Original Paper

Geostatistical Simulation with a Trend Using Gaussian Mixture Models

Jianan Qu [1,2] **and Clayton V. Deutsch**[1]

Geostatistics applies statistics to quantitatively describe geological sites and assess the uncertainty due to incomplete sampling. Strong assumptions are required regarding the location independence of statistical parameters to construct numerical models with geostatistical tools. Most geological data exhibit large-scale deterministic trends together with short-scale variations. Such location dependence violates the common geostatistical assumption of stationarity. The trend-like deterministic features should be modeled prior to conventional geostatistical prediction and accounted for in subsequent geostatistical calculations. The challenge of using a trend in geostatistical simulation algorithms for the continuous variable is the subject of this paper. A stepwise conditional transformation with a Gaussian mixture model is considered to provide a stable and artifact-free numerical model. The complex features of the regionalized variable in the presence of a trend are removed in the forward transformation and restored in the back transformation. The Gaussian mixture model provides a seamless bin-free approach to transformation. Data from a copper deposit were used as an example. These data show an apparent trend unsuitable for conventional geostatistical algorithms. The result shows that the proposed algorithm leads to improved geostatistical models.

KEY WORDS: Non-stationary regionalized variable, Stepwise conditional transformation, Sequential Gaussian simulation.

INTRODUCTION

Geostatistics provides tools to construct numerical models and support mineral and petroleum resource estimates. Geostatistical modeling requires parameters and assumptions based on the limited data available for a particular deposit. The parameters include global probability distributions, variograms and training images. A common assumption of geostatistical modeling is stationarity of these parameters, that is, they are independent of location. For example, the expected value or mean value is assumed constant for all locations within each domain. The decision of stationarity is made prior to any geostatistical prediction (Deutsch and Journel 1998; Davis and Sampson 2002; Wackernagel 2003; Pyrcz and Deutsch 2014). Most decisions of stationarity are implicitly made with the application of a particular algorithm.

Simulation plays an important role in geostatistical modeling. Simulation draws multiple realizations to characterize the geological heterogeneity and quantify the uncertainty of the regionalized variable. The realizations reflect the statistical characteristics of the observed data and unsampled locations. The central step in simulation is to draw

[1] Department of Civil and Environmental Engineering, Centre for Computational Geostatistics, 6-247 Donadeo Innovation Centre for Engineering, University of Alberta, 9211-116 Street, Edmonton, AB T6G 1H9, Canada.
[2] To whom correspondence should be addressed; e-mail: jqu2@ualberta.ca

simulated values from conditional distributions (Deutsch and Journel 1998). The variation between multiple realizations represents the geological uncertainty (Goovaerts 2001; Rossi and Deutsch 2014). This uncertainty is considered with other aspects of a project to support decision-making. Sequential Gaussian simulation is widely used (Deutsch and Journel 1998). Such simulation-based techniques make a strong assumption of stationarity. However, real geological data often exhibit trends or non-stationary location dependent features (Wang et al. 2012; Boisvert et al. 2013). This violates the assumption of stationarity and the local accuracy of the predicted uncertainty may be unreliable.

Non-stationary geostatistical methods have been developed. One approach is decomposition of the regionalized variable into a deterministic component with large-scale features and a stochastic component with small-scale variations. Several methods are available for modeling the deterministic component (Journel and Huijbregts 2003; Machuca-Mory 2010; Rossi and Deutsch 2014). The conventional geostatistical algorithms would only be applied to the stochastic component assuming that it is stationary and the modeled deterministic component would be added back to the simulated result (Wackernagel 2003; Chiles and Delfiner 2012). The naive approach of modeling with residuals and adding the trend model back in the final models is straightforward; however, the variogram from the residuals is biased downward compared with the underlying variogram (Delfiner 1976; Sabourin 1976; Chiles and Delfiner 2012). Additionally, some constraints should be considered to ensure nonnegative simulated values in the final model when adding the trend at the end (Leuangthong 2003). Another approach is a conditional transformation, such as the stepwise conditional transformation (Leuangthong and Deutsch 2003, 2004) or a locally varying transformation (Gonzales et al. 2006). The complex features of the trend can be removed in the forward transformation, but artifacts could also be introduced due to the bin selection or the use of few data in the transformation. Other approaches include the intrinsic random functions of order k (Matheron 1973), the non-stationary covariance function (Sampson and Guttorp 1992), the moving window averages (Brunsdon et al. 2002), the spatially varying linear model of coregionalization (Gelfand et al. 2004) and the local random function (Machuca-Mory 2010). However, these non-stationary techniques encounter difficulties in practice (Rossi and Deutsch 2014).

This paper develops a geostatistical modeling algorithm that accounts for the deterministic features of continuous regionalized variables in an artifact-free fashion. A methodology similar to the nonparametric stepwise conditional transformation proposed by Leuangthong and Deutsch (2003) is considered. The conditional distributions are calculated by a Gaussian mixture model fitted to the deterministic trend and the data. The trend-like features in the regionalized variable are removed by the conditional transformation. A porphyry copper deposit is considered where the grade shows an obvious trend. The trend is assumed known without uncertainty. A comparison to conventional geostatistical calculations is made. The results show that the geostatistical modeling with trend modeling outperforms the conventional geostatistical modeling with less error and better reproduction of important features of the regionalized variable.

BACKGROUND

The stepwise conditional transformation technique was first introduced by Rosenblatt (1952) as an extension of the normal score transformation. Leuangthong and Deutsch (2003) introduced this technique to geostatistics and developed practical applications (Leuangthong and Deutsch 2004). This technique removes some complex features from data.

Consider $\{Z_k(\boldsymbol{u}), k = 1, \ldots, K\}$ is a set of K stationary random functions. $\{\boldsymbol{u}_i, i = 1, \ldots, n\}$ represents a set of n data locations. The observations of the random function at location $\boldsymbol{u}_i$ are denoted by $z_i = \{z_{i,1}, \ldots, z_{i,K}\}$. The first variable $\{z_{i,1}, i = 1, \ldots, n\}$ is transformed independently to Gaussian units through a normal score transformation, the second variable $\{z_{i,2}, i = 1, \ldots, n\}$ is transformed conditional to the first variable and so on:

$$
\begin{aligned}
y_{i,1} &= G^{-1}\big(F_1\big(z_{i,1}\big)\big) \\
y_{i,2} &= G^{-1}\big(F_{2|1}\big(z_{i,2} \mid z_{i,1}\big)\big) \\
&\vdots \\
y_{i,k} &= G^{-1}\big(F_{k|1,\ldots,k-1}\big(z_{i,k} \mid z_{i,1}, \ldots, z_{i,k-1}\big)\big) \quad (1) \\
&\vdots \\
y_{i,K} &= G^{-1}\big(F_{K|1,\ldots,K-1}\big(z_{i,K} \mid z_{i,1}, \ldots, z_{i,K-1}\big)\big) \\
&\quad i = 1, \ldots, n
\end{aligned}
$$

where $\{y_{i,k}, i = 1, \ldots, n \text{ and } \mathrm{k} = 1, \ldots, \mathrm{K}\}$ are transformed multivariate Gaussian variables. $G^{-1}(\cdot)$
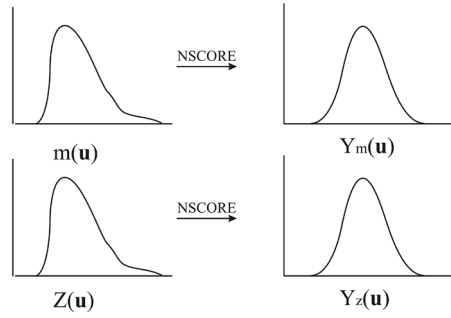
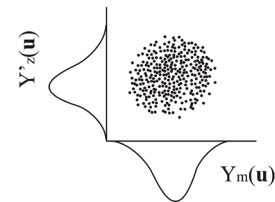**(a)** Normal score transform the exhaustive trend component m(**u**) to get $Y_m(\mathbf{u})$
Normal score transform data component Z(**u**) to get $Y_z(\mathbf{u})$
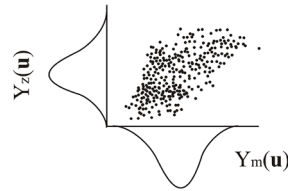
(b) Crossplot of the normal scored data $Y_z(\mathbf{u})$ and trend component $Y_m(\mathbf{u})$
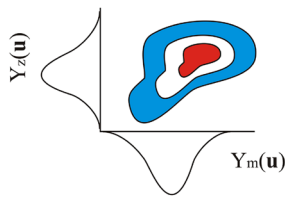
**(c)** Fit the crossplot by the multivariate probability density distribution

(d) Data $Y_z(\mathbf{u})$ are transformed conditional to the fitted Gaussian mixture model to get $Y'_z(\mathbf{u})$

Figure 1. Overview of the proposed approach: **(a)** normal score transform the trend model and data individually; **(b)** crossplot of the transformed data and the co-located trend; **(c)** fit with a Gaussian mixture model; and **(d)** transform the normal scored data with the Gaussian mixture model.

represents the inverse Gaussian cumulative distribution, and $F(\cdot)$ indicates a cumulative distribution function derived from the data. The co-located transformed values are independent although there is no guarantee of decorrelation at nonzero lag distances (Leuangthong and Deutsch 2003). The co-located complex features are removed in the for-

ward transformation and are brought back in the back transformation. The transformed variables are simulated, then back transformed in reverse order.

The original stepwise proposal considered nonparametric conditional distributions for the transformation. This approach, however, suffers from artifacts due to the bins used for the condi-
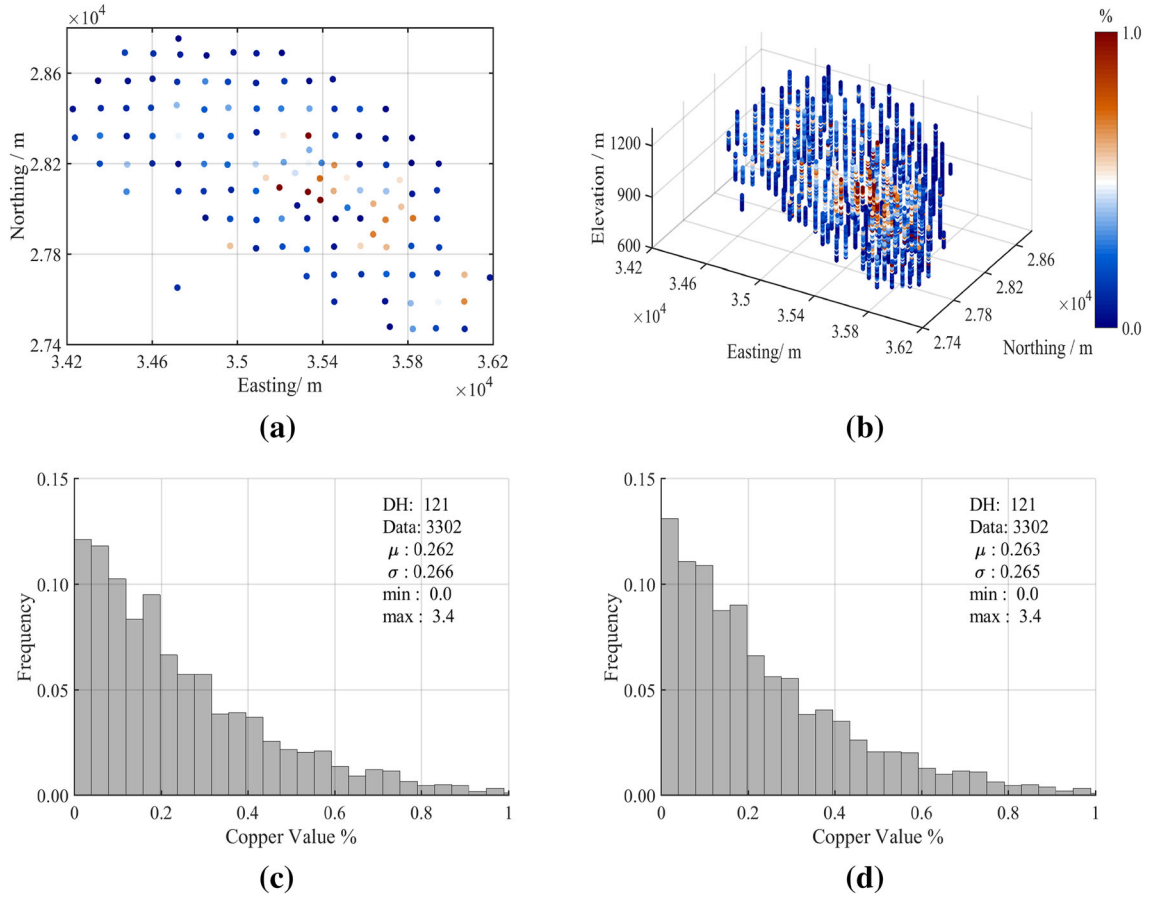
Figure 2. The location maps and histograms of 3302 data. (**a**) Location map in a top view, (**b**) Location map in a 3D view, (**c**) Histogram with raw data and (**d**) Histogram with despiking.
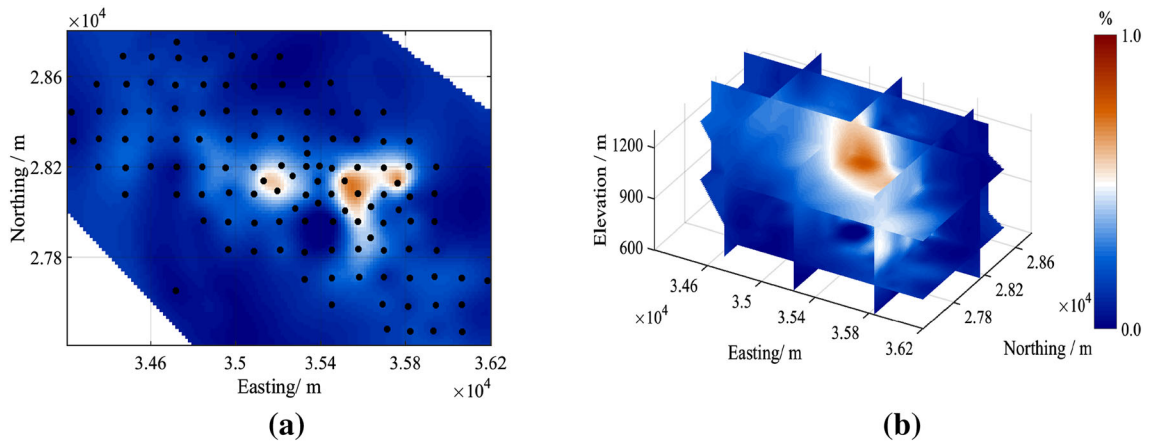
Figure 3. Visualization of the global kriging estimates of 3302 data. (**a**) Estimation at 775.5 meters and (**b**) Estimation in a 3D view.
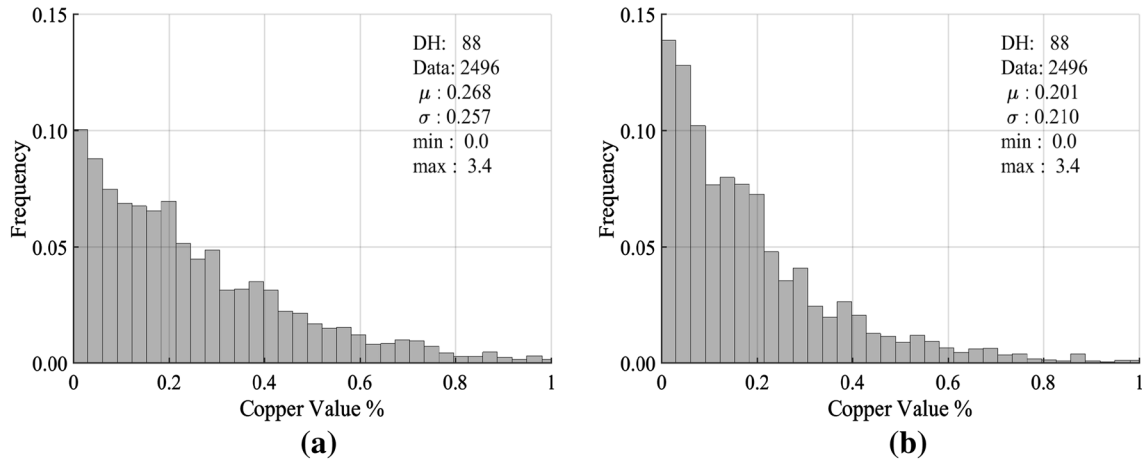
Figure 4. Histograms of 2496 data with different weights. (**a**) Histogram with despiking and (**b**) Histogram with declustering.
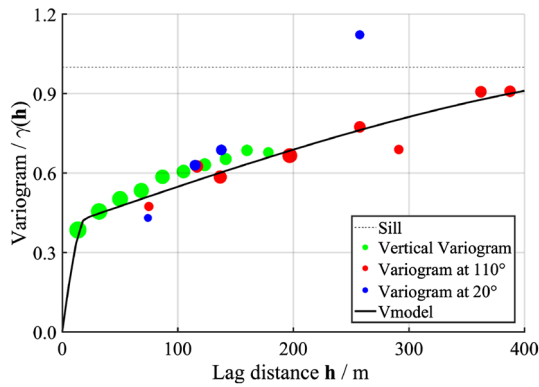

Figure 5. The isotropic variogram model of 2496 data in normal score units. The sizes of the dots represent the relative number of pairs in each direction.

tional distributions and becomes difficult to apply with more than $K = 3$ variables; there are rarely enough data to reliably inform the conditional distributions.

The Gaussian distribution is fully parameterized by a mean vector and a covariance matrix. One single Gaussian model cannot capture all the complex features of geological data, while one Gaussian kernel per observation is computationally expensive with a large number of data (Silverman 1986; Gray and Moore 2003). Mixture models with a small number of Gaussian kernels could be considered. Pearson (1894) proposed the initial approach of mixture models. A number of authors including Gilardi et al. (2002) and Silva and Deutsch (2016) have used them in geostatistics. The expectation

maximization algorithm is considered to fit a Gaussian mixture model (McLachlan and Peel 2004; McLachlan and Krishnan 2007; Silva and Deutsch 2016). The benefits of Gaussian mixture models are that complex features can be captured and any conditional distribution can be easily calculated.

Consider the same set of K variables at n data locations $z_k = \left\{ z_{1,k}, \ldots, z_{n,k} \right\}^T$. $y_k = \left\{ y_{1,k}, \ldots, y_{n,k} \right\}^T$ represents the set of the normal score transformed observations where each variable is transformed independently. The Gaussian mixture model is a multivariate probability density function. The probability density function is written as a sum of g components or mixtures:

$$f^{'}(y_k; \boldsymbol{\Psi}) = \sum_{j=1}^{g} \pi_j \phi \left(y_k; \boldsymbol{\mu}_j; \boldsymbol{\Sigma}_j \right) \quad k = 1, \ldots, K \quad (2)$$

here $f^{'}(\cdot)$ is the estimated distribution. $\boldsymbol{\Psi}$ is the set of unknown parameters $\left\{ \pi_1, \ldots, \pi_g, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_g \right\}$. $\left\{ \pi_1, \ldots, \pi_g \right\}$ are the nonnegative weights assigned to each mixture, $\left\{ \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_g \right\}$ indicates the mean vector of all variables and $\left\{ \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_g \right\}$ refers to the set of covariance matrices between variables for each mixture. $\phi(\cdot)$ is the multivariate Gaussian probability density function. The expectation maximization algorithm maximizes the log likelihood, $\log\{L(\boldsymbol{\Psi})\}$. The parameters of the mixtures would be iteratively fitted so that $f^{'}(\cdot)$ closely fits the experimental data. Any marginal or conditional distribution is easy to compute once $f^{'}(\cdot)$ is fit.
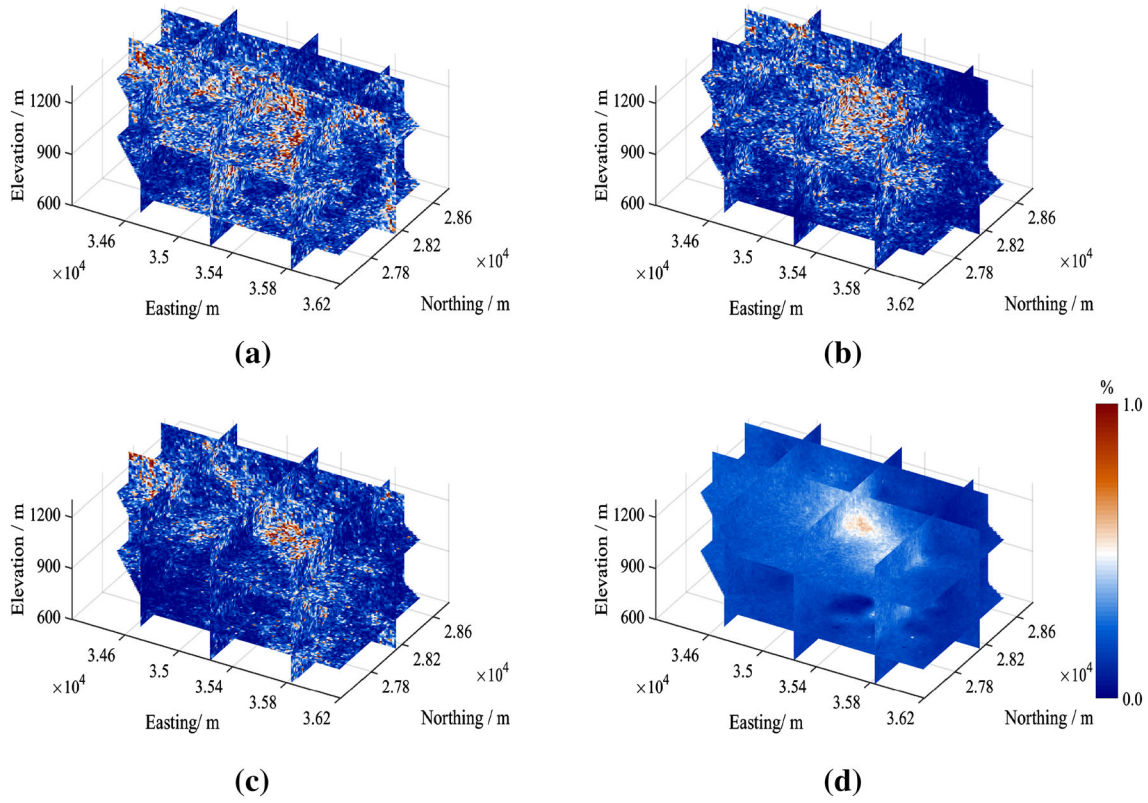
Figure 6. The first three realizations and the average over one hundred realizations in original units. (**a**) Realization 1, (**b**) Realization 2, (**c**) Realization 3 and (**d**) Etype.
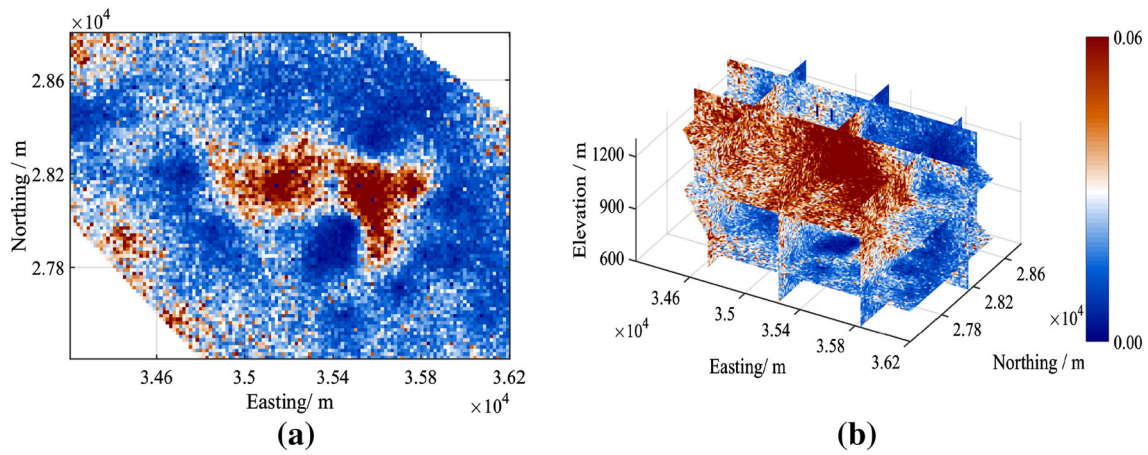

Figure 7. The variance over one hundred realizations in original units. (**a**) A slice at 775.5 meters and (**b**) A volume in a 3D view.

PROPOSED METHOD

The stepwise conditional transformation proposed by Leuangthong and Deutsch (2003) trans-

forms the residuals from the trend conditional to the trend. This approach has binning artifacts due to the nonparametric conditional distributions and creates a small number of negative estimates due to varia-
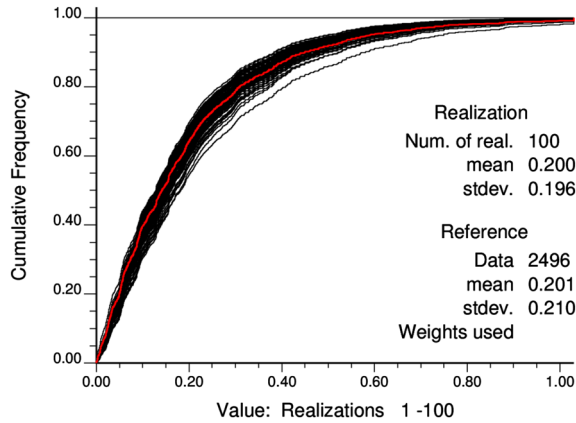
Figure 8. The histogram reproduction of 2496 data with the conventional geostatistical modeling in original units.
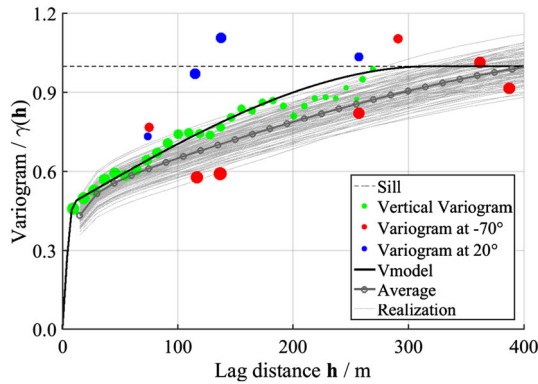
Figure 9. The variogram reproduction of 2496 data with the conventional geostatistical modeling in original units. Directional experimental variograms are plotted with points. Light gray lines are the variograms of each realization, while the dark gray line represents the average of all realizations. The black line is the directional variogram from the original 2496 values.

tions within the bins. A revised methodology is proposed. The first change is to transform the variable conditional to the trend, not the residual conditioned to the trend. The second change is to use a Gaussian mixture model to avoid any binning artifacts. The objective is to remove the trend-like features from data in a bin-free manner that accounts for the spatial structure and multivariate relationship between the data and the trend.

Consider a set of n observations, $\{z_i, i = 1, \ldots, n\}$. The trend is assumed exhaustive and known that it is represented by $\{m_i, i = 1, \ldots, N\}$. Figure 1 shows a schematic illustration of the proposed transformation sequence.

The steps for the stepwise conditional transformation using the Gaussian mixture model are as follows:

1. Normal score transformation: the trend and data are transformed into standard normal score units through the normal score transformation individually. The trend model is exhaustive, and there is no need to consider the declustering, while the data should be transformed with declustering weights if they are unequally sampled. The normal score transformations are written as:

$$y_{m_i} = G^{-1}(F_m(\boldsymbol{m_i})) \quad i = 1, \ldots, N$$
$$y_{z_i} = G^{-1}(F_z(\boldsymbol{z_i})) \quad i = 1, \ldots, n$$
(3)

here $\{y_{m_i}, i = 1, \ldots, N\}$ denotes the Gaussian transformed trend value. Such trend values are known everywhere. N is the number of grid nodes from the exhaustively sampled trend. $F_m(\cdot)$ represents the cumulative distribution function of the exhaustive trend. $\{y_{z_i}, i = 1, \ldots, n\}$ is the Gaussian transformed data, while $F_z(\cdot)$ represents its cumulative distribution function. n represents the number of data where $n \leq N$.

2. Review the transformed variables: the transformed data and the co-located transformed trend are crossplotted. This crossplot is used to help choose the number of Gaussian mixture components, g, for the bivariate fitting and conditional transformation. Too many mixture models will over-fit the complexity of the data, while too few mixture models would fail to reproduce the important complexity. The number of the Gaussian mixture models should be reasonable, such that it gives reliable conditional distributions from the bivariate distribution of the data and the trend in an artifact-free fashion. It is common to choose between 2 and 5.

3. Multivariate density estimation: the expectation maximization algorithm is considered to fit the bivariate distribution of the transformed variables. The estimated multivariate density function is calculated by Eq. 2.

4. Conditional transform the normal score data: the normal score data, $\{y_{z_i}, i = 1, \ldots, n\}$, are transformed by the conditional distribution of y_{z_i} given $\{y_{m_i}, i = 1, \ldots, n\}$. The equation is given as:

(a)

(b)
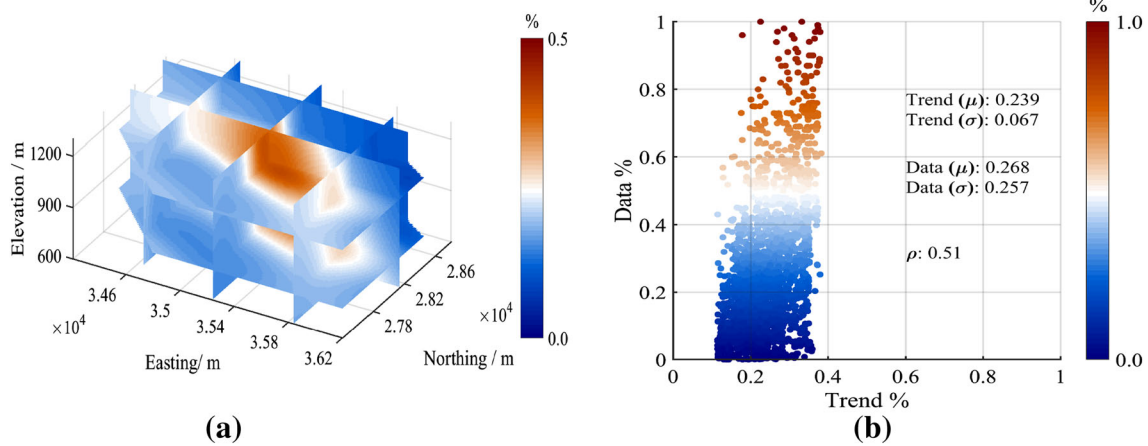
Figure 10. The trend model and the scatter plot of 2496 data in original units. (**a**) Trend model and (**b**) Scatter plot.

(a)

(b)
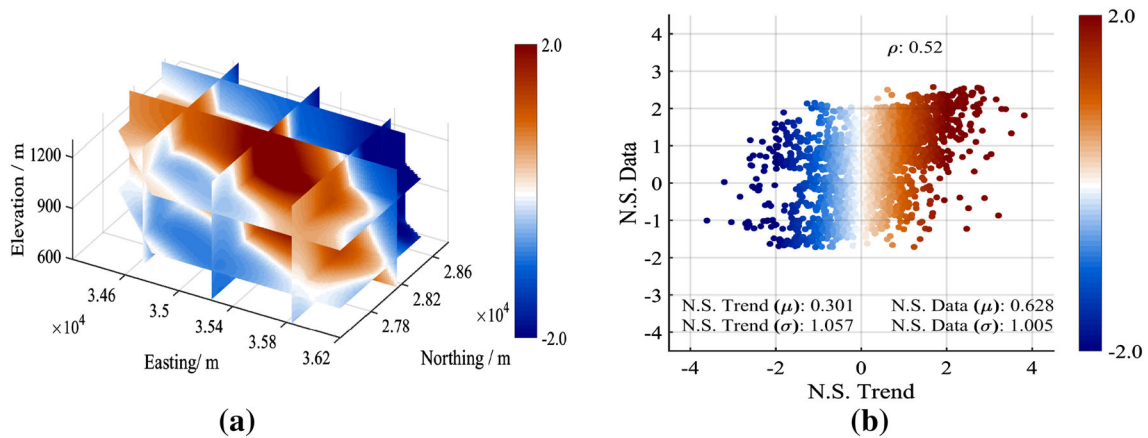
Figure 11. The transformed trend model and the scatter plot of 2496 data in normal score units. (**a**) Trend model and (**b**) Scatter plot.

$$y'_{z_i} = G^{-1}\left(F_{z|m}(y_{z_i} \mid y_{m_i})\right) \quad i = 1, \ldots, n$$

where the random variable y'_{z_i} indicates the transformed data by the Gaussian mixture models. $F_{z|m}(\cdot)$ represents the cumulative distribution function of the data given the exhaustive trend. The cumulative distribution function of the trend does not enter any calculations, but the transformed data consider the trend at each location. The bivariate distribution of the transformed data and the co-located normal score trend has no correlation.

The proposed parametric conditional transformation removes the trend-like features that

may be problematic in the modeling of the raw data directly. Gaussian simulation can now be used and several realizations are generated with the transformed data. The back transformation will ensure that the trend model is used everywhere. The trend is reproduced in original units.

APPLICATION

The data shown in Figure 2a (top view) and b (3D view) comprise 121 drillholes with 3302 grade measurements from a porphyry copper deposit. The location coordinates range from 34, 200 to 36, 200 meters East, from 27, 400 to 28, 800 meters North,
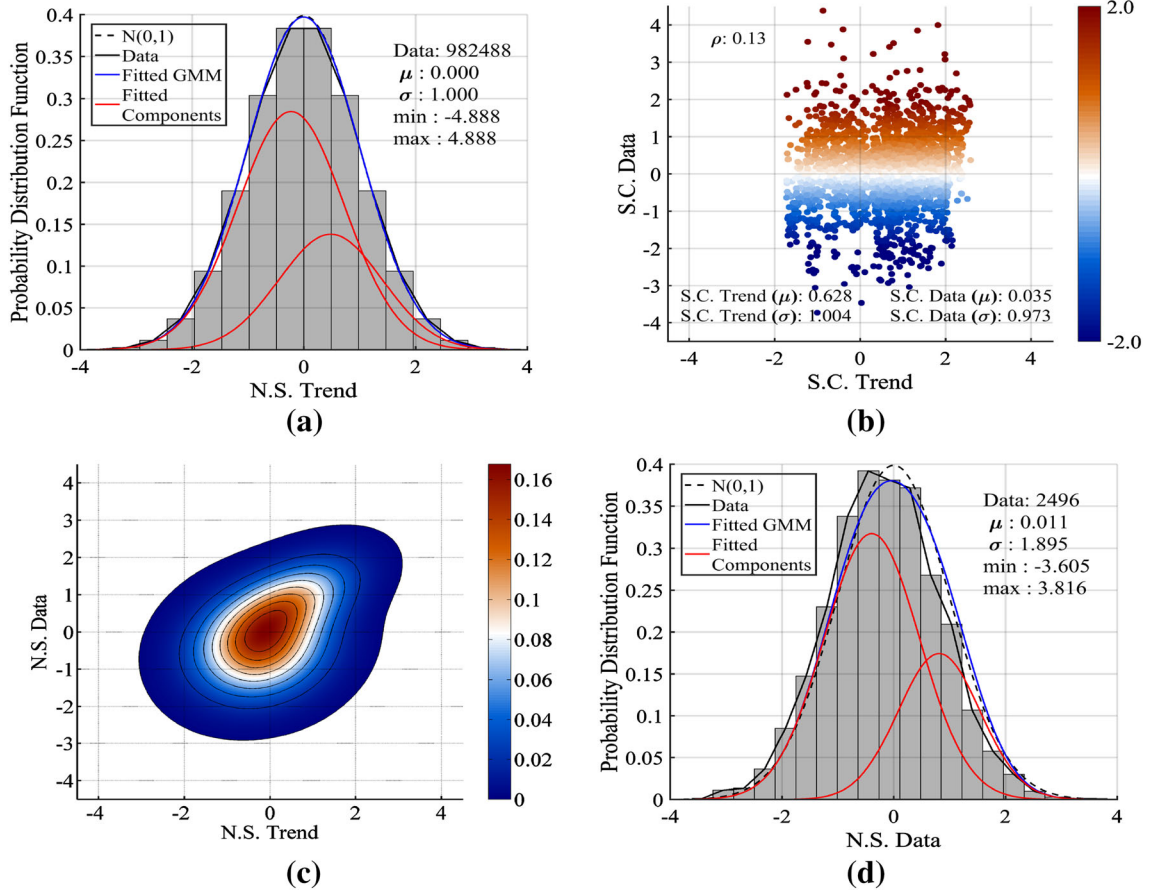
Figure 12. The bivariate and univariate distributions of the Gaussian mixture model. (**a**) Univariate distribution of the exhaustive trend, (**b**) Scatter plot of transformed variable, (**c**) Bivariate distribution and (**d**) Univariate distribution of data.
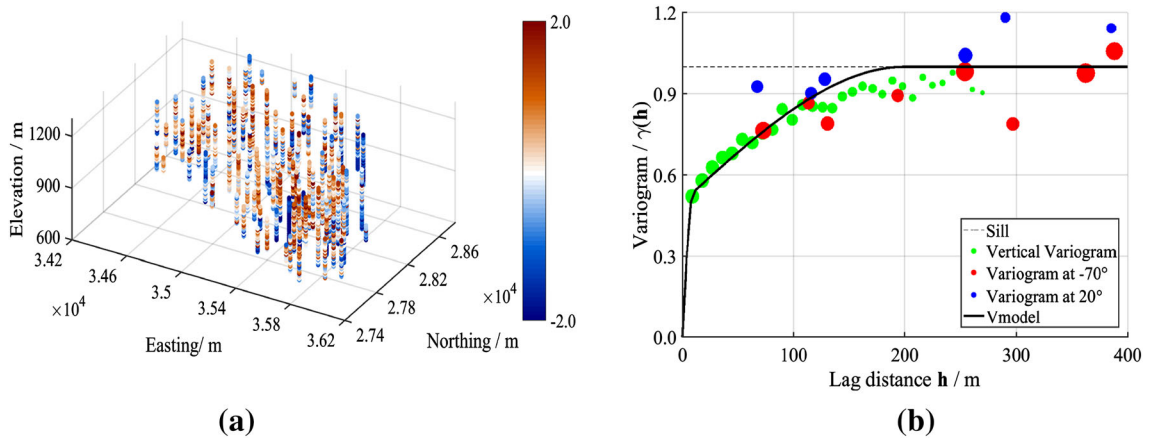
Figure 13. The location map and the variogram model of 2496 data in stepwise units. The sizes of the dots in the variogram model represent the relative number of pairs in each direction. (**a**) Location map and (**b**) Variogram model.
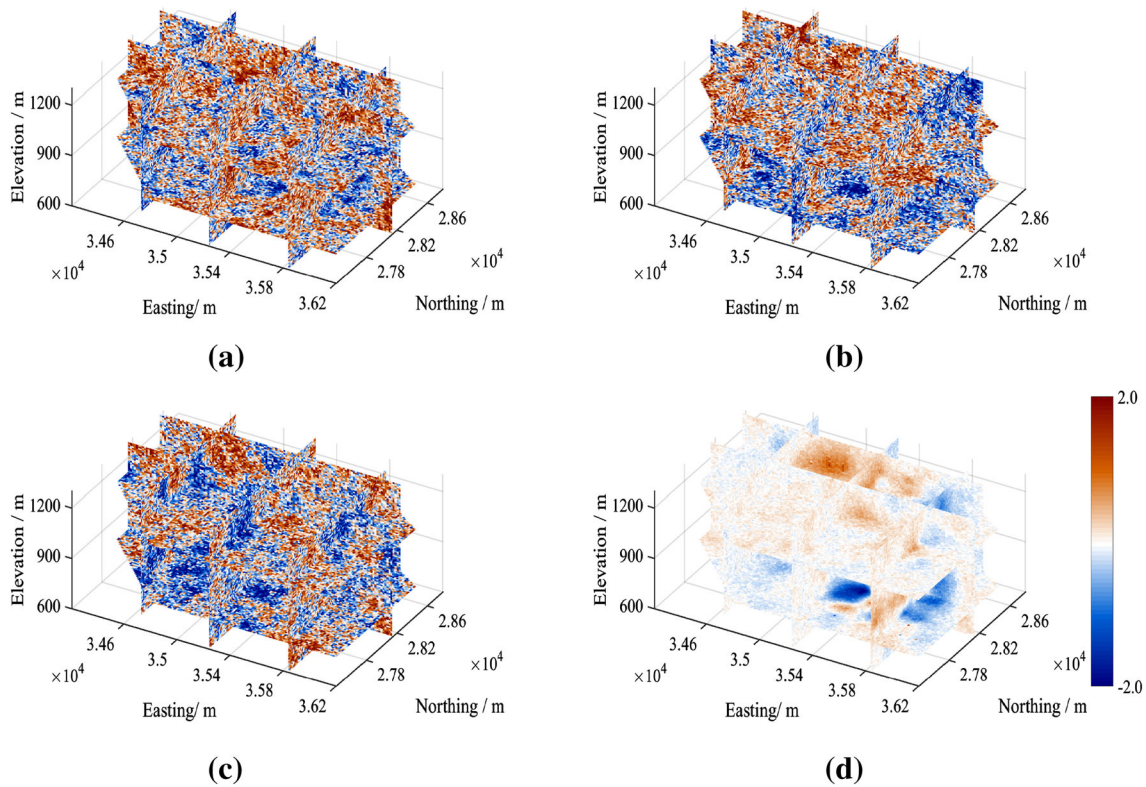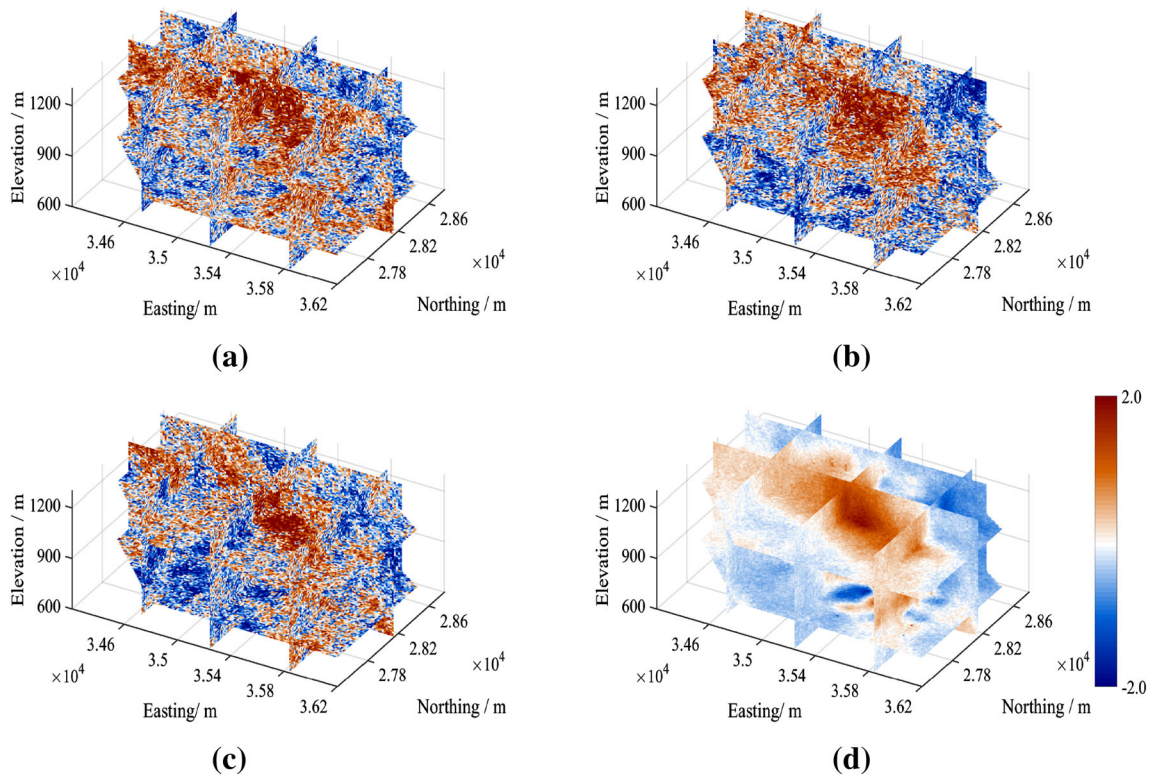
Figure 14. The first three realizations and the average over one hundred realizations in stepwise units. (**a**) Realization 1, (**b**) Realization 2, (**c**) Realization 3 and (**d**) Etype.

and from 600 to 1, 300 meters Elevation with 9 meters intervals. The grade of copper ranges from 0.0 to 3.4% with a mean of 0.262% and a standard deviation of 0.266% and the histogram is shown in Figure 2c. Despiking was considered due to grades of constant values (Rossi and Deutsch 2014). The histogram after despiking is shown in Figure 2d. The mean is 0.263% and the standard deviation is 0.265%.

Although kriging strongly depends on the decision of stationarity, it can still be used for mapping the large-scale trend-like features. A global kriging was performed with a variogram with a 20% nugget effect and a range of 1000 meters. The global kriging result in Figure 3 reveals the obvious trend where high values are concentrated in the center. The most continuous direction in the plane direction is at an azimuth of 110°.

The 3302 copper grade data from 121 drillholes were divided randomly into a modeling set of 2496 data from 88 drillholes and a test set of 806 data from 33 drillholes. The modeling data were used for geostatistical modeling, and the test data were used to check the simulated results.

The histogram of the modeling 2496 data is shown in Figure 4a. Data are clustered together, so the declustering was needed (Deutsch and Journel 1998). Cell declustering was performed with a 400-meter cell size and the corrected histogram with a mean of 0.201% and a standard deviation of 0.210% is shown in Figure 4b. The data were transformed into a normal distribution with the declustering weights. The directional variograms in normal score units were plotted with an isotropic variogram model in Figure 5. Sequential Gaussian simulation was considered and 100 realizations were generated. A normal score back transformation was considered to bring all realizations back to the original units. Figure 6 shows the back transformed results with the first three realizations and the average of all 100 realizations in original units. The variance of all 100 values at each location is shown in Figure 7. There is low variance in the low-valued zones and high variance in high-valued zones as expected with a positively skewed distribution. The variance is high around the margins because of the few conditioning data.
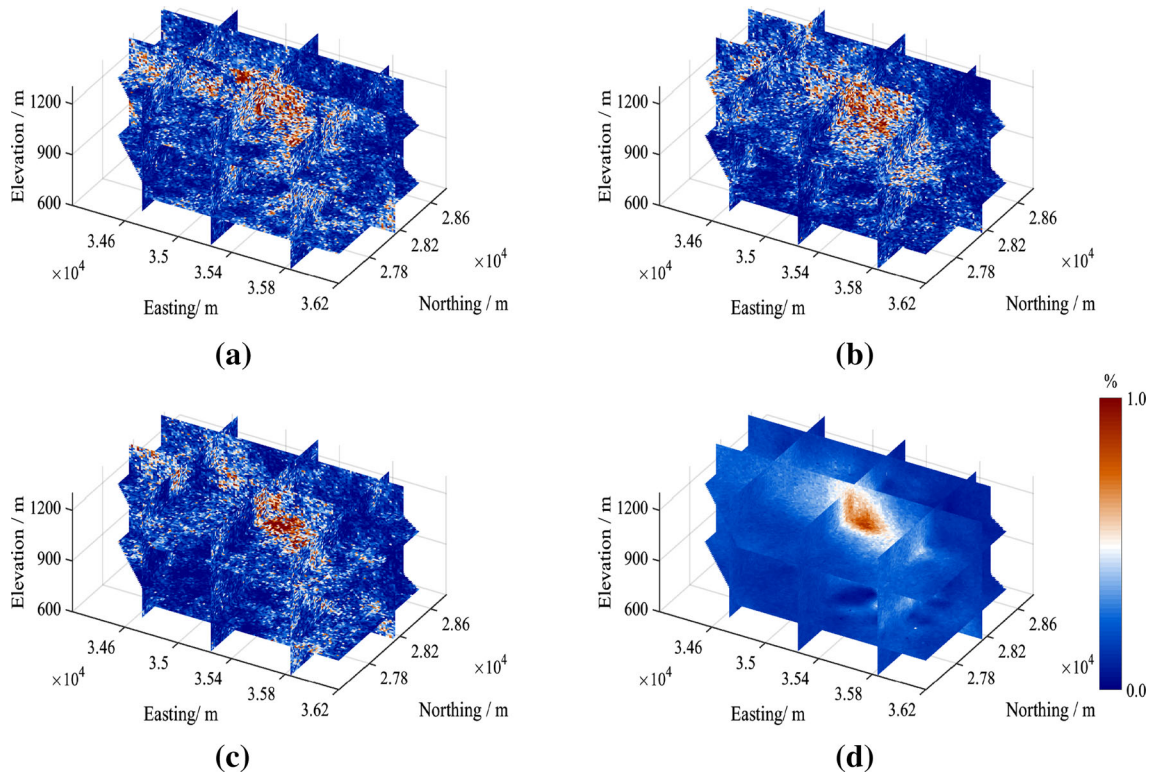
Figure 15. The first three realizations and the average over one hundred realizations in normal score units. (**a**) Realization 1, (**b**) Realization 2, (**c**) Realization 3 and (**d**) Etype.

Histogram reproduction can be checked that geostatistical realizations are intended to reproduce the input histogram. The histograms of 100 realizations are modeled and shown with *black lines*, while the 2496 conditioning data are shown with a *red line* in Figure 8. The mean of the realizations, 0.200%, is close to the reference mean, 0.201%; the standard deviation, 0.196%, is lower than the conditioning variance, 0.210%. The realizations successfully reproduce the global mean and the global distribution of the data. The histogram reproduction appears reasonable. Variogram reproduction should be theoretically honored in simulation that checks the spatial correlation in the final model. Figure 9 shows the variogram reproduction in original units. The variograms of all realizations are slightly more continuous than the original isotropic variogram.

The proposed methodology was implemented with a trend model. The trend model is constructed to avoid under- or over-fitting to the data. The trend model contains the large-scale variability and is shown in Figure 10a. The scatter plot is shown in Figure 10b, indicating that the correlation between

the trend and the data is 0.51. The exhaustive trend model was transformed into normal score units, while the 2496 data were transformed into normal score units with the declustering weights independently. Figure 11 shows the transformed results indicating a direct relationship with a correlation of 0.52 between the trend and the data in normal score units. The stepwise conditional transformation with a Gaussian mixture model was considered to remove the complexity of the data. The decision of the number of mixture components is subjective. In this case study, two components were determined by visual inspection to fit the scatter plot. Figure 12 shows the Gaussian mixture model. The univariate distributions of the trend model and the data are shown in Figure 12a and d, respectively. The marginal distributions from the Gaussian mixture models are not exactly normal; however, the deviation appears to be very small in Figure 12a where the combined mixture distribution and an exact normal distribution are almost perfectly overlapping. The bivariate distribution is shown on a 2D probability density plot in Figure 12c. The transformed vari-

Figure 16. The first three realizations and the average over one hundred realizations in original units (**a**) Realization 1, (**b**) Realization 2, (**c**) Realization 3 and (**d**) Etype.
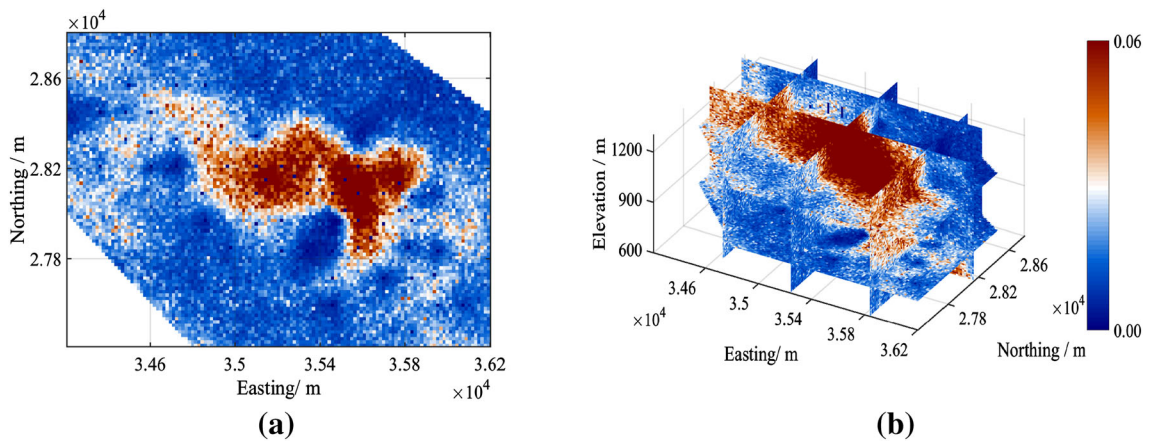
Figure 17. The variance over one hundred realizations in original units. (**a**) A slice at 775.5 m and (**b**) A volume in a 3D view.

ables are uncorrelated (Fig. 12b). The data after the stepwise conditional transformation in Figure 13a show a randomness, and the trend is removed. The directional variograms in stepwise units are fitted with an isotropic variogram model and shown in Figure 13b. Sequential Gaussian simulation was conducted on the transformed variable. Figure 14 shows the first three realizations and the average of 100 realizations. No trend-like features exist in the simulated results. A stepwise conditional back transformation with the trend was performed. Figure 15 shows the first three realizations and the
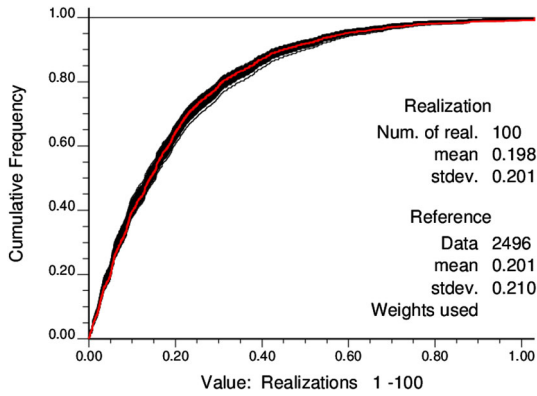
Figure 18. The histogram reproduction of 2496 data with the proposed geostatistical modeling in original units.
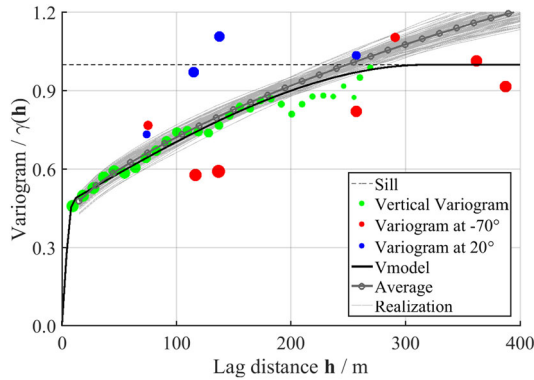

Figure 19. The variogram reproduction of 2496 data with the proposed geostatistical modeling in original units. Directional experimental variograms are plotted with points. Light gray lines are the variograms of each realization, while the dark gray line represents the average of all realizations. The black line is the directional variogram from the original 2496 values.

average of 100 realizations in normal score units. The simulated results show that the trend-like features are restored from the back transformation. The initial normal score transformation was also reversed. Figure 16 shows the first three realizations and the average of 100 realizations in original units. The local variance is calculated and shown in Figure 17. The map shows the high variance in the central and low variance around the margins.

The histogram of realizations must be consistent with the histogram of 2496 conditioning data. The realizations over all locations are considered. The histogram is reasonably reproduced in original units (Fig. 18). The mean over 100 realizations is 0.198%. The value is slightly lower than the conditioning mean, 0.201%. The standard deviation is

0.201%, which is lower than that of the conditioning data, 0.210%, but it is higher than that of 0.196% in the conventional method. Figure 19 shows the variogram reproduction. The overall variogram reproduction from the realizations appears better than that from the conventional method in Figure 9.

The first validation step was to compare 806 true values with the simulated average values in normal score units. The test data were transformed into a normal distribution with the reference distribution of 2496 data. The locations of the test data were labeled with the drillhole IDs and shown in Figure 20. The distributions of the local uncertainty were specified by a conditional mean and variance in normal score units. The plot in Figure 21 shows the accuracy of the simulated distributions of the uncertainty with the conventional method and the developed method in normal score units. The mean of the variance over 100 realizations from 806 checking locations represents the local uncertainty of the model. The local uncertainty is 0.589 that is underestimated by the conventional method, while the local uncertainty is fair, 0.742, with the developed method. It highlights that the numerical model with the developed method contains more variance than the model with the conventional method due to the values with the conventional method is smooth and close to the global mean. The accuracy of the developed method is better than that of the conventional method.

The second validation step was to compare 806 true values with the simulated average values with mean squared error values in original units. Figure 22 shows the location maps of the test data labeled with drillhole IDs. The mean of the developed method, 0.247%, is close to the true mean, 0.247%. The standard deviation of the average measures the smoothing effect. The standard deviation of the average values with the developed method, 0.128% contains more variability than that with the conventional method, 0.108%. The mean squared error value measures the difference between the truth and what is being estimated and, further, summarizes the prediction performance. The minimized mean squared error is used to identify the best method for modeling with a trend. The mean squared error values between the true values and the average values are 0.0631 and 0.0618, respectively. It shows a 2.06% improvement in the developed method. Three drillholes extracted from high-, medium- and low-valued zones are compared and shown in Figure 23, indicating 8.53, 2.89, and 46.21% improvements,
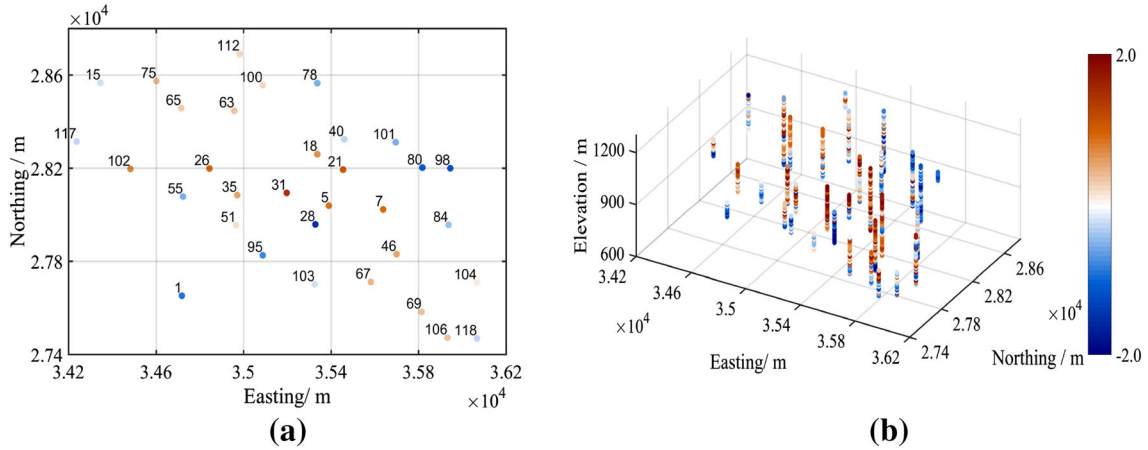
Figure 20. The location maps with the test data in normal score units. (**a**) Location map in a top view and (**b**) Location map in a 3D view.
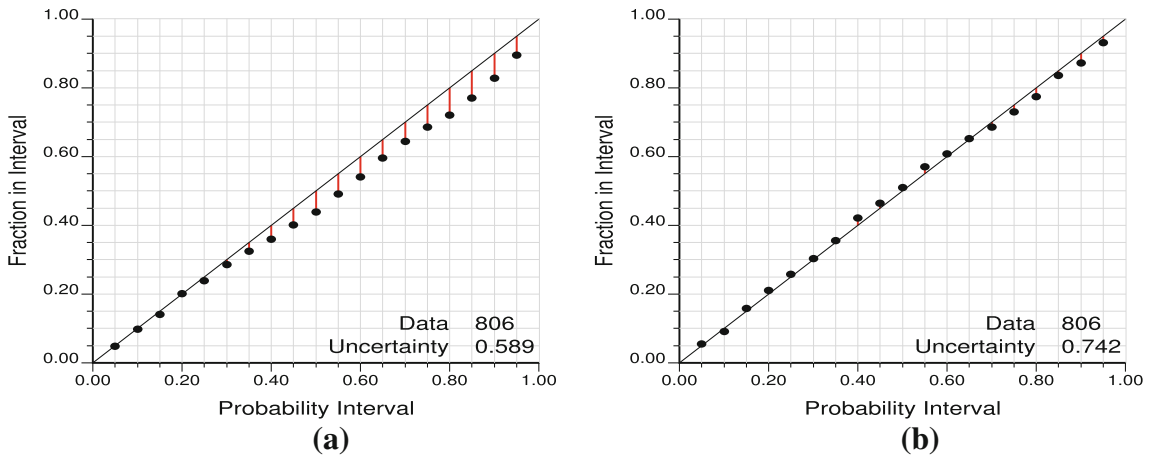
Figure 21. The cross-validations of the test data with different methods in normal score units. The grid of light lines shows the probability intervals, while the red lines and bullets show the deviations of the actual proportions from the predicted probability intervals (Deutsch 2010). (**a**) Conventional method and (**b**) Developed method.
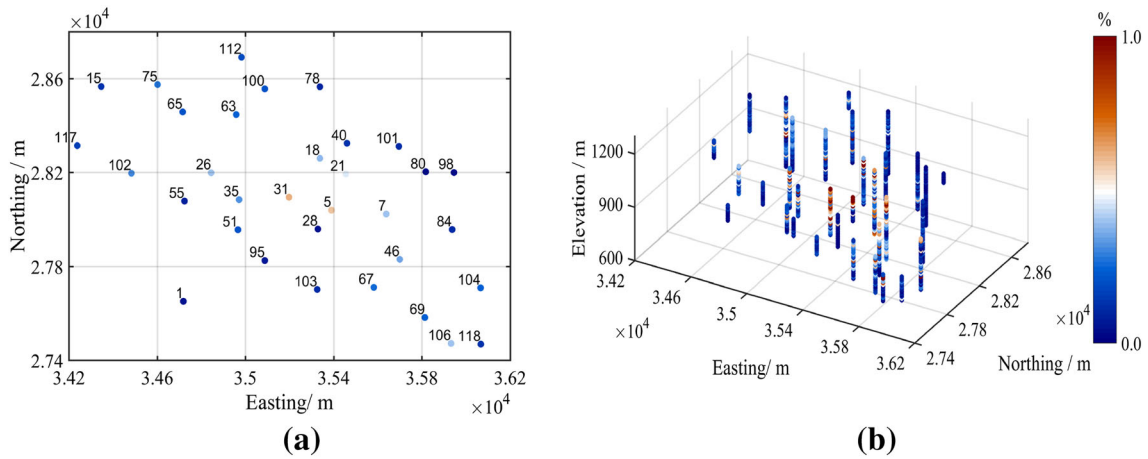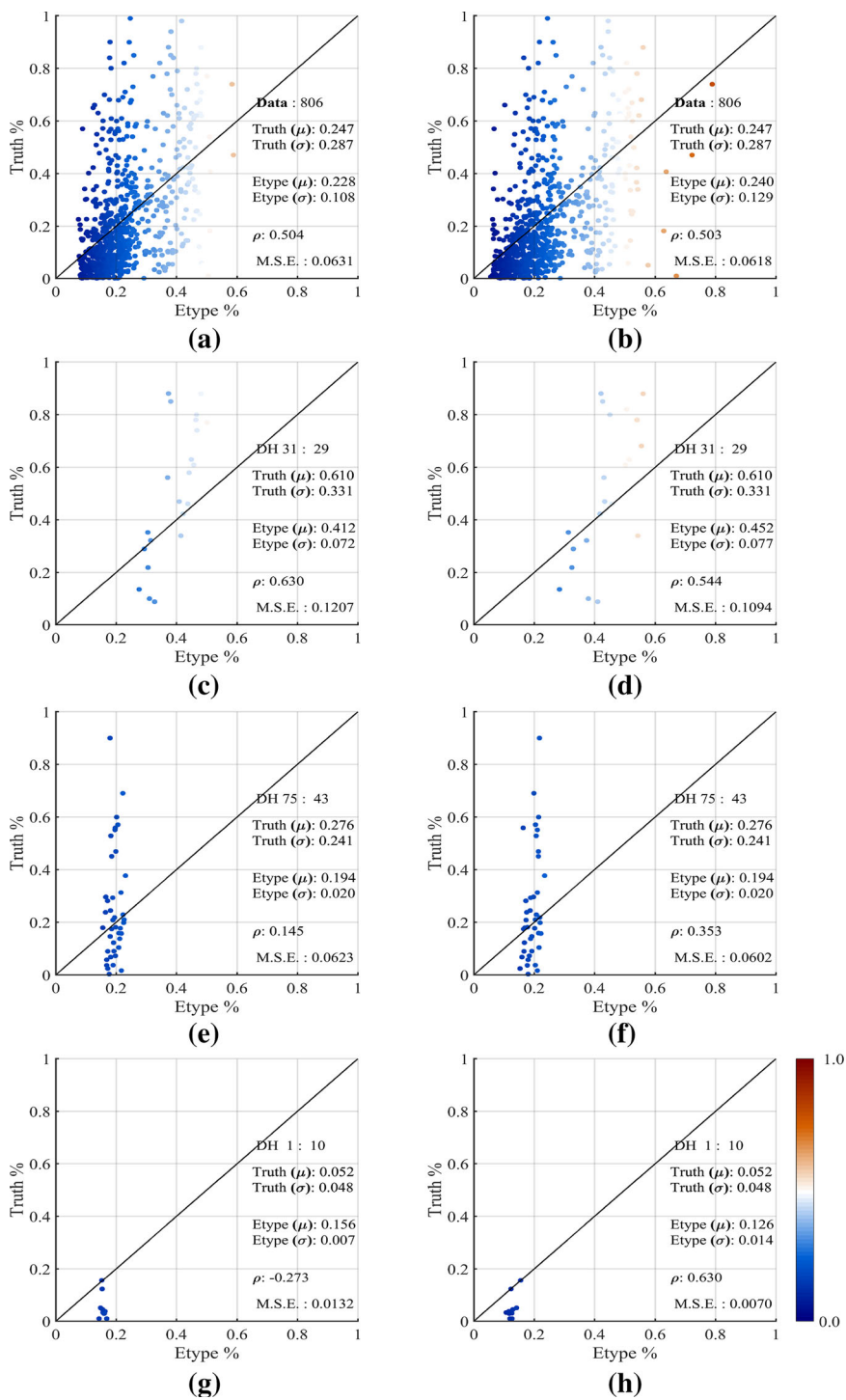

Figure 22. The location maps with the test data in original units. (**a**) Location map in a top view and (**b**) Location map in a 3D view.

Figure 23. The Comparisons from drillholes with different methods in original units: (**a**, **c**, **e**, **g**) the mean squared error value between true data and simulated values with the conventional method; and (**b**, **d**, **f**, **h**) the mean squared error value between true data and simulated values with the proposed method. (**a**) 806 data from 33 DHs, (**b**) 806 data from 33 DHs, (**c**) 29 data from DH 31, (**d**) 29 data from DH 31, (**e**) 43 data from DH 75, (**f**) 43 data from DH 75, (**g**) 10 data from DH 1 and (**h**) 10 data from DH 1.

respectively. The developed method shows a significant improvement.

DISCUSSION

A practical framework for non-stationary geostatistical techniques using a Gaussian mixture model was established. The data were divided into a modeling set and a test set. The modeling set was used for proceeding geostatistical modeling, and the test data were used for checking the results. The assumption of stationarity is made in the conventional geostatistical prediction and relaxed in the developed method. The proposed method is more accurate but with greater uncertainty. The mean squared error comparisons show a modest yet important 2.06% improvement in the developed method. Drillholes close the margins of the deposit show the greatest improvement.

A significant assumption in the case study is that the trend model is assumed optimal and known. The trend model is a part of characterizing the natural resources. The uncertainty in trend model is ignored so that the overall uncertainty might be underestimated. Data with an apparent trend were transformed conditional to the trend that the trend is important in the stepwise transformation. The parameterization and optimization of the trend is an important area of future work.

Another assumption is that two components for the Gaussian mixture modeling are optimal. A visual inspection is a common approach, but this decision is subjective and depends on the practitioner. A criterion for the number of Gaussian mixtures should be proposed in future research.

The approach of modeling with residuals using Gaussian mixture models and adding the trend model back in final models, that is, $R(\boldsymbol{u}) = Z(\boldsymbol{u}) - m(\boldsymbol{u})$ then $Z(\boldsymbol{u}) = m(\boldsymbol{u}) + R(\boldsymbol{u})$, was also implemented. The mean squared error value between the truth and the simulated results is 0.0639 in Figure 24, indicating a 3.29% loss. The performance of modeling with residuals using Gaussian mixture models was not as good as the proposed method, which models the data more accurately. In addition, the constraint for nonnegative simulated values ($Z(\boldsymbol{u}) \geq 0.0$) is not required in the proposed method.

The improvements of the stepwise conditional transformation with Gaussian mixture model still exist. The covariance after the stepwise conditional
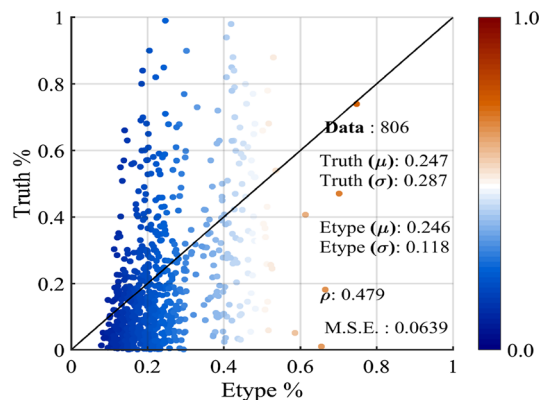
Figure 24. The mean squared error value of 806 data from 33 drillholes in the final model with the naive approach of modeling with residuals using Gaussian mixture models and adding the trend model back in final models.

transformation is zero at the lag distance $\boldsymbol{h} = 0$ and may not be zero at other lag distances that could affect the result (Leuangthong and Deutsch 2003). The use of minimum/maximum autocorrelation factors (MAF) (Desbarats and Dimitrakopoulos 2000) may be considered on the transformed variables if remnant cross-spatial correlation is present. A MAF could assist with variogram fitting, and further, it could help a better performance of mixture models and lead to a better result.

Multiple non-stationary variables could also be considered simultaneously in a hierarchical workflow. Each variable could be processed according to the proposed workflow in Figure 1, and then another Gaussian mixture model could be fit to the detrended variables. A second stepwise conditional transform would remove the dependency between the variables. Gaussian simulation of the independent factors would proceed; then, the back transformation would be performed in reverse order to account for multivariate dependencies and the nonstationary trend models.

CONCLUSION

Geostatistics has been used for predicting spatial variability. Geostatistical methods depend on stationary statistics. Real geological data often exhibit trend-like features that represent the large-scale variability of the regionalized variable. The assumption of stationarity is not satisfied with the variable in presence of trends. The trend should lead to more accurate estimates than if the trend is ignored.

A modified stepwise conditional transformation for geostatistical modeling is proposed. Data with an apparent trend were transformed conditional to the trend by a parametric transformation. The use of the Gaussian mixtures removes the trend-like features from the regionalized variable, eliminates the artifacts from the data binning of the conventional stepwise conditional transformation, and brings more variation to numerical models. The improved performance of the geostatistical algorithm is attributed to the stationarity of the transformed result.

A real dataset with an obvious trend was used to demonstrate the proposal. Comparisons between the conventional prediction and the developed prediction were made. The performances of numerical models, the reproduction of geological characterizations and the analysis of the local uncertainty were compared. The case study shows that the geostatistical modeling with trend modeling performs better than conventional geostatistical modeling, especially around the margins of the domain.

ACKNOWLEDGMENTS

REFERENCES

Boisvert, J. B., Rossi, M. E., Ehrig, K., & Deutsch, C. V. (2013). Geometallurgical modeling at Olympic dam mine, South Australia. *Mathematical Geosciences*, *45*(8), 901–925.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (2002). Geographically weighted summary statistics—a framework for localised exploratory data analysis. *Computers, Environment and Urban Systems*, *26*(6), 501–524.

Chiles, J. P., & Delfiner, P. (2012). *Geostatistics: Modeling spatial uncertainty* (2nd ed.). New York: Wiley.

Davis, J. C., & Sampson, R. J. (2002). *Statistics and data analysis in geology* (3rd ed.). New York: Wiley.

Delfiner, P. (1976). Linear estimation of non-stationary spatial phenomena. In M. Guarascio, M. David, C. J. Huijbregts (Eds.), *Advanced Geostatistics in the Mining Industry*. NATO Advanced Study Institutes Series (Series C — Mathematical and Physical Sciences) (Vol. 24, pp. 49–68). Dordrecht: Springer.

Desbarats, A. J., & Dimitrakopoulos, R. (2000). Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors. *Mathematical Geology*, *32*(8), 919–942.

Deutsch, C. V. (2010). Display of cross validation/jackknife results. *Centre for Computational Geostatistics Annual Report*, *12*(406), 1–4.

Deutsch, C. V., & Journel, A. G. (1998). *GSLIB: Geostatistical software library and user's guide* (2nd ed.). New York: Oxford University Press.

Gelfand, A. E., Schmidt, A. M., Banerjee, S., & Sirmans, C. F. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization. *Test*, *13*(2), 263–312.

Gilardi, N., Bengio, S., & Kanevski, M. (2002). Conditional Gaussian mixture models for environmental risk mapping. In *Proceedings of the 2002 IEEE international workshop on neural networks for signal processing* (pp. 777–786).

Gonzales, E., McLennan, J. A., & Deutsch, C. V. (2006). A new approach to sequential Gaussian simulation with a trend: Non-stationary transformation tables. *Centre for Computational Geostatistics Annual Report*, *08*(120), 1–13.

Goovaerts, P. (2001). Geostatistical modelling of uncertainty in soil science. *Geoderma*, *103*(1), 3–26.

Gray, A. G., & Moore, A. W. (2003). Nonparametric density estimation: Toward computational tractability. In *Proceedings of the 2003 society for industrial and applied mathematics international conference on data mining* (pp. 203–211).

Journel, A. G., & Huijbregts, C. J. (2003). *Mining geostatistics*. London: Academic Press.

Leuangthong, O. (2003). *Stepwise conditional transformation for multivariate geostatistical simulation*. Doctoral dissertation, University of Alberta.

Leuangthong, O., & Deutsch, C. V. (2003). Stepwise conditional transformation for simulation of multiple variables. *Mathematical Geology*, *35*(2), 155–173.

Leuangthong, O., & Deutsch, C. V. (2004). Transformation of residuals to avoid artifacts in geostatistical modelling with a trend. *Mathematical Geology*, *36*(3), 287–305.

Machuca-Mory, D. F. (2010). *Geostatistics with location-dependent statistics*. Doctoral dissertation, University of Alberta.

Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, *5*(3), 439–468.

McLachlan, G., & Krishnan, T. (2007). *The EM algorithm and extensions* (3rd ed.). New York: Wiley.

McLachlan, G., & Peel, D. (2004). *Finite mixture models*. New York: Wiley.

Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London*, *185*, 71–110.

Pyrcz, M. J., & Deutsch, C. V. (2014). *Geostatistical reservoir modeling* (2nd ed.). New York: Oxford University Press.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The annals of mathematical statistics*, *23*(3), 470–472.

Rossi, M. E., & Deutsch, C. V. (2014). *Mineral resource estimation*. Berlin: Springer.

Sabourin, R. (1976). Application of two methods for the interpretation of the underlying variogram. In M. Guarascio, M. David, C. J. Huijbregts (Eds.), *Advanced Geostatistics in the Mining Industry*. NATO Advanced Study Institutes Series (Series C — Mathematical and Physical Sciences) (Vol. 24, pp. 101–109). Dordrecht: Springer.

Sampson, P. D., & Guttorp, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, *87*(417), 108–119.

Silva, D. S. F., & Deutsch, C. V. (2016). Multivariate data imputation using Gaussian mixture models. *Spatial Statistics*. doi:10.1016/j.spasta.2016.11.002.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis* (Vol. 26). Boca Raton: CRC Press.

Wackernagel, H. (2003). *Multivariate geostatistics: An introduction with applications* (3rd ed.). Berlin: Springer.

Wang, G., Carranza, E. J. M., Zuo, R., Hao, Y., Du, Y., Pang, Z., et al. (2012). Mapping of district-scale potential targets using fractal models. *Journal of Geochemical Exploration*, *122*, 34–46.