

# Predictive Geological Mapping Using Closed-Form Non-stationary Covariance Functions with Locally Varying Anisotropy: Case Study at El Teniente Mine (Chile)

Francky Fouedjio<sup>1,3</sup> and Serge Séguret<sup>2</sup>

Received 11 January 2016; accepted 24 February 2016  
Published online: 12 March 2016

This paper is concerned with the problem of predicting the surface elevation of the Braden breccia pipe at the El Teniente mine in Chile. This mine is one of the world's largest and most complex porphyry-copper ore systems. As the pipe surface constitutes the limit of the deposit and the mining operation, predicting it accurately is important. The problem is tackled by applying a geostatistical approach based on closed-form non-stationary covariance functions with locally varying anisotropy. This approach relies on the mild assumption of local stationarity and involves a kernel-based experimental local variogram a weighted local least-squares method for the inference of local covariance parameters and a kernel smoothing technique for knitting the local covariance parameters together for kriging purpose. According to the results, this non-stationary geostatistical method outperforms the traditional stationary geostatistical method in terms of prediction and prediction uncertainty accuracies.

**KEY WORDS:** Mining, Geostatistics, Non-stationarity, Locally varying anisotropy, Covariance function, Kriging, Simulation.

## INTRODUCTION

One of the most important goals in geostatistical applications is the prediction of a physical quantity of interest over the entire study domain from measurements in some locations. This problem is fundamentally based on modeling and estimation of the spatial dependence structure of data. Describing the spatial dependence structure of data is usually carried out using statistical tools such as the variogram or covariogram, which is calculated over the entire study domain under the stationarity assumption. However, this assumption that states

that the spatial dependence structure of data remains constant over the entire study domain is driven more by mathematical convenience than by reality. In practice, it can be doubtful due to some local influences or localized effects, which can be reflected by computing local stationary variograms whose characteristics may vary spatially. The use of a stationary geostatistical approach in such cases is not suitable because it could produce less accurate predictions, including an incorrect assessment of the prediction error.

Various approaches have been proposed over the years to deal with non-stationarity of the spatial dependence structure of data (Guttorp and Schmidt 2013; Sampson et al. 2001; Guttorp and Sampson 1994). One of the most interesting is the classes of explicit non-stationary covariance functions with locally varying anisotropy developed by Paciorek and Schervish (2006) and Stein (2005). The covariance

<sup>1</sup>CSIRO Mineral Resources, 26 Dick Perry Ave, Kensington, WA 6151, Australia.

<sup>2</sup>MINES ParisTech, PSL Research University, Centre de Géosciences, 77300, Fontainebleau, France.

<sup>3</sup>To whom correspondence should be addressed; e-mail: francky.fouedjio@csiro.au

functions belonging to these classes are locally stationary and their parameters can vary spatially yielding local variances, ranges, geometric anisotropies, and other. The estimation of these spatially varying parameters is a challenging problem. Paciorek and Schervish (2006) suggest to estimate these parameters as in the moving windows non-stationary geostatistical approach based on the variogram (Haas 1990a, b; Harris et al. 2010; Magneron et al. 2010; Machuca-Mory and Deutsch 2013). Anderes and Stein (2011) propose a weighted local likelihood approach where the influence of faraway observations is smoothly down-weighted. Fouedjio et al. (2016) propose a distribution-free approach involving a local stationary variogram kernel estimator, a weighted local least-squares method in combination with a kernel smoothing method.

In this article, we address the problem of predicting the surface elevation of the breccia pipe called Braden at the El Teniente mine in Chile. As described by Skewes et al. (2002), Maksaev et al. (2004), and Spencer et al. (2015), the supergiant El Teniente deposit is one of the world's largest and most complex porphyry-copper ore bodies, containing an estimated premining resource of approximately 95 million metric tons Cu and 2.5 million metric tons Mo. The mine is located approximately 70 km southeast of Santiago on the western margin of the Andean Cordillera and within the confines of the central Chilean porphyry Cu belt. The center of the deposit is composed of a late-stage diatreme known as the Braden pipe, which is 1200 m in diameter at the surface and close to 600 m at a depth of 1800 m. The pipe is poorly mineralized and surrounded by different kinds of mineralized geological units. Knowing the exact location of the pipe surface is important because it constitutes the internal limit of the deposit and the mining operation. To do this, a geostatistical approach based on closed-form non-stationary covariance functions with locally varying anisotropy proposed by Fouedjio et al. (2016) is used. Previously, Fouedjio (2015) applied a non-stationary geostatistical approach based on space deformation to address this problem. The purpose of this present contribution is to highlight the added value of a non-stationary geostatistical approach compared to a stationary one.

The remainder of paper is structured as follows. The geostatistical approach based on closed-form non-stationary covariance functions with locally varying anisotropy is described in “[Closed-Form Non-stationary Covariances Based Approach](#)” sec-

tion. The application of this non-stationary geostatistical approach on the breccia pipe elevation dataset is presented in “[Application to Breccia Pipe Elevation](#)” section. Comparisons are made to the conventional stationary method. Concluding remarks are provided in “[Concluding Remarks](#)” section.

## CLOSED-FORM NON-STATIONARY COVARIANCES BASED APPROACH

### Modeling

Consider  $Y(\cdot) = \{Y(x) : x \in G \subseteq R^p, p \geq 1\}$  a random field defined on a fixed continuous study domain  $G$  and reflecting the studied phenomenon. We consider that  $Y(\cdot)$  is governed by the following equation:

$$Y(x) = m(x) + \sigma(x)Z(x), \quad \forall x \in G, \quad (1)$$

where  $m(\cdot) : R^p \rightarrow R$  is an unknown fixed function,  $\sigma(\cdot) : R^p \rightarrow R^+$  is an unknown positive fixed function, and  $Z(\cdot)$  is a zero expectation, unit variance random field with correlation function defined by

$$R^{NS}(x, y) = \phi_{xy} R^S\left(\sqrt{Q_{xy}(x-y)}\right), \quad \forall (x, y) \in G \times G \quad (2)$$

with  $\phi_{xy} = |\Sigma_x|^{1/4} |\Sigma_y|^{1/4} \left| \frac{\Sigma_x + \Sigma_y}{2} \right|^{-1/2}$ ,  $Q_{xy}(h) = h^T \left( \frac{\Sigma_x + \Sigma_y}{2} \right)^{-1} h$ ,  $\forall h \in R^p$ .  $\Sigma(\cdot) : R^p \rightarrow PD_p(R)$ ,  $x \mapsto \Sigma_x$  is a mapping from  $R^p$  to  $PD_p(R)$  the set of real-valued positive definite  $p$ -dimensional square matrices;  $R^S(\cdot)$  is a continuous isotropic stationary correlation function, positive definite on  $R^p$ , for all  $p \in \mathbb{N}^*$ .

The class of closed-form non-stationary covariance functions depicted by Eq. (2) is that developed by Paciorek and Schervish (2006). Fouedjio et al. (2016) shows that this class is obtained by convolving an orthogonal random measure with a spatially varying random weighting function. The intuition behind this class is that at each location  $x$  is assigned a matrix  $\Sigma_x$  interpreted as a locally varying geometric anisotropy matrix. The correlation between two locations  $x$  and  $y$  is obtained by averaging the two matrices at  $x$  and  $y$ . In this way, the local characteristics at both locations influence the correlation of the corresponding target values, allowing to account for the non-stationarity. One will note that this class provides non-stationary versions of some well-known stationary correlation functions, for a specific choice of the isotropic stationary correlation function  $R^S(\cdot)$

(Gaussian, exponential, power exponential, and Matérn). However, this class does not enable the use of isotropic stationary correlation function with compact support like the spherical model (Paciorek and Schervish 2006; Fouedjio et al. 2016). Indeed, the correlation function  $R^S(\cdot)$  specified in Eq. (2) must be positive definite in all dimensions.

From model defined in Eq. (1), the two first moments of the random field  $Y(\cdot)$  are given by

$$E(Y(x)) = m(x), \quad \forall x \in G, \quad (3)$$

$$\begin{aligned} \text{Cov}(Y(x), Y(y)) &= \sigma(x)\sigma(y)R^{NS}(x, y) \\ &\equiv C^{NS}(x, y), \quad \forall (x, y) \in G \times G. \end{aligned} \quad (4)$$

Then, the non-stationarity of the random field  $Y(\cdot)$  is characterized by the spatially varying parameters  $m(\cdot)$ ,  $\sigma(\cdot)$ , and  $\Sigma(\cdot)$  defined at any location of the study domain. Under the local stationarity assumption (Matheron 1971), parameters are smooth functions varying slowly in space so that at any location  $x_0 \in G$ , we can define a neighborhood  $V_{x_0} = \{x \in G, \|x_0 - x\| \leq b\}$  wherein the mean  $m(\cdot)$  and the covariance function  $C^{NS}(\cdot, \cdot)$  are approximately stationary. Thus,  $\forall (x, y) \in V_{x_0} \times V_{x_0}$  Eqs. (3) and (4) are reduced as follows:

$$m(x) \approx m(y) \approx m(x_0), \quad (5)$$

$$\begin{aligned} C^{NS}(x, y) &\approx \sigma^2(x_0)R^S\left(\sqrt{(x-y)^T \Sigma_{x_0}^{-1}(x-y)}\right) \\ &\equiv C^S(x-y; x_0). \end{aligned} \quad (6)$$

Thus, locally, the mean  $m(\cdot)$  is constant as shown in Eq. (5) and the non-stationary covariance function depicted by Eq. (4) is reduced to the anisotropic stationary covariance function given in Eq. (6). The anisotropy matrix  $\Sigma_{x_0}$  at any location  $x_0 \in G$  is parameterized through the spectral decomposition, thus ensuring its positive definiteness. Formally, we have  $\Sigma_{x_0} = \Psi_{x_0} \Lambda_{x_0} \Psi_{x_0}^T$ , where  $\Lambda_{x_0}$  is the diagonal matrix of eigenvalues and  $\Psi_{x_0}$  is the eigenvector matrix formulated in 2d as follows:

$$\Lambda_{x_0} = \begin{pmatrix} \lambda_1^2(x_0) & 0 \\ 0 & \lambda_2^2(x_0) \end{pmatrix}, \quad \Psi_{x_0} = \begin{pmatrix} \cos\psi(x_0) & \sin\psi(x_0) \\ -\sin\psi(x_0) & \cos\psi(x_0) \end{pmatrix},$$

where  $\lambda_1(x_0), \lambda_2(x_0) > 0$  controls the local ranges and  $\psi(x_0) \in [0, \pi)$  specifies the local azimuth.

## Inference

Suppose that  $n$  data  $Y(s_1), \dots, Y(s_n)$  are collected, at known spatial locations,  $s_1, \dots, s_n \in G$ . The goal is to estimate the mean function  $m(\cdot)$ , the standard deviation function  $\sigma(\cdot)$ , and the anisotropy function  $\Sigma(\cdot)$  characterized by  $\lambda_1(\cdot), \lambda_2(\cdot)$  and  $\psi(\cdot)$ . This is achieved by using a step-by-step estimation procedure developed by Fouedjio et al. (2016) and similar to one proposed by Machuca-Mory and Deutsch (2013). First, a local stationary variogram kernel estimator is defined at any location of interest. Then, this is used in a weighted local least-squares procedure to estimate parameters at a representative set of locations referred to as anchor locations. Finally, a kernel smoothing technique is used to make available the parameter values at any location of interest.

### Step 1

A non-parametric kernel moment estimator of the local stationary variogram  $\gamma(h; x_0) = \sigma^2(x_0) - C^S(h; x_0)$  at a fixed location  $x_0 \in G$  and for a spatial lag  $h \in R^p$ ,  $\|h\| \leq b$  is defined as follows (Fouedjio et al. 2016):

$$\hat{\gamma}_\epsilon(h; x_0) = \frac{\sum_{V(h)} K_\epsilon^\star(x_0, s_i) K_\epsilon^\star(x_0, s_j) [Y(s_i) - Y(s_j)]^2}{2 \sum_{V(h)} K_\epsilon^\star(x_0, s_i) K_\epsilon^\star(x_0, s_j)}, \quad (7)$$

where  $K_\epsilon^\star(x_0, s_i) = K_\epsilon(x_0, s_i) / \sum_{r=1}^n K_\epsilon(x_0, s_r)$  are standardized weights and  $K_\epsilon(\cdot, \cdot)$  is a non-negative, symmetric kernel on  $R^p \times R^p$ , with bandwidth parameter  $\epsilon > 0$ .  $V(h) = \{(s_i, s_j) \in G \times G : s_i - s_j = h\}$  is the set of all pairs of locations separated by vector  $h$ . For irregularly spaced data where there are usually not enough observations separated by exactly  $h$ ,  $V(h)$  is commonly modified by  $\{(s_i, s_j) \in G \times G : s_i - s_j \in T(h)\}$ , where  $T(h)$  is a tolerance region surrounding  $h$ .

The role of the kernel function in Eq. (7) is to smoothly down-weight the squared differences (for each spatial lag) according to the distance of these paired values from a target location. A pair of locations receives a weight that is proportional to the product of the individual weights. Geographically weighted estimators similar to one described by

Eq. (7) have been proposed by Harris et al. (2010) and Machuca-Mory and Deutsch (2013) for the moving windows non-stationary geostatistical approach. Regarding the choice of the kernel function, Fouedjio et al. (2016) suggest a Gaussian kernel whose support is non-compact. Indeed, a kernel function with non-compact support includes all observations, then allows to avoid artifacts caused by using only observations close to the target location. It also reduces instability of the kernel estimator of the local stationary variogram at regions with low sampling density. Furthermore, the use of a Gaussian kernel provides a smooth parameter estimates and then it is compatible with the local stationarity assumption. The size of the local stationarity neighborhood  $b$  is set with respect to the bandwidth  $\epsilon$ :  $b = \sqrt{3}\epsilon$  so that the standard deviation of the isotropic stationary Gaussian kernel matches the isotropic stationary uniform kernel, which has a compact support. Another possible choice for  $b$  is a quantile of the isotropic stationary Gaussian kernel (e.g.,  $b \approx 2\epsilon$ ).

Step 2

Let  $A$  be a set of representative locations referred to as anchor locations over the study domain  $G$ . Given the local stationary variogram kernel estimator defined in Eq. (7), the estimation of the parameters vector  $\theta(x_0) = (\sigma(x_0), \lambda_1(x_0), \lambda_2(x_0), \psi(x_0))$  at an anchor location  $x_0 \in A$  is given by following minimization problem (Fouedjio et al. 2016):

$$\hat{\theta}(x_0) = \arg \min_{\theta_0 \in \Theta} \|w_\epsilon(x_0) \odot (\gamma(\theta_0) - \hat{\gamma}_\epsilon(x_0))\|, \quad (8)$$

where  $\theta_0$  is the vector of unknown parameters and  $\Theta$  is an open parameter space;  $\gamma^T(\theta_0) = [\gamma(h_l; \theta_0)]_{l=1 \dots L}$ ;  $\hat{\gamma}_\epsilon^T(x_0) = [\hat{\gamma}_\epsilon(h_l; x_0)]_{l=1 \dots L}$ ;  $w_\epsilon^T(x_0) = [w_\epsilon(h_l; x_0)]_{l=1 \dots L}$ , with  $w_\epsilon(h_l; x_0) = \left[ \left( \sum_{V(h_l; x_0)} K_\epsilon^{\star\star}(x_0, s_i) K_\epsilon^{\star\star}(x_0, s_j) \right) / \|h_l\| \right]^{1/2}$ ;  $\{h_l \in \mathbb{R}^p\}_{l=1, \dots, L}$  is a collection of lag vectors;  $\odot$  is the product term by term.

Given the parameter vector estimate  $\hat{\theta}(x_0)$  at anchor location  $x_0 \in A$ , which characterizes the local stationary spatial dependence structure of data, the mean  $m(x_0)$  is estimated explicitly by a local stationary kriging of the mean (Matheron 1971). As highlighted by Fouedjio et al. (2016), the estimation of parameters  $\sigma(\cdot), \lambda_1(\cdot), \lambda_2(\cdot)$  and  $\psi(\cdot)$  governing the non-stationary spatial dependence structure of data does not require the prior estimation of the mean  $m(\cdot)$ . Moreover, no model is specified for the

latter. Indeed, the local stationary variogram kernel estimator given in Eq. (7) filters out the mean  $m(\cdot)$  at short distances, because the mean is approximately equal to a constant in the quasi-stationarity neighborhood. For distances up to the radius of the quasi-stationarity neighborhood, the local stationary variogram kernel estimator thus estimates well the local stationary spatial dependence structure, whose parameters are estimated as shown in Eq. (8).

Step 3

For the kriging or simulation purpose, it is necessary to get the estimates of the spatially varying parameters at any location of interest, especially at unobserved and observed locations. In practice, it is unnecessary to solve the minimization problem depicted by Eq. (8) at each location of interest. Indeed, doing so is computationally intensive and redundant for close locations because their estimates are highly correlated. To reduce the computational burden, Fouedjio et al. (2016) propose to perform the parameter estimation as shown in step 2. Then, using the resulting parameter estimates at anchor locations, a kernel smoothing method is used to make them available at any location of interest. This idea was also proposed by Machuca-Mory and Deutsch (2013) for the moving windows non-stationary geostatistical approach. Fouedjio et al. (2016) suggest the Nadaraya-Watson kernel smoother (Wand and Jones 1995), which is suitable and relatively simple. However, other smoothers can be used as well (local polynomials, splines, etc.). The set of anchor locations may be chosen as a grid over the study domain. The number of anchor locations may depend on the complexity of the true underlying non-stationarity and especially it is a trade-off between the computing time and the accuracy of the estimation. Indeed, it is possible to choose the set of anchor locations as the large grid of locations to predict but it will be computationally demanding.

The estimation of the spatially varying parameters depends on the bandwidth parameter  $\epsilon$  used in the computation of the local stationary variogram kernel estimator defined in Eq. (7). The estimation of the spatial dependence structure of data being rarely a goal per se but an intermediate step before kriging, the bandwidth parameter is selected by a data-driven method consisting for choosing the

bandwidth value that gives the best cross-validation mean square prediction error (Fouedjio et al. 2016):

$$\text{MSPE}(\epsilon) = \frac{1}{n} \sum_{i=1}^n \left( Z(s_i) - \widehat{Z}_{-i}(s_i; \epsilon) \right)^2, \quad (9)$$

where  $\widehat{Z}_{-i}(s_i; \epsilon)$  is the kriging at location  $s_i$  using all measured data excluding  $\{Z(s_i)\}$ . The prediction method is described in “Kriging” section.

## Prediction

### Kriging

Given measured data  $Y(s_1), \dots, Y(s_n)$ , the point predictor for the unknown value of the random field  $Y(\cdot)$  at an unsampled location  $s_0 \in G$  is given by the optimal linear predictor:

$$\widehat{Y}(s_0) = m(s_0) + \sum_{i=1}^n \eta_i(s_0)(Y(s_i) - m(s_i)), \quad (10)$$

where the kriging weight vector  $\boldsymbol{\eta}(s_0) = [\eta_i(s_0)]$  and the corresponding kriging variance  $Q(s_0)$  are given by  $\boldsymbol{\eta}(s_0) = C^{-1}C_0$  and  $Q(s_0) = \sigma^2(s_0) - C_0^T C^{-1}C_0$ , with  $C_0 = [C^{NS}(s_i, s_0)]$ ;  $C = [C^{NS}(s_i, s_j)]$ .

### Conditional Simulation

The unconditional simulation of the random field  $Y(\cdot) = m(\cdot) + \sigma(\cdot)Z(\cdot)$  involves the unconditional simulation of the random field  $Z(\cdot)$  with zero expectation, unit variance and non-stationary correlation function  $R^{NS}(\cdot, \cdot)$ . Since, we are in a non-stationary framework, conventional simulation methods such as spectral method or turning bands method (Lantuejoul 2002) are not adapted because they rely on the stationarity assumption. In the Gaussian framework, the unconditional simulation of the random field  $Z(\cdot)$  can be carried out using a propagative version of the Gibbs sampler proposed by Lantuejoul and Desassis (2012). This algorithm allows to simulate a Gaussian vector at a large number of locations compared to existing classical algorithms such as Cholesky method or Gibbs sampler. This algorithm does not rely on a Markov assumption and requires neither the inversion nor the factorization of a covariance matrix. From an unconditional simulation of the random field  $Y(\cdot)$ , a conditional simulation of the latter can be obtained

through the conditioning by kriging approach as in the stationary framework (Lantuejoul 2002).

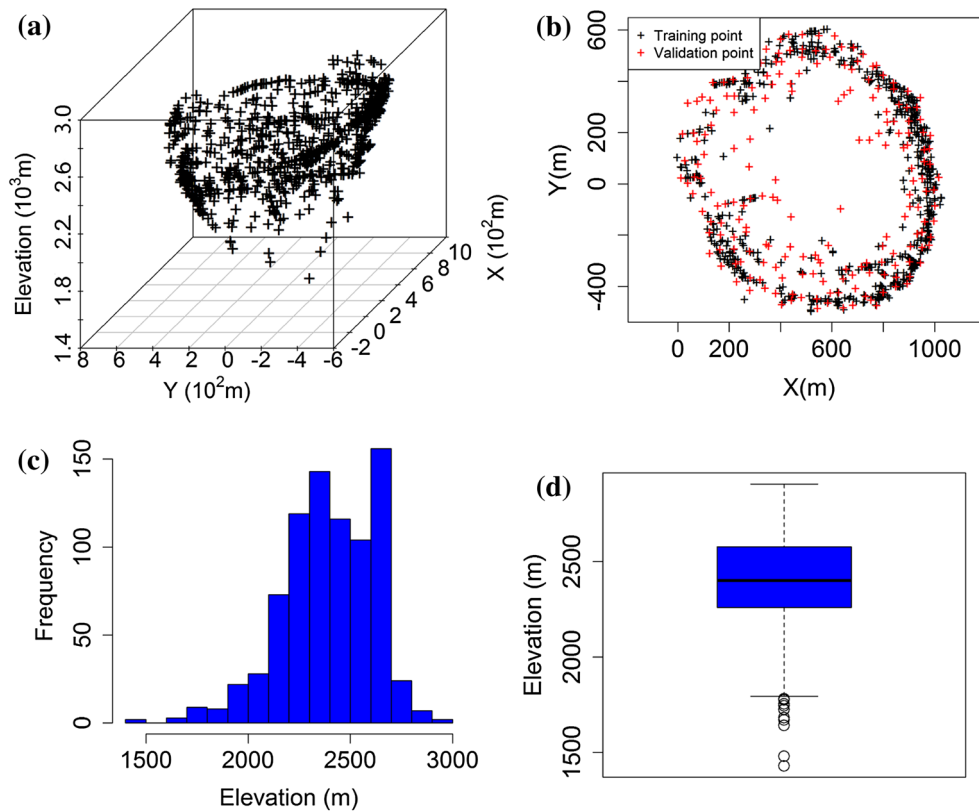
## APPLICATION TO BRECCIA PIPE ELEVATION

### Data Description

A visualization of the breccia pipe elevation data is given in Figure 1a. The geometric configuration of data is circular with high values located at margins of the mine, whereas the central part exhibits low values. There are more samples at the margin than at the center. The dataset comprises  $n = 816$  measurements partitioned into a training set ( $n_1 = 616$  measurements) and a validation set ( $n_2 = 200$  measurements) as shown in Figure 1b. The training data serve to estimate the model and the validation data serve to evaluate the prediction performances. A summary statistics of training, validation, and whole data are given in Table 1. The histogram and boxplot of data are slightly skewed with values ranging from 1429 to 2906 m, a mean of 2392 m and a median of 2401 m (Fig. 1c, d). The data present some outliers corresponding to the lowest values, which are located at the center of the study domain.

### Exploring Evidence of Variogram Non-stationarity

An exploration of the non-stationarity of the spatial dependence structure of data is carried out through the local stationary variogram (Lloyd 2010). The latter is computed at some locations across the study domain as shown in Figure 2. There is clear evidence of variogram non-stationarity, as the variographic parameters (sill and range) vary spatially. Specifically, the East margin area (including the locations 1, 2, and 3) is an area of relatively low spatial correlation, while the West margin area (comprising the locations 4, 5, and 6) is an area of relatively high spatial correlation. Indeed, the former area has a short range ( $\sim 300$  m) and high variance ( $\sim 60,000$  m<sup>2</sup>) compared to the latter area, which has a long range ( $\sim 400$  m) and low variance ( $\sim 45,000$  m<sup>2</sup>). This difference between the two sub-areas may be related to lithologic conditions. Moreover, the representation of data in Figure 2a



**Figure 1.** Elevation data, El Teniente mine, Chile.

**Table 1.** Summary Statistics of Measured Elevation Data, El Teniente Mine, Chile

	Training ( $n_1 = 616$ )	Validation ( $n_2 = 200$ )	Whole ( $n = 816$ )
Minimum	1755	1429	1429
1st quartile	2284	2162	2259
Median	2418	2354	2401
Mean	2417	2315	2392
3rd quartile	2602	2526	2577
Maximum	2906	2817	2906
SD	200	262	221

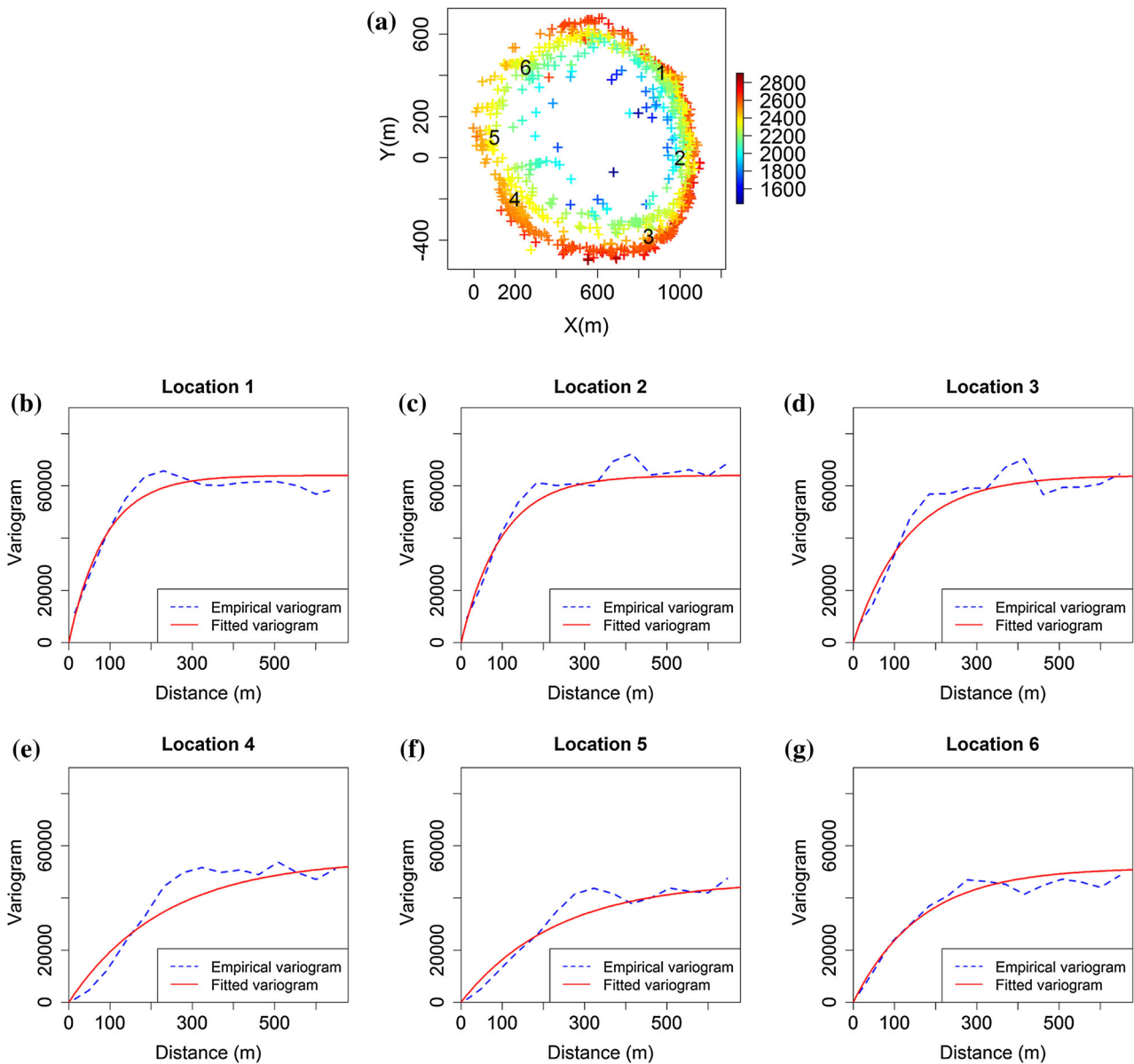
shows a spatially varying azimuth, which follows circular geometric configuration of data.

### Model Estimation Results

Following the estimation procedure described in “Inference” section, raw estimates of spatially varying parameters  $m(\cdot)$ ,  $\sigma^2(\cdot)$  and  $\Sigma(\cdot)$  at anchor locations are shown, respectively, in Figure 3b–d, thus demonstrating the non-stationarity in the data.

One can observe for example in Figure 3c the locally varying geometric anisotropy. As mentioned in “Conditional Simulation” section, such directional effects are also quite visible in the data. Note that the stationary approach has not detected a global geometric anisotropy. Raw estimates of parameters are based on the non-stationary exponential covariance function with locally varying anisotropy. The geometric configuration of anchor locations is given in Figure 3a. Despite the fact that there is almost no data in the center of the domain, the use of a non-compact kernel function (Gaussian kernel) in the computing of the local stationary variogram kernel estimator defined in Eq. (7), makes possible the estimation because it considers all data.

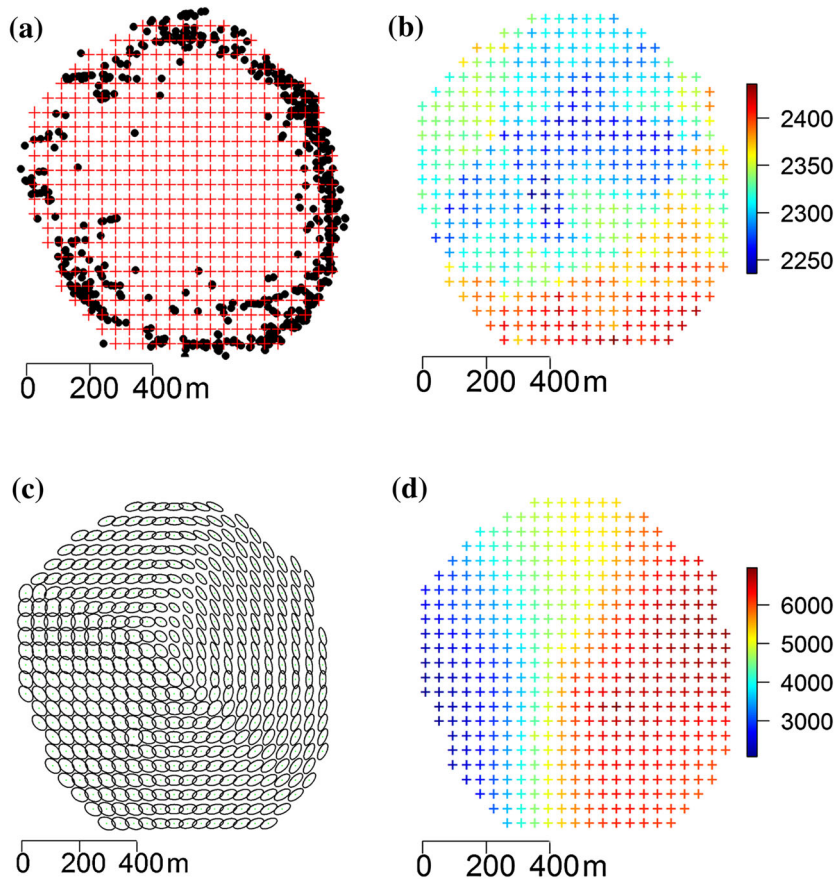
According to the bandwidth selection given by Eq. (9), Figure 4a, b shows, respectively, the mean square prediction error for cross-validation and external validation. The optimal value in cross-validation corresponds to  $\epsilon = 302$  m. This optimal value is consistent with those given by the external validation. Figure 5 shows the maps of smoothed parameters over the whole study domain for mean,



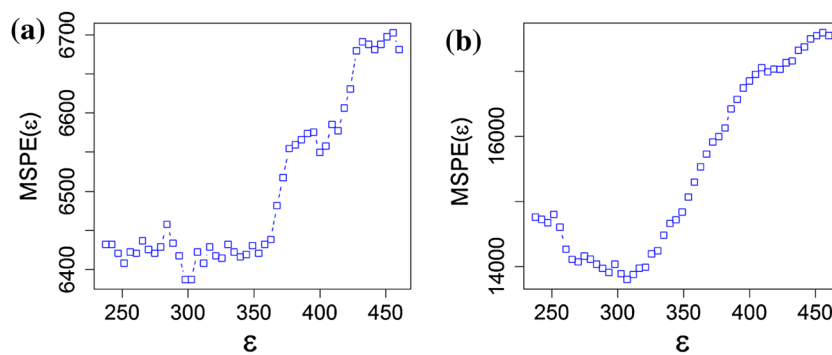
**Figure 2.** The study domain: (a) measured data and (b–d) local stationary variograms at some specified locations.

variance, anisotropy ratio, and azimuth. The variance map portrays the difference between an area of relatively high variance (East margin region) and low variance (West margin region), thereby confirming the result obtained from the exploratory analysis of non-stationarity in “Exploring Evidence of Variogram Non-stationarity” section. Moreover, the variance map is consistent with results provided by the space deformation non-stationary geostatistical approach applied by Fouedjio (2015): areas of low variance are contracted, while those of high variance are stretched.

A visualization of the covariance function at certain locations (with all other locations) via the level contours under the estimated stationary and non-stationary models is given in Figure 6. The change of shape in the non-stationary spatial dependence structure of data from one location to another one is quite visible. The stationary model is a nested isotropic model (small nugget effect, exponential, and spherical), while the non-stationary model corresponds to the non-stationary exponential covariance function with locally varying anisotropy. This difference between the estimated



**Figure 3.** The study domain: (a) data locations (dots) and anchor locations (cross), (b) estimated mean function  $\widehat{m}(\cdot)$  at anchor locations, (c) estimated anisotropy function  $\widehat{\Sigma}(\cdot)$  at anchor locations (the ellipses were scaled to ease visualization) and (d) estimated variance function  $\widehat{\sigma}^2(\cdot)$  at anchor locations.



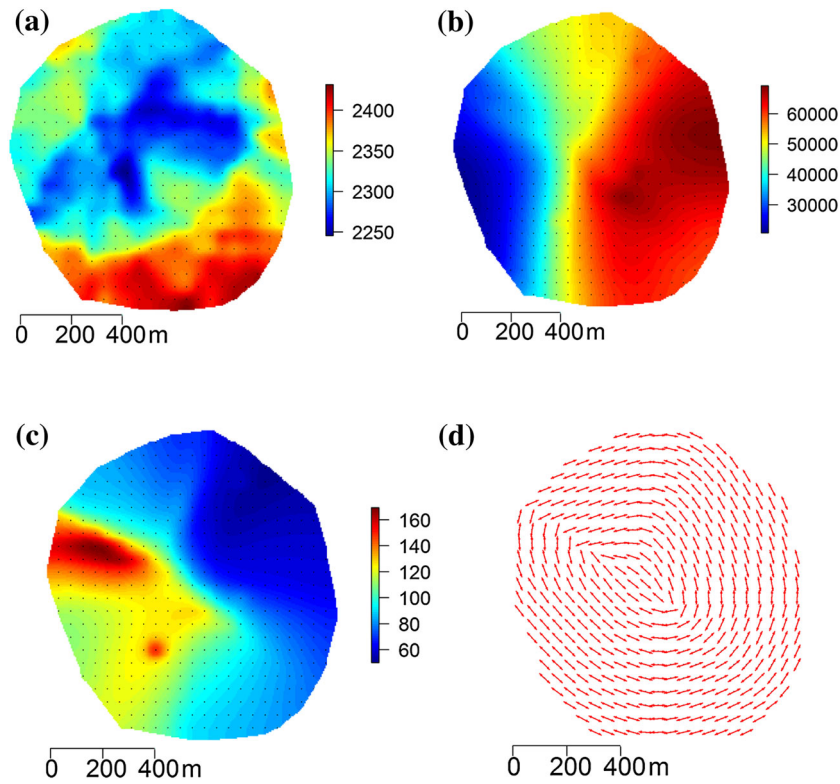
**Figure 4.** Bandwidth parameter selection based on the mean square prediction error  $MSPE(\epsilon)$  for (a) cross-validation and (b) external validation.

stationary and non-stationary models is reflected by both the prediction accuracy and the prediction uncertainty accuracy (“[Evaluation of Predictions](#)” section).

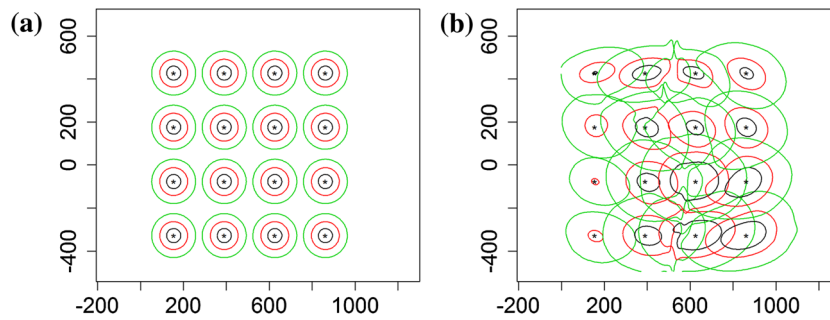
**Evaluation of Predictions**

The evaluation of the predictive performance of the non-stationary method relative to the stationary





**Figure 5.** The study domain: smoothed parameters for (a) mean, (b) variance, (c) anisotropy ratio, and (d) azimuth.



**Figure 6.** Covariance level contours at few locations for (a) the estimated stationary model and (b) the estimated non-stationary model. Level contours correspond to 30,000 m<sup>2</sup> (black), 20,000 m<sup>2</sup> (red), and 10,000 m<sup>2</sup> (green).

method is carried out using some well-known statistics calculated in external validation, namely mean absolute error (*MAE*), root mean square error (*RMSE*),

normalized mean square error (*NMSE*), logarithmic score (*LogS*), and continued ranked probability score (*CRPS*). These criteria are formulated as follows:

$$\begin{aligned}
 MAE &= \frac{1}{n_2} \sum_{i=1}^{n_2} |\hat{Z}(s_i) - Z(s_i)|, \\
 RMSE &= \left[ \frac{1}{n_2} \sum_{i=1}^{n_2} (\hat{Z}(s_i) - Z(s_i))^2 \right]^{\frac{1}{2}}, \\
 NMSE &= \frac{1}{n_2} \sum_{i=1}^{n_2} \left( \frac{\hat{Z}(s_i) - Z(s_i)}{\hat{\sigma}(s_i)} \right)^2, \\
 LogS &= \frac{1}{n_2} \sum_{i=1}^{n_2} \left( \frac{1}{2} \log(2\pi\hat{\sigma}^2(s_i)) + \frac{1}{2} \left( \frac{\hat{Z}(s_i) - Z(s_i)}{\hat{\sigma}(s_i)} \right)^2 \right), \\
 CRPS &= \frac{1}{n_2} \sum_{i=1}^{n_2} \int_{-\infty}^{+\infty} (F_i(z) - 1\{Z(s_i) \leq z\})^2 dz, \\
 F_i(z) &= P(Z(s_i) \leq z | \text{training data}),
 \end{aligned}
 \tag{11}$$

where  $\hat{Z}(s_i)$  is the kriging at a validation data location  $s_i$  computed from all training data and  $\hat{\sigma}^2(s_i)$  the corresponding kriging variance.  $n_2$  is the number of validation data locations.

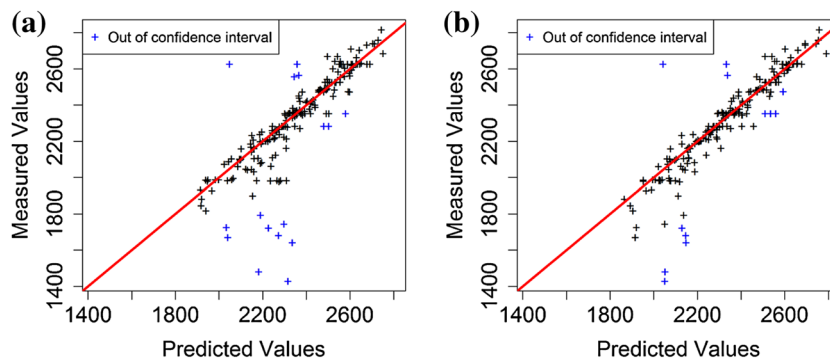
For *MAE*, *RMSE*, *LogS* and *CRPS*, the smaller, the better; for *NMSE*, the nearer to one the better. The prediction accuracy is measured by *MAE* and *RMSE*. The prediction uncertainty accuracy is as-

essed by *NMSE*, *LogS*, and *CRPS*, which combine the kriging and the kriging variance. The *MAE*, *RMSE*, and *NMSE* do not depend on the distribution of measured data. The *LogS* is equivalent to the pseudo-likelihood in the Gaussian framework. The *CRPS* corresponds to the distance between the distribution function of the predicted variable and the measured data (itself expressed as a distribution function). It is generally calculated in the Gaussian setting where it admits a closed-form expression. Although the *LogS* and *CRPS* scores are usually calculated in the Gaussian framework, they are quite robust. The probability Gaussian-type confidence interval is calculated also at each validation location (i.e., using  $\hat{Z}(s_i) \pm 1.96\hat{\sigma}(s_i)$ ) and the proportion of validation locations where the 95% confidence interval (*PCI*) actually includes the true value is computed. This proportion should be near 95% for an accurate modeling of uncertainty. The correlation between true and predicted values (*Rho*) is computed also, the closer to one the better. A description of these criteria can be found in Chilès and Delfiner (2012), Zhang and Wang (2010), and Gneiting and Raftery (2007).

Scatter plots of kriged versus measured values for the stationary and non-stationary methods are presented in Figure 7. Globally, the non-stationary approach provides a more accurate prediction compared to the stationary approach. This is demonstrated by a reduced *MAE* and *RMSE*, and by an increased *Rho* (Table 2). For example, in terms of *RMSE*, the improvement is about 23% with respect to the stationary approach. The reliability of the kriging variances by *NMSE*, *LogS*, *CRPS*, and *PCI* (Table 2) shows that the non-stationary approach is more accurate for modeling of uncertainty

**Table 2.** Validation Statistics of the Prediction Performance

Criterion	Stationary Approach	Non-stationary Approach
<i>MAE</i>	79.7	61.5
<i>RMSE</i>	154.4	117.9
<i>Rho</i>	0.82	0.90
<i>NMSE</i>	0.98	0.73
<i>LogS</i>	2439	2315
<i>CRPS</i>	123.6	99.1
<i>PCI</i>	0.92	0.94



**Figure 7.** Scatter plots of predicted versus measured values for (a) the stationary approach and (b) the non-stationary approach.

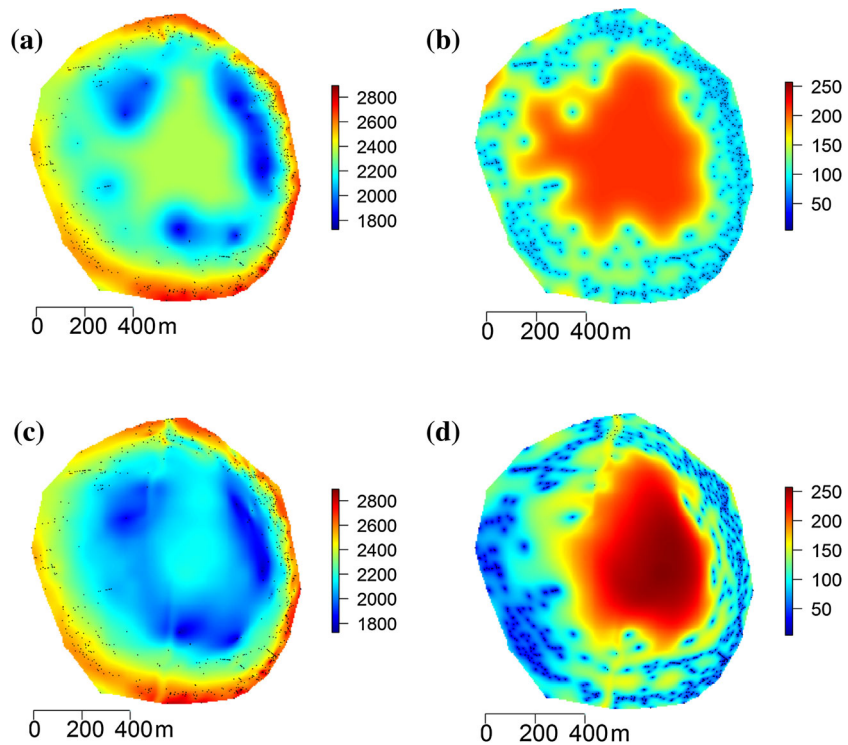


Figure 8. Predictions and prediction standard deviations based on (a, b) the estimated stationary model and (c, d) the estimated non-stationary model.

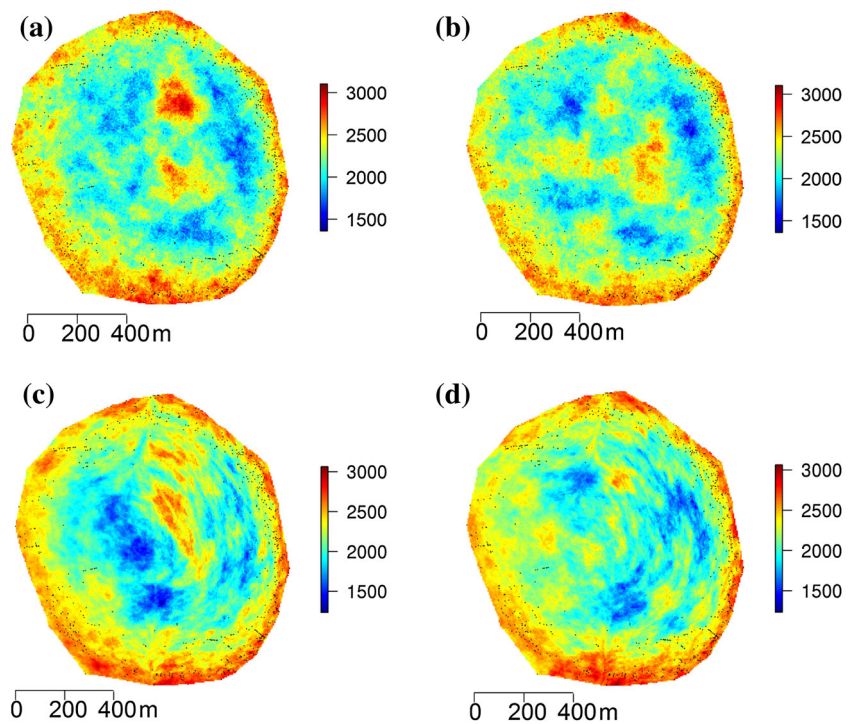


Figure 9. Two conditional simulations based on (a, b) the estimated stationary model and (c, d) the estimated non-stationary model.

compared to the stationary approach. When considering the proportion of validation locations included in the 95% confidence interval, the non-stationary approach shows 12 locations outside (94% of locations included in that interval), while the stationary approach shows 16 locations outside (92% of locations included in that interval) as shown in Figure 7 and reported in Table 2 (we expect about  $200 \times 0.05 = 10$  locations outside). Specifically, the stationary approach has more difficulty in predicting the lowest values (located at the center) compared to the non-stationary approach. Furthermore, we note that this non-stationary approach gives slightly better results compared to the space deformation non-stationary geostatistical approach (Fouedjio 2015, Table 2).

The kriging results for the stationary and non-stationary methods are shown in Figure 8. The general appearance of the map of kriged values associated with each method differs significantly (Fig. 8a, c). The difference is particularly marked at the center of the domain where there are not enough data locations. The stationary and non-stationary methods differed notably in describing the uncertainty associated with the predictions (Fig. 8b, d). One can see that the non-stationary approach tends to provide a low prediction standard deviation in the area of low spatial variability (West margin area), while it tends to give a high prediction standard deviations in the area of high spatial variability (East margin area). Thus, prediction standard deviations reflect not only the samples configuration and availability around target locations, but also the local variability. On the other hand, kriging standard deviations map for the stationary approach shows slight differences over the margin areas of the mine, which were dependent on the sampling intensity. Such a pattern was expected for a stationary approach because it is based on identical global structural parameters throughout the study domain, while the non-stationary approach adapts to locally varying structure of data. Figure 9 shows some conditional simulations in the Gaussian framework, based on the estimated stationary and non-stationary models. Conditional simulations under the non-stationary model differ from one under the stationary model, especially in terms of the locally varying anisotropy.

## CONCLUDING REMARKS

This paper demonstrated the added value of using the geostatistical approach based on explicit

non-stationary covariance functions with locally varying anisotropy, to predict the surface elevation of the Braden breccia pipe at the El Teniente mine, Chile. This non-stationary approach is capable of capturing some varying spatial features such as locally varying geometric anisotropy in the data that cannot be detected by a stationary method. Thus, the non-stationary approach outperforms the stationary one in terms of prediction and prediction uncertainty accuracies. Moreover, as an exploratory tool for the non-stationarity, it allows to identify areas of high and low spatial variability, and to detect the existence of relationship between the mean and the variance known as proportional effect.

The geostatistical approach based on explicit non-stationary covariance functions with locally varying anisotropy requires enough observations to be able to properly capture the non-stationarity as any non-stationary geostatistical approach. It is also based on local stationarity assumption, then it works well for smoothly varying non-stationarity. Thus, it can be difficult to apply on sparse data or data with abrupt spatial structure variations. In such cases, it may be advisable to proceed under a stationary framework or to divide the study domain if possible. Since this non-stationary method relies on the local stationary assumption, it should not be applied to intrinsically locally stationary spatial processes with unbounded local stationary variograms. Thus, it is important to check that the local stationary variogram kernel estimator used to describe local spatial variation has a sill, before deciding to use this non-stationary method. This is done by visualizing the local stationary variogram kernel estimator at anchor locations. This non-stationary approach is time-consuming compared to a stationary approach. For this case study, the computational time is 10 times more than that of the stationary method. However, if there are enough data to allow reliable inference, this non-stationary method will outperform a stationary method in terms of prediction and prediction uncertainty accuracies as in this case study.

## ACKNOWLEDGMENTS

We would like to thank the company CO-DELCO-Chile for providing the data used in this paper. We acknowledge the editor-in-chief of NARR for expert handling of our manuscript and

we thank the anonymous referees for their constructive comments.

## REFERENCES

- Anderes, E. B., & Stein, M. L. (2011). Local likelihood estimation for nonstationary random fields. *Journal of Multivariate Analysis*, 102(3), 506–520.
- Chilès, J. P., & Delfiner, P. (2012). *Geostatistics: Modeling spatial uncertainty*. New York: Wiley.
- Fouedjio, F. (2015). Space deformation non-stationary geostatistical approach for prediction of geological objects: Case study at El Teniente mine (Chile). *Natural Resources Research*. doi:10.1007/s11053-015-9287-7.
- Fouedjio, F., Desassis, N., & Rivoirard, J. (2016). A generalized convolution model and estimation for non-stationary random functions. *Spatial Statistics*, 16, 35–52.
- Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378.
- Guttorp, P., & Sampson, P. D. (1994). *Methods for estimating heterogeneous spatial covariance functions with environmental applications* (Vol. 12, pp. 661–689). New York: Elsevier.
- Guttorp, P., & Schmidt, A. M. (2013). Covariance structure of spatial and spatiotemporal processes. *Wiley Interdisciplinary Reviews: Computational Statistics*, 5(4), 279–287.
- Haas, T. C. (1990a). Kriging and automated variogram modeling within a moving window. *Atmospheric Environment. Part A. General Topics*, 24(7), 1759–1769.
- Haas, T. C. (1990b). Lognormal and moving window methods of estimating acid deposition. *Journal of the American Statistical Association*, 85(412), 950–963.
- Harris, P., Charlton, M., & Fotheringham, A. S. (2010). Moving window kriging with geographically weighted variograms. *Stochastic Environmental Research and Risk Assessment*, 24(8), 1193–1209.
- Lantuejoul, C. (2002). *Geostatistical simulation: Models and algorithms*. Berlin: Springer.
- Lantuejoul, C. & Desassis, N. (2012). Simulation of a gaussian random vector: A propagative version of the gibbs sampler. In *The 9th international geostatistics congress*.
- Lloyd, C. (2010). *Local models for spatial analysis* (2nd ed.). Boca Raton: CRC Press.
- Machuca-Mory, D., & Deutsch, C. (2013). Non-stationary geostatistical modeling based on distance weighted statistics and distributions. *Mathematical Geosciences*, 45(1), 31–48.
- Magneron, C., Jeanne, N., Le Moine, O., & Bourillet, J. F. (2010). *Integrating prior knowledge and locally varying parameters with moving-geostatistics: Methodology and application to bathymetric mapping, geoENV VII—Geostatistics for Environmental Applications* (Vol. 16). Dordrecht: Springer.
- Maksaev, V., Munizaga, F., McWilliams, M., Fanning, M., Matheron, R., Ruiz, J., et al. (2004). New chronology for El Teniente, Chilean Andes, from U-Pb, 40Ar/39Ar, Re-Os, and fission-track dating: Implications for the evolution of a supergiant porphyry Cu-Mo deposit. *Society of Economic Geologists Special Publication*, 11, 15–54.
- Matheron, G. F. (1971). *The theory of regionalized variables and its applications. Les cahiers du Centre de Morphologie Mathématique de Fontainebleau* (Vol. 5). Paris: École Nationale Supérieure des Mines de Paris.
- Paciorek, C. J., & Schervish, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17(5), 483–506.
- Sampson, P. D., Damian, D., & Guttorp, P. (2001). *Advances in modeling and inference for environmental processes with nonstationary spatial covariance. Geoenv III—Geostatistics for environmental applications* (Vol. 11). Dordrecht: Springer.
- Skewes, M. A., Arévalo, A., Floody, R., Zuñiga, P. H., & Stern, C. R. (2002). The giant El Teniente breccia deposit: Hypogene copper distribution and emplacement. *Society of Economic Geologists Special Publication*, 9, 299–332.
- Spencer, E. T., Wilkinson, J. J., Creaser, R. A., & Seguel, J. (2015). The distribution and timing of molybdenite mineralization at the El Teniente Cu-Mo porphyry deposit, Chile. *Economic Geology*, 110(2), 387–421.
- Stein, M. (2005). Nonstationary spatial covariance functions. Unpublished report.
- Wand, M., & Jones, C. (1995). *Kernel smoothing*. Monographs on statistics and applied probability. London: Chapman and Hall.
- Zhang, H., & Wang, Y. (2010). Kriging and cross-validation for massive spatial data. *Environmetrics*, 21(3–4), 290–304.