies, Idaho National Laboratory,
Idaho Falls, ID 83415, USA

however, semivariogram inference is data intensive, necessitating areally extensive field sampling campaigns that allow numerous two-point differences to be evaluated over variable distances and directions. In fact, it has been suggested that a minimum of 150–200 samples are required for sampling isotropic processes and substantially more may be required in the presence of anisotropy (Webster and Oliver 1992). In addition to data quantity, the geometric distribution of sample locations must be considered; previous investigations have shown regular hexagonal and triangular sampling networks are generally preferred on the basis of minimum kriging errors (Olea 1984; Yfantis et al. 1987; Christakos and Olea 1992). Iterative strategies have also been proposed for optimal sampling locations based on minimizing point distances for *a priori* selected semivariogram lag intervals (Russo 1984; Warrick and Myers 1987), or, when faced with physical boundaries, minimizing a mean shortest distance criteria by spatial simulated annealing (Van Groenigen and Stein 1998).

Despite the theoretical appeal of triangular, hexagonal, and iterative field sampling strategies for semivariogram inference, these methods assume easily accessible terrain suitable for locating a sample network by gps or rod-and-level surveys. Although reasonable for many spatial sampling applications, e.g. siting monitoring wells, this assumption is tenuous when the field area is hazardous or difficult-to-access. Furthermore, in most real-life situations data collection is constrained by available resources (i.e., time and money). As a result, an optimum sampling strategy should minimize costs, while ensuring the required spatial correlation information is secured. In this study, we develop a field sampling strategy aimed at minimizing the cost of collecting data for semivariogram inference in a vertical study area. In particular, we are interested in a model of spatial variability for fractured regions within an analog outcrop that may be assumed to exhibit orders of magnitude greater permeability than the host rock. We derive our sample spacing and density requirements by estimating experimental and model semivariograms using data available in the literature [as suggested by Lark (2002) and implemented by Caeiro et al. (2003)] and this information is translated into a line survey sampling technique (ODOT 2010) modified for optimal spatial sampling.

2 Site description

The proposed field sampling strategy was developed for characterizing fracture distributions of a basalt outcrop in the Box Canyon, which is located on the east Snake River Plain (ESRP) in southeast Idaho (Fig. 1). The ESRP is a layered assemblage of low-volume vesicular basalt flows that exhibit a columnar joint structure with column normal (horizontal) and column bounding (vertical) fractures being the most prevalent fracture types (Schaefer and Kattenhorn 2004). Fracture densities are highest in the upper colonnades and lateral flow margins due to elevated thermal gradients from convective heat transfer to the atmosphere during cooling (Lore et al. 2001; Schaefer and Kattenhorn 2004). In plan view, the edges of column bounding fractures define a polygonal network of columns that are cross-cut by column normal fractures in the upper colonnades, see, e.g., Lore et al. (2001, Fig. 3). To support our studies of carbon dioxide sequestration in reactive basalts, we elected to represent this morphologically controlled and interconnected fracture geometry using a SC model.

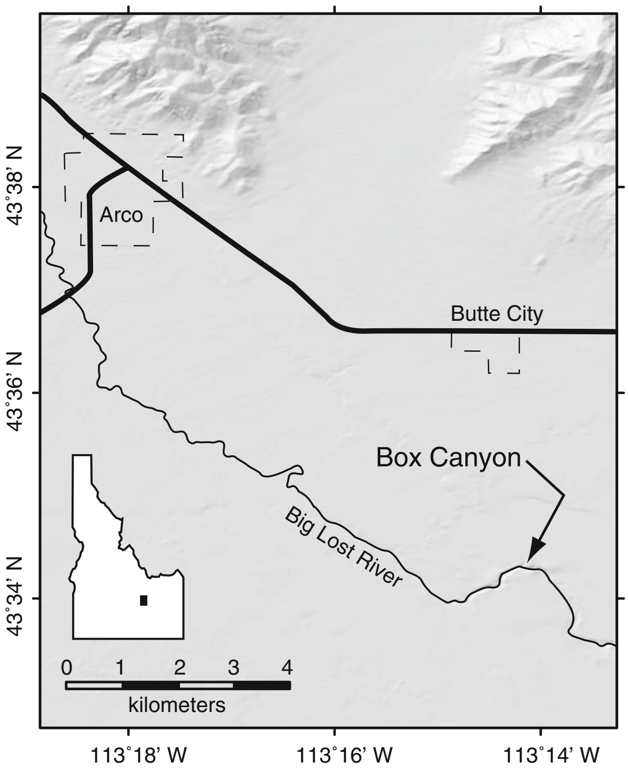

Fig. 1 Location map of the Box Canyon field site

3 Sampling plan design

A fracture map of a Box Canyon basalt exposure, taken from the literature (Schaefer and Kattenhorn 2004, Fig. 3a), is enlarged and overlayed with a locally-referenced 1 m regular grid containing two randomly located 0.5 m subgrids (Fig. 2a). The nested subgrids are included to allow additional short range comparisons in the event that small scale processes contribute significantly to the sample variance and to provide support against estimation error that may arise if the sample grid coincides with periodic behavior of the fracture distribution. Sampling the gridded fracture map is done by inspecting a scaled 0.25 m circle are around each grid node and recording the presence or absence of a fracture; in total, 550 grid nodes were sampled.

The isotropic experimental semivariogram for this data set are evaluated using the program `gamv` (Deutsch and Journel 1998). The semivariogram is defined as half the average square difference between data separated by a distance class (or lag); for a continuous variable, this is given by:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (z_i - z_{i+h})^2 \tag{1}$$

where, z_i is a datum at location i, z_{i+h} is a datum at distance h from location i, and $N(h)$ is the number of data pairs in the lag class (Deutsch and Journel 1998). In the present case, z is a binary categorical variable such that

$$z = \begin{cases} 1 & \text{fracture} \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

so that direct application of the Eq. 1 is equivalent to an indicator semivariogram (Deutsch and Journel 1998, p. 46). The correlation range parameter is inferred by fitting the positive-definite (i.e., allowable) exponential model (Goovaerts 1997) to the experimental semivariogram using a range parameter of 3.5 m and nugget contribution of 65 % (Fig. 2a, inset).

Field sampling was conceived as a modified line survey in which field personnel would rappel the rock face to obtain samples. In a traditional rock slope line survey, sampling lines are distributed across the outcrop and data are obtained for every fracture intersecting the sample line (ODOT 2010). Because this method only samples fractures, data is not provided about the spatial relationship of fractured and unfractured regions. To obtain data of this type, the traditional line survey method is modified to obtain samples on a regular interval. With this approach, a sampling pattern (i.e., lines of descent and vertical sampling interval) is required that minimizes sampling frequency while reproducing, as closely as possible, the semivariogram correlation model derived from the fully gridded fracture map. We provisionally select a "nested" distribution of seven sample lines consisting of a center line flanked on each side by sample lines spaced 0.5-, 2-, and 5-m from center. This staggered seven line cluster allows a greater proportion of two-point comparisons within the known semivariogram correlation range, while accommodating longer range comparisons over the distal

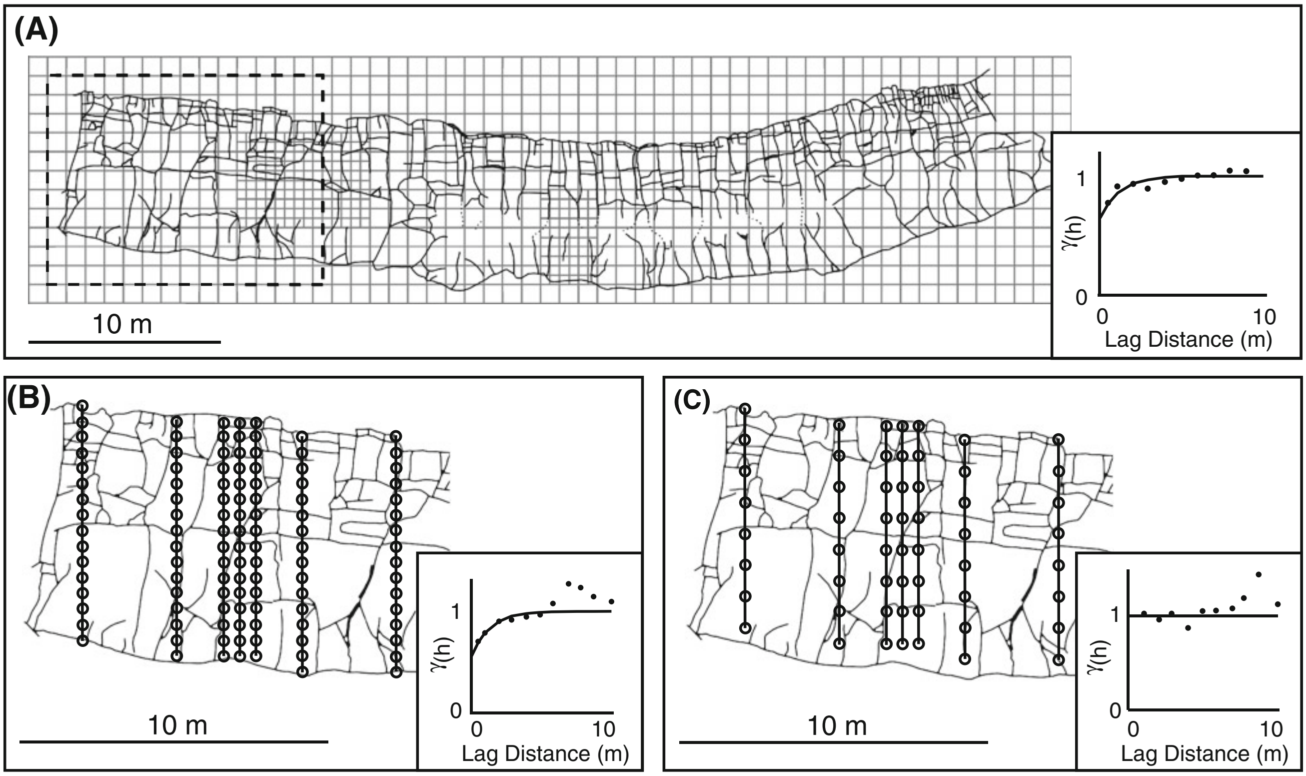

Fig. 2 The Box Canyon fracture map used for this analysis (Schaefer and Kattenhorn 2004, Fig. 3a). **a** 1 m regular grid with two randomly placed 0.5 m subgrids. The *dashed box* indicates blow up section for (**b**) and (**c**). **b** Vertical sample lines spaced 0.5-, 2-, and 5-m from the center line with samples collected every 0.5-m. **c** Same as (**b**) with samples collected every 1-m. Normalized semivariograms for each sampling are *inset*; points represent experimental semivariogram values and *solid line* represents model semivariogram

lines of the array. The effects of using different sample spacings in the vertical direction are tested on the fracture map by evaluating semivariograms sampled at 0.5- and 1.0-m intervals at three randomly located line clusters; the resulting sampling patterns for one line cluster location are shown in Fig. 2a, b, respectively.

Samples were obtained and semivariograms were evaluated for each line cluster by the same methods described for the fully gridded map. In all three samplings the line clusters with 0.5-m vertical spacing returned semivariogram models closely resembling that of the fully gridded map (see, e.g., Fig. 2b, inset). In contrast, the semivariograms for line clusters sampled vertically every 1 m returned pure nugget effect models suggesting inadequate sample density for two-point semivariogram analysis (see, e.g., Fig. 2c, inset). The final field sampling strategy incorporates as many seven line clusters as materials (i.e., rope anchors) would allow, while sampling vertically every 0.5 m.

4 Discussion

Field sampling was performed over a two week period on a basalt outcrop within the Box Canyon that was selected on the basis of exposure and accessibility. The selected outcrop is ~ 12 m vertical by 90 m horizontal. This exposure permits access to two adjacent basalt flows separated by a rubbly flow margin, as well as the upper colonnades of several underlying basalt flows (Fig. 3a). Sufficient materials, i.e., rope anchors, were available to place five seven-line clusters and one additional three-line cluster with 0.5-m line spacing (Fig. 3a, white lines). The survey lines were spatially referenced on the canyon rim with a plane table and alidade and sample locations on the rock face were referenced with a weighted surveyor's tape suspended from the canyon rim. Samples were collected by rappelling the exposure (Fig. 3b), inspecting the rock behind a 0.25-m ring (Fig. 3c), and recording the rock "texture" as fractured, massive, or rubble and the "morphology" as upper colonnade, lower colonnade, entablature, or flow boundary. In total, 755 samples were obtained.

A binary categorical transform is applied to the dataset on the basis of the assumed higher hydraulic conductivity of densely fractured upper colonnades and rubble-dominated lateral basalt flow boundaries, and the less fractured, lower conductivity materials:

$$i(\mathbf{u}) = \begin{cases} 1 & \left(\begin{array}{c} \text{flow boundary or} \\ \text{rubble, or} \\ \text{upper colonnade and} \\ \text{fractured} \end{array} \right) \\ 0 & \text{otherwise} \end{cases} . \qquad (3)$$

In Eq. 3, i represents the binary value at each location $\mathbf{u}$ assigned as a function of the "texture" and "morphology" field data. This binary transform is an explicit assumption that translates independent morphology and fracture data into spatially dependent reservoir permeability structures, i.e., this binary transform is an expression of the generating process (thermal contraction resulting in mode I fracture propagation) that is responsible for the densely fractured upper colonnades and rubbly flow margins. This assumption is our decision of stationarity (Deutsch and Journel 1998, p. 12) and forms the basis of our two-continuum model for fractured ESRP basalt.

The high permeability continuum was shown to exhibit spatial autocorrelation using the isotropic (direction independent) experimental semivariogram; however, a cross validation test returned spatially correlated kriging residuals, suggesting the presence of anisotropy in the data

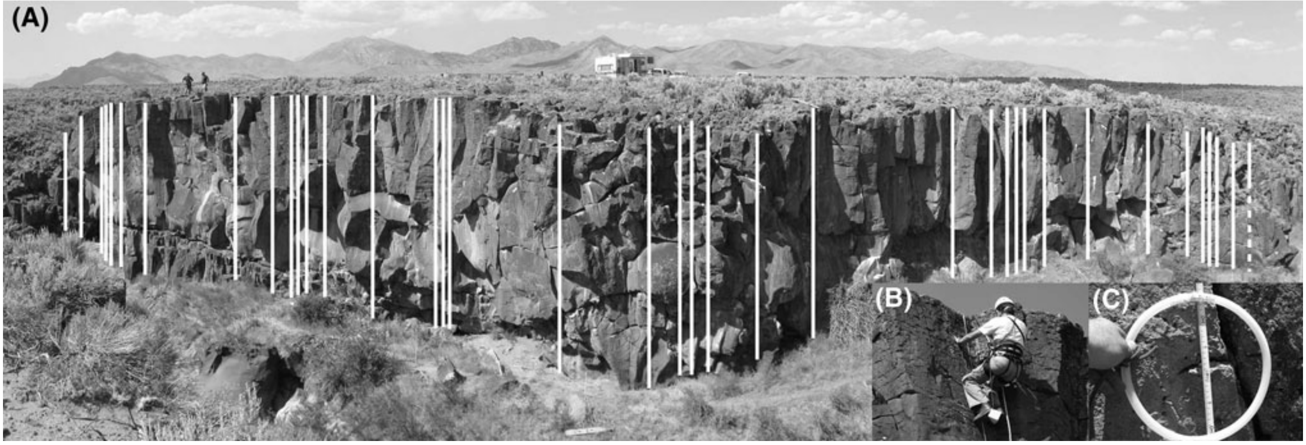

Fig. 3 **a** Image of the Box Canyon outcrop sampled for this investigation. *White lines* represent sampling lines. *Dashed line* indicates obstructed view of the actual sample line. Note that image is a composite of several images resulting in distorted scale. **b** Sample locations were accessed while descending the rock face. **c** Samples were obtained by inspecting the rock behind a 0.25-m ring

(Kitanidis 1997, p. 90). The cross validation test was satisfied with an anisotropic semivariogram model (Fig. 4) in which the experimental semivariogram was computed by limiting the search with angular tolerances of 22.5° parallel to the horizontal and vertical outcrop dimensions. Although the geometry of the sampled outcrop permits only a two dimensional correlation model, this modified line survey method may be extended into three dimensions in situations where (1) outcrop geometry is more accommodating, (2) bare rock is exposed on the flat lying surface above the canyon rim, or (3) sufficient borehole data is available for inferring a long range correlation model. More generally, this sample optimization strategy may be used in any application where prior information is available to test different sampling strategies ahead of field mobilization; potential examples that may be of interest to the reader include, sample plan design for site characterization efforts related to (1) unexploded ordinance remediation efforts [see e.g., Ostrouchov et al. (2009)] and (2) monitoring well network planning at contaminated groundwater sites [see e.g., Bierkens (2006)]. In addition, and while outside the scope of this study, it should be mentioned that semivariogram models of the type developed here are the foundation upon which geostatistical reservoir models are built; common methods for geostatistical reservoir modeling include ordinary and simple kriging, sequential Gaussian simulation, sequential indicator simulation, and spatial simulated anealing (Deutsch 2002).

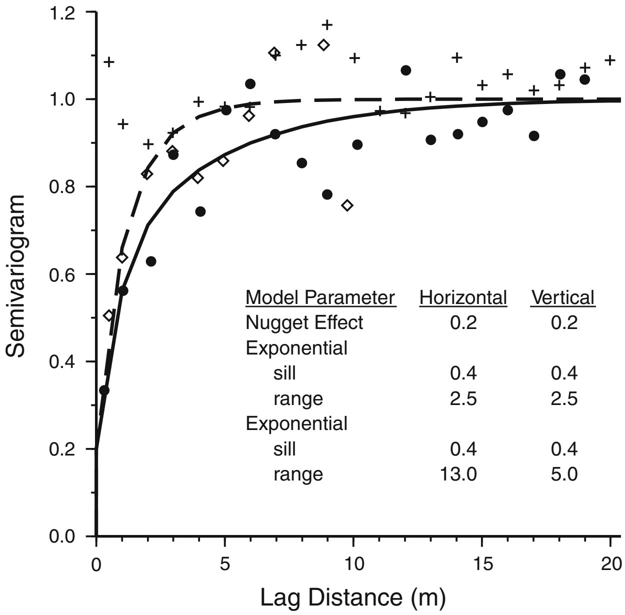

Fig. 4 Anisotropic semivariogram of fracture locations at the Box Canyon outcrop (Fig. 3). The direction of maximum spatial correlation occurs parallel to the outcrop face, denoted here as the horizontal direction. The direction of minimum spatial correlation is vertical. Experimental semivariograms are denoted for the horizontal and vertical directions with *points* and *diamonds*, respectively. Model semivariograms are denoted for the horizontal and vertical directions as *solid* and *dashed lines*, respectively. Model parameters are listed. Units for the range parameter are in meters. *Plus symbols* represent the experimental variogram for cross-validation kriging errors

5 Conclusion

While field sampling is rarely, if ever, the end result of hydrogeologic investigation, it is axiomatic that field data are the fundamental link between the physical and simulated environments. In this paper we developed a field sampling strategy for inferring the spatial correlation of permeability structures in a vertical outcrop of ESRP basalt. The difficulty presented by the vertical study area was addressed by seeking information about the semivariogram range parameter of fracture distributions in similar basalts. Prior to field mobilization, the semivariogram range parameter was estimated by sampling a gridded fracture map from a previous ESRP investigation. Field sampling was then simulated using a modified line survey technique on progressively fewer subsets of the fracture map until a minimum sample distribution was found with a correlation model similar to that of the full distribution. This sampling pattern was executed on a basalt exposure at the Box Canyon in southeast Idaho and the data analysis resulted in a cross-validated, anisotropic semivariogram model that is compatible with a variety of different reservoir simulation techniques. The method of sample plan described here is one possible approach for sampling difficult-to-access environments or in the face of resource constraints, and may be applied to any sampling campaign in which a proxy can be sampled in advance of field mobilization.

Acknowledgments The authors thank Professor Simon Kattenhorn for generously providing a digital image of the Box Canyon fracture map used in this analysis. The authors also thank Jennifer Hinds for her assistance preparing graphics and Scott Brinton for his tireless help in the field. This manuscript was greatly improved by suggestions from two anonymous reviewers. Financial support was provided by the Laboratory Directed Research and Development Program of Idaho National Laboratory, administered by the Center for Advanced Energy Studies, under the US Department of Energy Idaho Operations Office Contract DE-AC07-05ID14517.

References

Berkowitz B (2002) Characterizing flow and transport in fractured geological media: a reivew. Adv Water Resour 25:861–884

Bierkens M (2006) Designing a monitoring network for detecting groundwater pollution with stochastic simulation and a cost model. Stoch Environ Res Risk Assess doi:10.1007/s00477-005-0025-2

Caeiro S, Painho M, Goovaerts P, Costa H, Sousa S (2003) Spatial sampling design for sediment quality assessment in estuaries. Environ Model Softw 18:853–859

Christakos G, Olea RA (1992) Sampling design for spatially distributed hydrogeologic and environmental processes. Adv Water Res 15:219–237

Deutsch CV (2002) Geostatistical reservoir nodeling. Oxford University Press, New York

Deutsch CV, Journel AG (1998) GSLIB: geostatistical software library and user's guide, 2nd edn. Oxford University Press, New York

Finsterle S (2000) Using the continuum approach to model unsaturated flow in fractured rock. Water Resour Res 36(8):2055–2066

Goovaerts P (1997) Geostatistics for natural resource evaluation. Oxford University Press, New York

Kitanidis PK (1997) Introduction to geostatistics: applications in hydrogeology. Cambridge University Press, New York

Lark R (2002) Optimized spatial sampling of soil for estimation of the variogram by maximum likelihood. Geoderma 10(1–2):49–80

Long J, Remer J, Wilson C, Witherspoon P (1982) Porous media equivalents for networks of discontinuous fractures. Water Resour Res 18(3):645–658

Lore J, Aydin A, Goodson K (2001) A deterministic methodology for prediction of fracture distribution in basaltic flows. J Geophys Res 106(B4):6447–6489

Neuman SP (2005) Trends, prospects and challenges in quantifying flow and transport through fractured rocks. Hydrogeol J 13:124–147

ODOT (2010) Rock slope design manual. Ohio Department of Transportation

Olea RA (1984) Sampling design optimization for spatial functions. Math Geol 16 (4):369–392

Ostrouchov G, Doll W, Beard L, Morros M, Wolf D (2009) Multiscale structure of UXO site characterization: spatial estimation and uncertainty. Stoch Environ Res Risk Assess 23:215–225 doi:10.1007/s00447-007-0212-4

Russo D (1984) Design of an optimal sampling network for estimating the variogram. Soil Sci Soc Am J 48:708–716

Schaefer CJ, Kattenhorn SA (2004) Kattenhorn characterization and evolution of fractures in low-volume pahoehoe lava flows, eastern Snake River Plain, Idaho. Geol Soc Am Bull 116(3/4):322–336

Van Groenigen J, Stein A (1998) Constrained optimization of spatial sampling using continuous simulated annealing. J Environ Qual 27:1078–1086

Warrick A, Myers D (1987) Optimizaiton of sampling locations for variogram calculations. Water Resour Res 23(3): 496–500

Webster R, Oliver M (1992) Sample adequately to estimate variograms of soil properties. J Soil Sci 43:177–192

Yfantis EA, Flatman GT, Behar JV (1987) Efficiency of kriging estimation for square triangular, and hexagonal grids. Math Geol 19(3):183–205