



Journal of Mining and Environment (JME)

journal homepage: [www.jme.shahroodut.ac.ir](http://www.jme.shahroodut.ac.ir)



## A Practical Comparison between Gaussian and Direct Sequential Simulation Algorithms using a 3D Porosity Dataset

Hamid Sabeti<sup>1,2</sup> and Farzad Moradpouri<sup>3</sup>

1- Department of Mining Engineering, Birjand University of Technology, Birjand, Iran,

2- CERENA, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal

3- Department of Mining Engineering, Faculty of Engineering, Lorestan University, Khoramabad, Iran

### Article Info

Received 1 May 2022

Received in Revised form 22 May 2022

Accepted 1 June 2022

Published online 1 June 2022

DOI: [10.22044/jme.2022.11874.2180](https://doi.org/10.22044/jme.2022.11874.2180)

### Keywords

Geo-statistics

Sequential Gaussian Simulation

Direct Sequential Simulation

Variogram, Porosity

### Abstract

The geo-statistical simulation algorithms for continuous spatial variables have been used widely in order to generate the statistically-honored models. There are two main algorithms doing the continuous variable simulation, Sequential Gaussian Simulation (SGS) and Direct Sequential Simulation (DSS). The main advantage of the DSS algorithm against the SGS algorithm is that in the DSS algorithm no Gaussian transformation of the original data is made. In this work, these two simulation algorithms are explained, and their applications to a 3D spatial dataset are deeply investigated. The dataset consists of the porosity values of 16 vertical wells extracted from an actual cube obtained by a seismic inversion process. One well data is excluded from the simulation process for the blind well test. Comparison between the histograms show that the histogram reproduction is slightly better for the SGS algorithm, although the population reproductions are the same for both SGS and DSS results. The DSS algorithm reproduce the mean of input data closer to the mean of well data compared to that of the SGS algorithm. Considering one realization from each simulation algorithm, the RMS error corresponding to all simulated cells against the real values is approximately equal for both algorithms. On the other hand, the error show a slightly less value when the mean of 100 realizations of the DSS result is considered.

### 1. Introduction

The geo-statistical simulation algorithms have been widely used during the last decades in order to generate the equiprobable models of any spatial variable based on some conditioning data. Different types of variables have led to the development of various simulation algorithms. The most common and traditional algorithm for continuous variables is Sequential Gaussian Simulation or SGS [1]. The application of SGS to different geo-science fields has been recently developed by numerous researchers [2-10]. Sakizadeh *et al.* (2017) have presented a spatial risk assessment of heavy metals using SGS [7]. Talesh Hosseini *et al.* (2018) have used the geo-chemical data to model the zonality elements in the Baghqlloom area of Iran [6]. Wang and Zuo (2018)

have identified the geo-chemical anomalies by a combination of SGS and grid-based local singularity analysis [9]. Gholampour *et al.* (2019) have applied the SGS algorithm to drill the core data to model the alteration zones [10]. Metahni *et al.* (2019) have compared different interpolation methods and SGS algorithms in order to estimate the contaminated soil volumes [8]. Sotoudeh *et al.* (2020) have applied the SGS algorithm to grade values of copper, and have developed a new underground mine design based on grade uncertainty [5]. Shen *et al.* (2021) have compared the SGS algorithm with a positive matrix factorization model for risk assessment of soil heavy metals in the area of the Yellow River, China [3].

SGS uses a Gaussian transformation of original data to perform the simulation. However, Soares (2001) has proposed a new algorithm called Direct Sequential Simulation (DSS), in which no prior and posterior Gaussian transformations of the original variable were required [11]. Many researchers have used the DSS algorithms to model their spatial datasets. Ribeiro *et al.* (2016) have applied direct sequential simulation algorithms to precipitation time series in order to detect inhomogeneities in Portugal [12]. Sabeti *et al.* have used the DSS algorithm to generate different acoustic impedance models for a stochastic seismic inversion method [13]. Pereira *et al.* (2019) have analyzed the impact of a priori elastic models on seismic inversion algorithm using direct sequential simulation [14]. Madenoglu *et al.* (2020) have applied DSS to the soil surface samples to assess the uncertainty of soil erodibilities [4]. Almeida *et al.* (2020) have developed an integrated method for reducing the estimation uncertainty of reservoir properties using the DSS algorithm [15]. Horta *et al.* (2021) have applied the DSS algorithm to build soil contamination maps [16]. Otzen *et al.* (2022) have developed a new spherical linear inversion algorithm using direct sequential simulation [17].

These two main simulation algorithms for modeling many different variables from different fields are now routinely used. The algorithm selection has been always a difficult decision that comes from lacking the comparing research on these two algorithms. In this paper, the SGS and DSS algorithms are briefly described, and a practical comparison between them is presented using a pseudo-real 3D dataset.

**2. Sequential Gaussian Simulation**

Considering the desired variable in a Gaussian field, one can generate the joint distribution of random variables. These variables are then conditioned by a set of data of any type using the notation. The following expression presents the complementary cumulative distribution function (CCFD) of the variables [1]:

$$F_{(N)}(z_1, \dots, z_N | (n)) = Prob\{Z_i \leq z_i, i = 1, \dots, N | (n)\} \tag{1}$$

The simulation is sequentially done by drawing the values from the normal CCDF of each variable using the Monte Carlo simulation procedure. The conditioning data include all the original data (hard data), and all the previously simulated points are found within a neighborhood [18-20].

The conditioning simulation of a continuous variable  $z(u)$  proceeds as follows [18-20]:

1. Generate the univariate CCDF,  $F_z(z)$  for the entire grid.
2. Perform the normal transform of the  $z$  data into the  $y$  data with a standard normal CDF.
3. Define a random path that defines which node is going to be simulated first. At each node  $u$ , specify the number of surrounding conditioning data (both the hard data and previously simulated nodes).
4. Use simple kriging (SK) to determine the mean and variance of CCDF.
5. Draw a simulated value  $y^{(1)}(u)$  from CCDF using the Monte Carlo simulation.
6. Add the simulated value to the dataset.
7. Proceed to the next node based on the pre-defined random path, and repeat the steps until all node are simulated.
8. Transform the simulated values back into the original variable (reverse normal transform).

**3. Direct Sequential Simulation**

Soares (2001) has introduced a direct sequential simulation algorithm without any transformation of the original variable. The algorithm procedure is described in the following steps [11]:

1. Calculate the global CFD,  $F_z(z)$  of the continuous variable  $Z(x)$ .
2. Define a random path.
3. Estimate the local mean and variance of the visited node using simple kriging, presented as  $z(x_u)^*$  and  $\sigma_{sk}^2(x_u)$ , respectively.
4. Define the interval of  $F_z(z)$  to be sampled using Gaussian CDF.
5. Generate the value  $p$  from a uniform distribution  $U(0,1)$ .
6. Generate the value  $y^s$  from  $G(y(x_u)^*, \sigma_{sk}^2(x_u))$  using the following equation:

$$y = G^{-1}(y(x_u)^*, \sigma_{sk}^2(x_u), p) \tag{2}$$

where  $y(x_u)^*$  corresponds to the local estimate  $z(x_u)^*$  and  $G^{-1}$  is the inverse Gaussian CDF.

7. Return the simulated value using the following equation:

$$z^s(x_u) = \varphi^{-1}(y^s) \tag{3}$$

where  $\varphi^{-1}$  in the inverse normal score transform.

8. Repeat until all nodes have been visited.

Figure 1 shows the sampling of global distribution  $F_z(z)$  using the intervals that came from the local estimate [11].

Note that a Gaussian transformation is used for the sample interval preparation purpose only, and not for generating the local distribution as in the SGS algorithm.

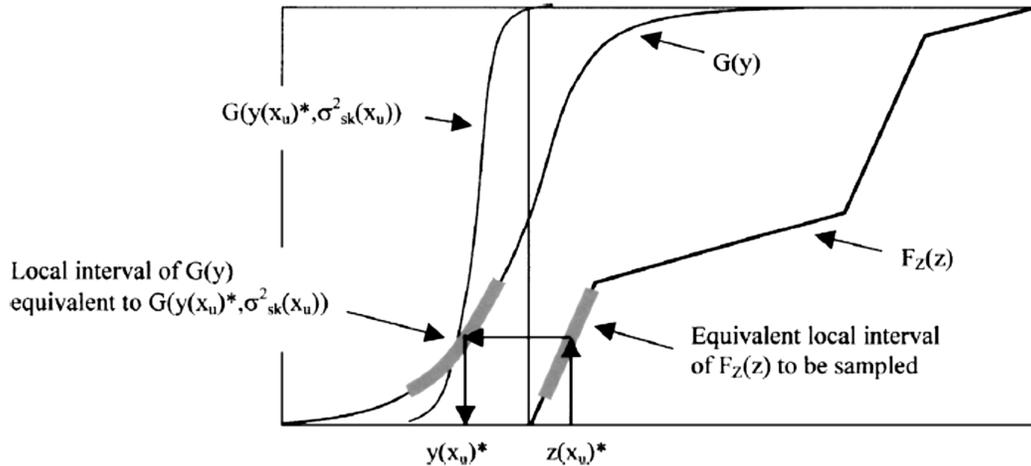


Figure 1. Sampling of global distribution  $F_z(z)$  [11].

#### 4. Application of algorithms to a 3D dataset

In order to compare the two simulation algorithms, a pseudo-real dataset from a deepwater turbidite reservoir was selected. The acoustic impedance values were available from a seismic inversion process. These values were later converted to the porosity values. This means that the actual porosity values in all areas are available. The grid has the dimension of  $100 \times 100 \times 70$  cells as a result of the seismic inversion procedure. Each

cell is 25 m long in each direction. These units came from the 3D seismic data acquisition. In order to build a dataset to be used for simulations, a static model was built by extracting 16 vertical wells from the 3D data. All wells were randomly chosen. One well remained unused for blind well purposes. Figure 2a shows a perspective view of wells in the grid. A 2D horizontal slice is illustrated in Figure 2b, showing the location of all wells including the blind well.

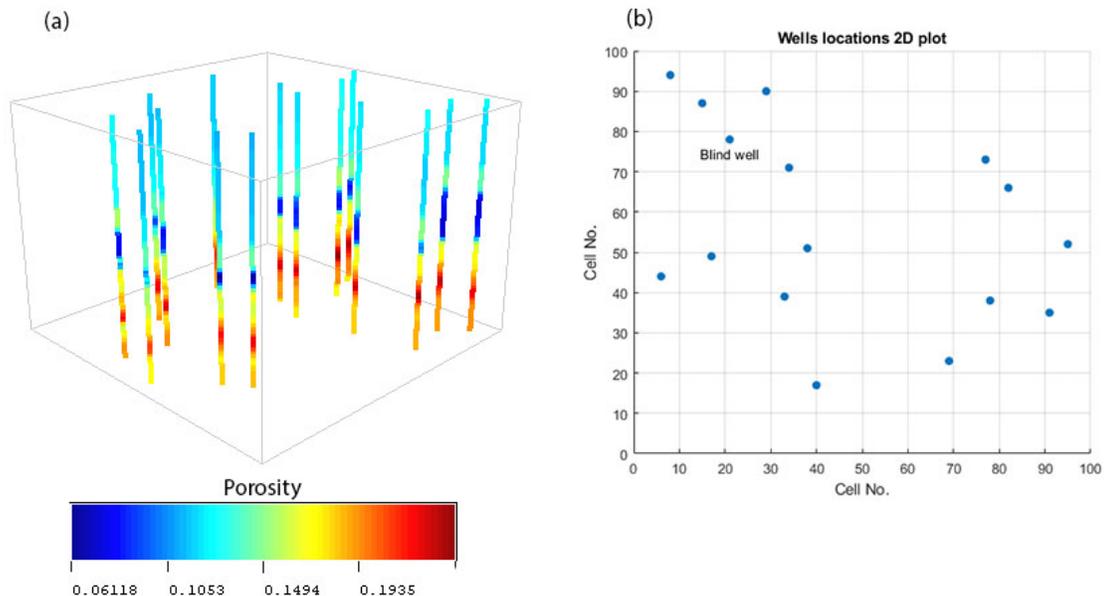


Figure 2. a) 3D view of porosity data wells, b) 2D horizontal slice of wells.

In order to do any geo-statistical simulation, first, we need to do the semivariogram modeling. In this research work, two directions were selected to calculate the variograms, horizontal and vertical. Figure 3 presents the experimental variograms and fitted spherical models as well. The horizontal

direction was considered omnidirectional since there were no considerable changes in different horizontal directions. The ranges for the horizontal and vertical directions were calculated 36 and 18 cells, respectively.

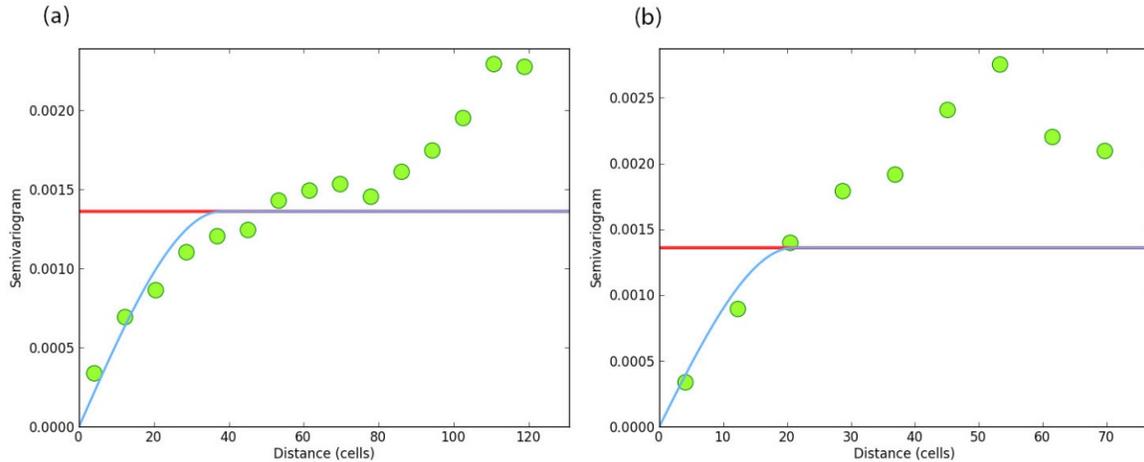


Figure 3. a) Horizontal variogram and b) vertical variogram modeling.

Using the variogram parameters and 15 well logs, the sequential Gaussian simulation algorithm was applied to the grid. Two vertical sections in both x and y directions are presented in Figure 4. The same parameters were used to run the direct sequential simulation algorithm. Figure 5 shows the same sections extracted from the DSS results. These figures are only for visualization purposes.

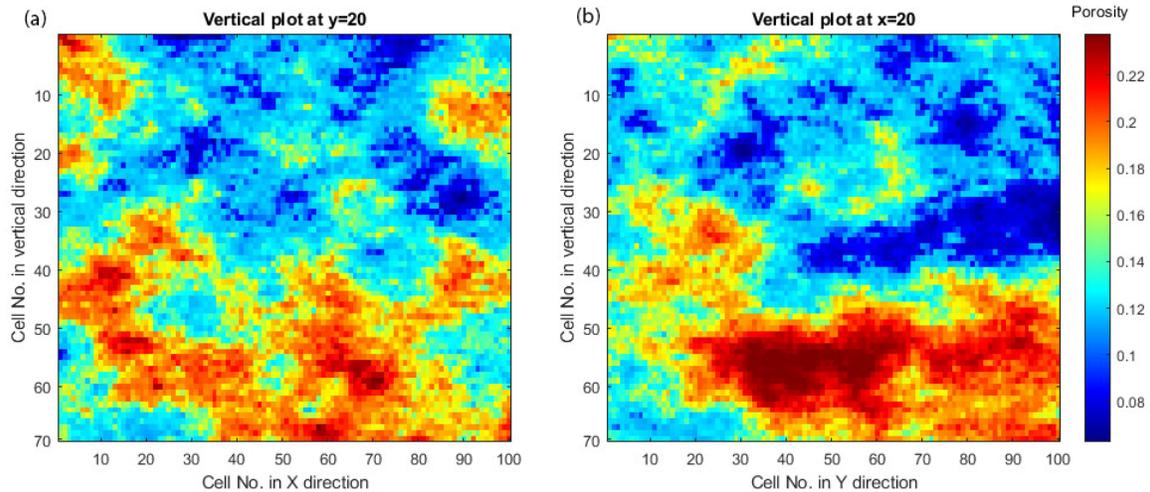


Figure 4. SGS results. (a) a vertical section in x direction and (b) a vertical section in y direction.

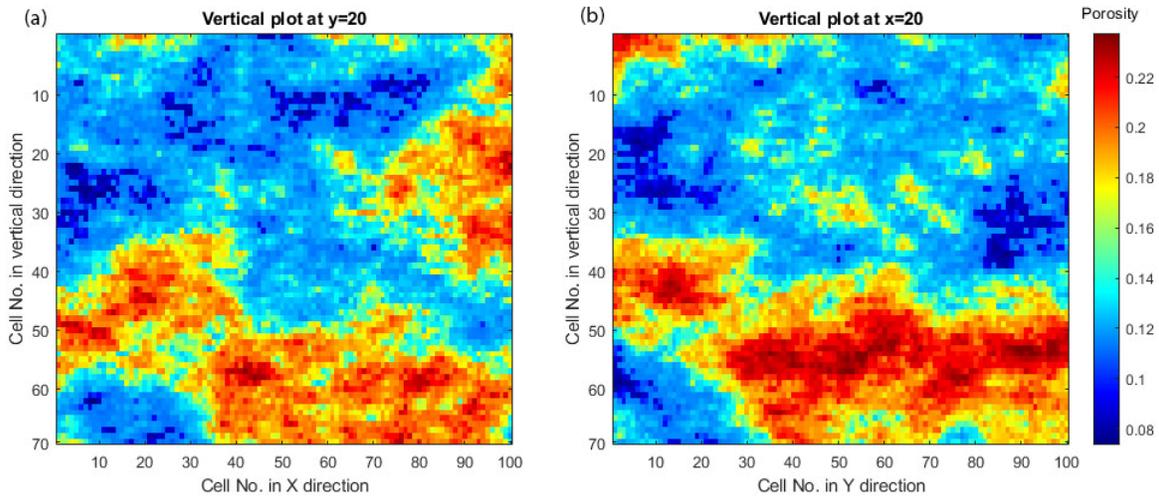


Figure 5. DSS results. (a) a vertical section in x direction and (b) a vertical section in y direction.

In order to see if the histogram is reproduced in the simulation results, the histograms are plotted in Figures 6 and 7 for the SGS and DSS algorithms, respectively. As seen, both algorithms are able to reproduce the histogram of the well data in the simulated values. The SGS algorithm seems to be slightly more successful in the reproduction of the histogram. Other important parameters that should be reproduced in any geo-statistical simulation algorithms are the mean and variance of the input data. Table 1 is created for the comparisons. As shown in this table, both algorithms reproduce the mean and variance of the well data in their outputs. Although the differences are negligible, the DSS algorithm is able to reproduce the mean more efficiently.

For the purpose of variogram reproductions, Figures 8 and 9 are presented. As shown in these figures, the variograms are reproduced in the simulation results for both the SGS and DSS simulation algorithms.

Since the porosity cube of the same grid dimension is available in the current dataset, it is possible to compare it to the simulation results. One realization from each simulation algorithm is randomly selected. In addition, the mean cube through 100 realizations for each simulation algorithm is included in this comparison. Root mean square error (RMSE) is calculated, and the results are shown in Table 2. This table confirms that both algorithms are approximately equal in the simulation of the porosity of all cells.

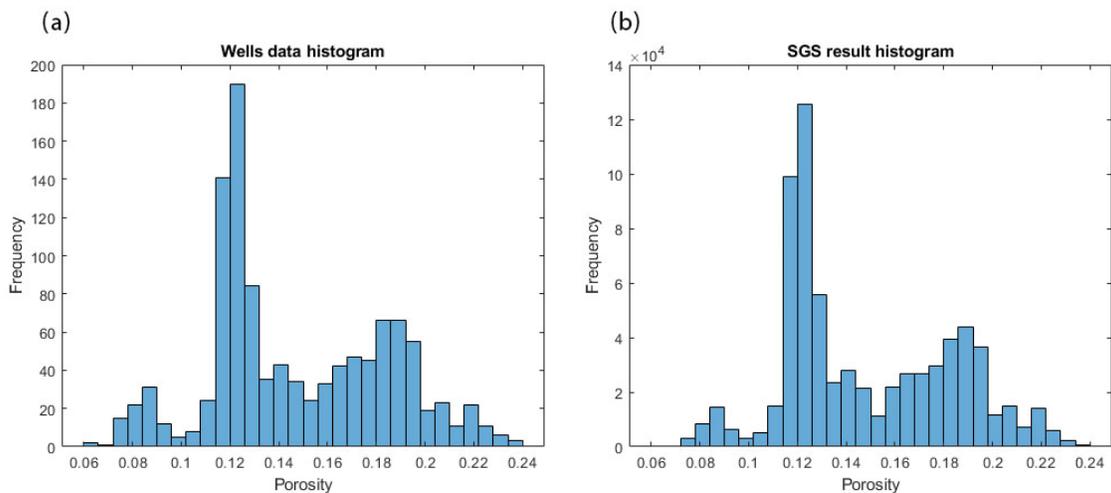


Figure 6. (a) Histogram of porosity from well data and (b) histogram of simulation result using SGS algorithm.

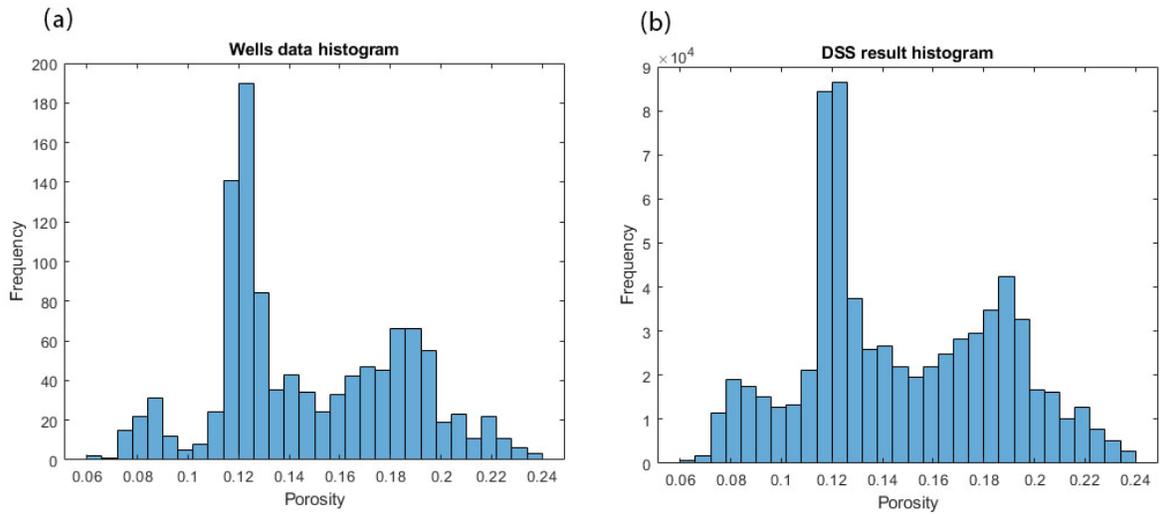


Figure 7. (a) Histogram of well porosity data and (b) histogram of simulation result using DSS algorithm.

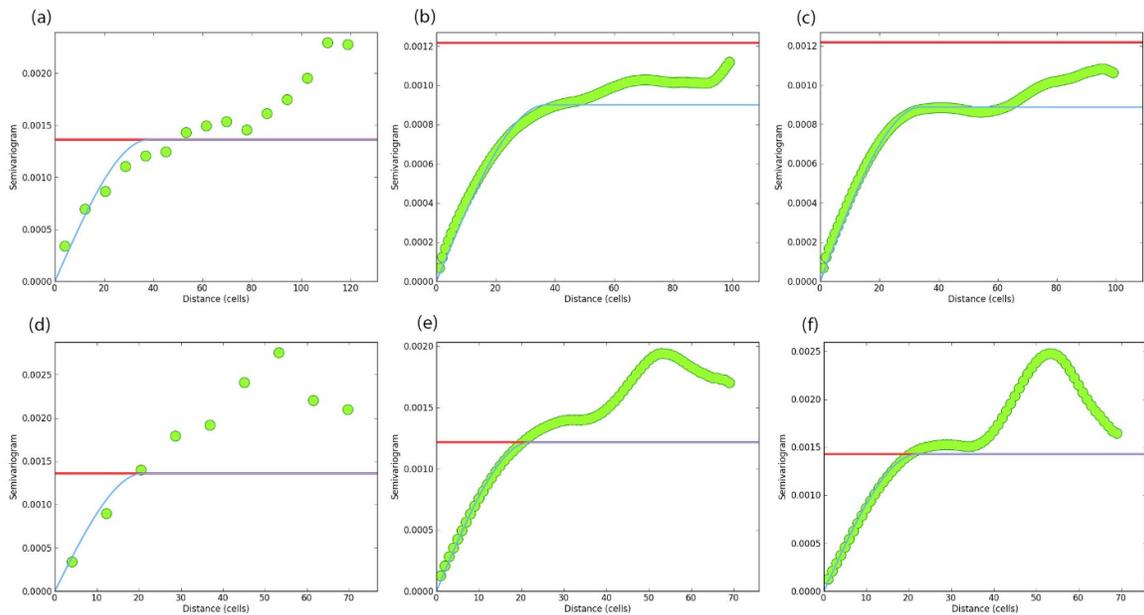


Figure 8. Variogram reproduction in SGS algorithm. (a) Horizontal variogram of well data. (b) and (c) Horizontal variograms of two randomly selected simulation results from SGS algorithm. (d) Vertical variogram of well data. (e) and (f) Vertical variograms of simulation results.

Table 1. Mean and variance comparisons for SGS and DSS results.

	Mean	Variance	Mean difference from well data	Variance difference from well data
Well data	0.1467	0.00136	-	-
SGS results	0.1479	0.00122	7.6 %	-10.7 %
DSS results	0.1468	0.00152	0.047 %	11.2 %

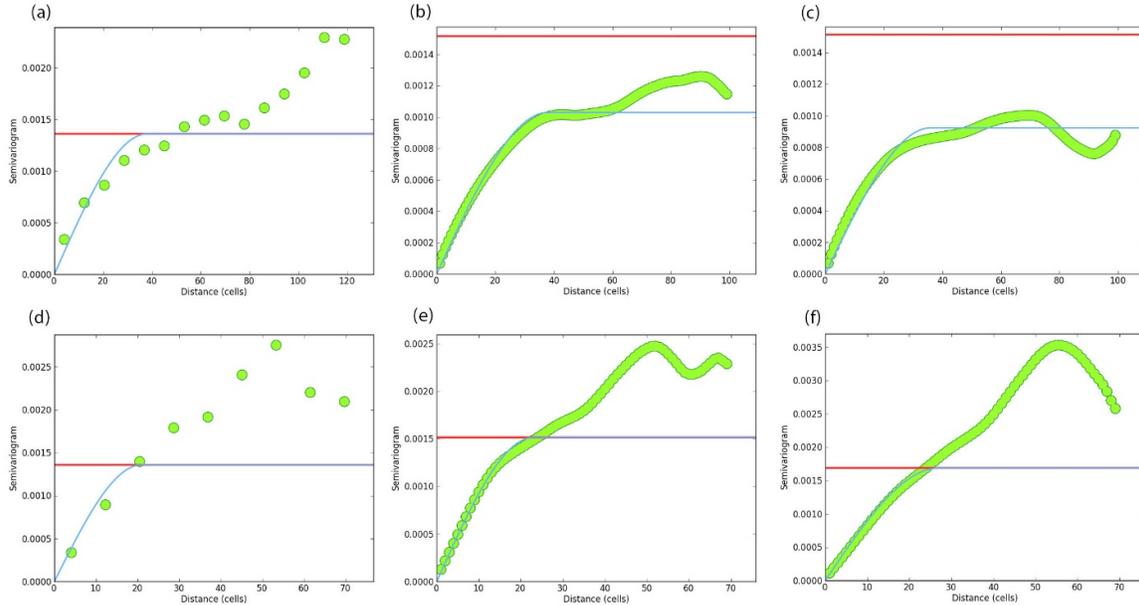
Table 2. Root mean square error (RMS error) between actual porosity values and simulated ones.

	One realization	Mean of 100 realizations
SGS algorithm	3.43 %	1.98%
DSS algorithm	3.50 %	1.68%

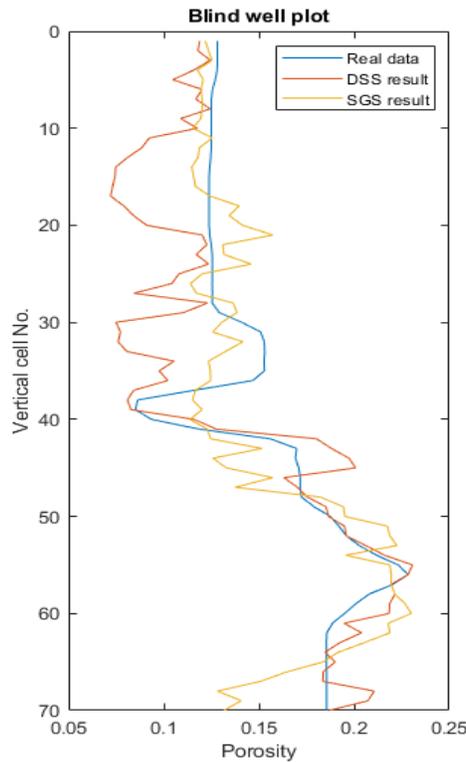
**5. Blind well analysis**

For vertical changes analysis, a blind well test was performed. Figure 10 presents the actual porosity data plot versus the SGS and DSS results.

Table 3 shows the correlation coefficient and RMS error between the actual data and the simulation results. Although the correlation coefficient for the DSS algorithm is higher than the SGS algorithm, the RMS errors show a reverse outcome.



**Figure 9. Variogram reproduction in DSS algorithm. (a) Horizontal variogram of well data. (b) and (c) Horizontal variograms of two randomly selected simulation results from DSS algorithm. (d) Vertical variogram of well data. (e) and (f) Vertical variograms of simulation results.**



**Figure 10. Blind well test. Comparing porosity values from actual data, SGS and DSS results along the blind well as presented in Figure 3.**

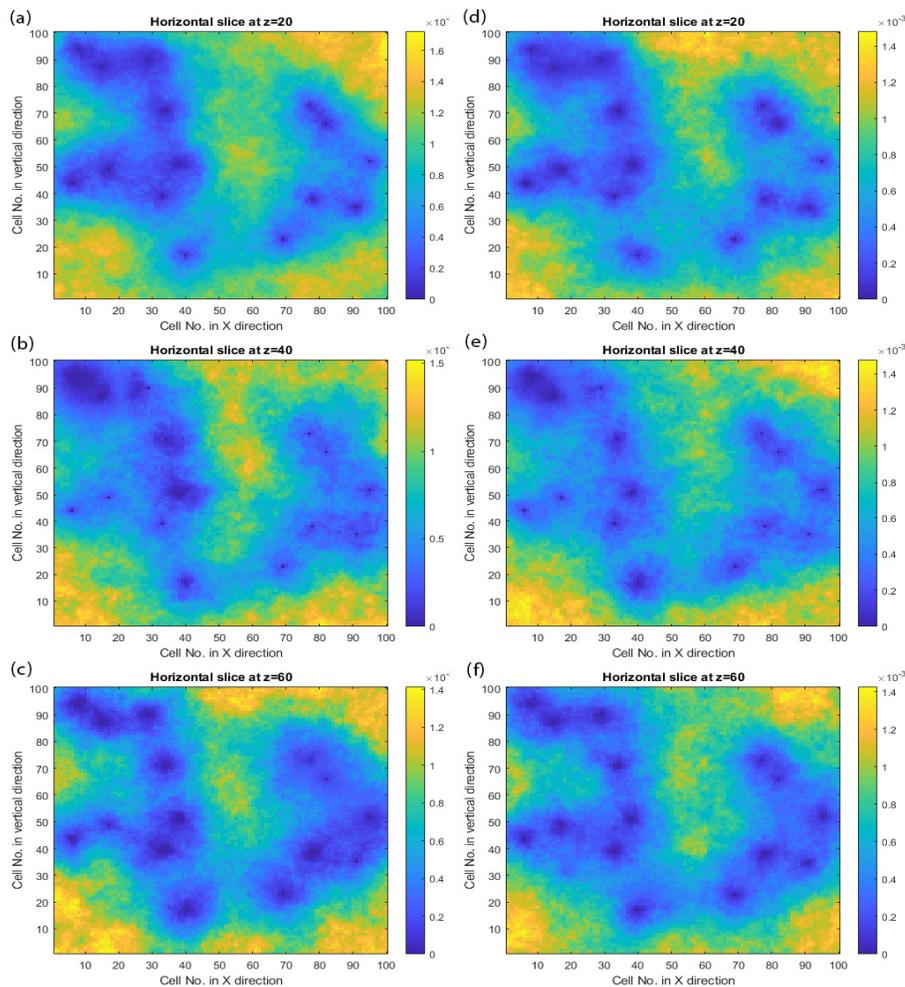
**Table 3. Blind well analysis. Correlation coefficients and RMS error between actual data and SGS and DSS results.**

	Correlation coefficient	RMS error
SGS algorithm	83 %	21.4 %
DSS algorithm	88 %	28.9 %

**6. Uncertainty analysis**

One of main advantages of the simulation algorithms compared to other deterministic estimation algorithms is the ability to do the uncertainty assessment. This assessment can be done by calculating the variance for each cells using various simulated results. In this research work, 100 realizations were used for generating a variance cube for each simulation algorithm. Higher variance values show that the uncertainty of simulated values to be more close to the actual ones is high, and vice versa. As the geo-statistical simulation algorithms are conditioned to the well

data, it is expected to have less uncertainty close to the wells locations. For the visualization purpose, three horizontal slices from different areas were extracted from variance cube and illustrated in Figure 11. As shown in this figure, the more distance from wells, the higher uncertainty in the simulation results. The averages of standard deviation values among all simulated cells from each simulation results have been calculated just to have an idea about the uncertainty. These averages were 0.0264 and 0.0252 for SGS and DSS algorithms, respectively, showing that the uncertainty corresponding to both simulation algorithms is approximately equal.



**Figure 11. Uncertainty analysis. Horizontal slices extracted from different areas of variance cube from SGS (a, b and c) and DSS algorithm results (d, e and f).**

## 7. Conclusion

In this work, two common geo-statistical simulation algorithms, SGS and DSS, were applied to a porosity dataset to compare the results. The sequential Gaussian simulation algorithm requires transformation of data into Gaussian distribution. Unlike the SGS algorithm, the direct sequential simulation uses the original data during the simulation procedure. Applying these two algorithms to the porosity data of 15 wells shows that both algorithms are able to generate simulated cubes without showing any errors simulating 700000 cells. The same random path and variogram models were used. The histogram reproduction analysis shows that the SGS algorithm is able to slightly better reproduce the histogram of well data. Note that the populations are perfectly reproduced for the algorithms in the same way. On the other hand, the DSS algorithm reproduce the mean value of well data with much less RMS error compared to the SGS algorithm. The reproduction of variance value was approximately equal. Considering one realization from each simulation algorithm, the RMS errors corresponding to all simulated cells were almost the same but the mean of 100 realizations showed a slightly less RMS error in the DSS algorithm. Regarding a blind well test correlation coefficient between the actual values and simulated ones was higher when using the DSS algorithm. However, the RMS error shows a less value when using the SGS algorithm.

## Acknowledgments

The authors would like to thank three anonymous reviewers for their useful comments. The authors are grateful to Professor Amilcar Soares for his valuable remarks.

## References

- [1]. Deutsch, CV and Journel, A.G. (1998). GSLIB Geostatistical Software Library and User's Guide. New York: Oxford University Press.
- [2]. Bai, T. and Tahmasebi, P. (2022). Sequential Gaussian simulation for geosystems modeling: A machine learning approach. *Geoscience Frontiers*. 13 (1): 101258.
- [3]. Shen, W., Hu, Y., Zhang, J., Zhao, F., Bian, P. and Liu, Y. (2021). Spatial distribution and human health risk assessment of soil heavy metals based on sequential Gaussian simulation and positive matrix factorization model: A case study in irrigation area of the Yellow River. *Ecotoxicology and Environmental Safety*. 225, 112752.
- [4]. Madenoglu, S., Atalay, F. and Erpul G. (2020). Uncertainty assessment of soil erodibility by direct sequential Gaussian simulation (DSIM) in semiarid land uses. *Soil and Tillage Research*. 204, 104731.
- [5]. Sotoudeh, F., Ataei, M., Kakaie, R. and Pourrahimian, Y. (2020). Application of Sequential Gaussian Conditional Simulation to Underground Mine Design under Grade Uncertainty. *Journal of Mining and Environment*. 11 (3): 695-709.
- [6]. Talesh Hosseini S., Asghari O, and Ghavami Riabi SR. (2018). Spatial modelling of zonality elements based on compositional nature of geochemical data using geostatistical approach: a case study of Baghqlloom area, Iran. *Journal of Mining and Environment*. 9 (1): 153-67.
- [7]. Sakizadeh M., Sattari M.T. and Ghorbani H. (2017). A new method to consider spatial risk assessment of cross-correlated heavy metals using geo-statistical simulation. *Journal of Mining and Environment*. 8 (3): 373-91.
- [8]. Metahni, S., Coudert, L., Gloaguen, E., Guemiza, K., Mercier, G. and Blais J-F. (2019). Comparison of different interpolation methods and sequential Gaussian simulation to estimate volumes of soil contaminated by As, Cr, Cu, PCP and dioxins/furans. *Environmental Pollution*. 252, 409-19.
- [9]. Wang, J. and Zuo, R. (2018). Identification of geochemical anomalies through combined sequential Gaussian simulation and grid-based local singularity analysis. *Computers and Geosciences*. 118, 52-64.
- [10]. Gholampour, O., Hezarkhani, A., Maghsoudi, A. and Mousavi, M. (2019). Application of sequential Gaussian simulation and concentration-volume fractal model to delineate alterations in hypogene zone of miduk porphyry copper deposit, SE Iran. *Journal of African Earth Sciences*. 150, 389-400.
- [11]. Soares A. (2001). Direct Sequential Simulation and Cosimulation. *Mathematical Geology*. 33(8), 911-26.
- [12]. Ribeiro, S., Caineta, J., Costa, A.C., Henriques, R., and Soares, A. (2016). Detection of inhomogeneities in precipitation time series in Portugal using direct sequential simulation. *Atmospheric Research*. 171, 147-58.
- [13]. Sabeti, H., Moradzadeh, A., Ardejani, F.D., Azevedo, L., Soares, A., Pereira, P. and Nunes, R. (2017). Geostatistical seismic inversion for non-stationary patterns using direct sequential simulation and co-simulation. *Geophysical Prospecting*. 65, 25-48.
- [14]. Pereira, P., Azevedo, L., Nunes, R. and Soares, A. (2019). The impact of a priori elastic models into iterative geostatistical seismic inversion. *Journal of Applied Geophysics*. 170, 103850.
- [15]. Almeida, F., Davolio, A. and Schiozer, D.J. (2020). Reducing uncertainties of reservoir properties in

an automatized process coupled with geological modeling considering scalar and spatial uncertain attributes. *Journal of Petroleum Science and Engineering*. 189, 106993.

[16]. Horta, A., Azevedo, L., Neves, J., Soares, A. and Pozza, L. (2021). Integrating portable X-ray fluorescence (pXRF) measurement uncertainty for accurate soil contamination mapping. *Geoderma*. 382, 114712.

[17]. Otzen, M., Finlay, C.C. and Hansen, T.M. (2022). Direct Sequential Simulation for spherical linear inverse problems. *Computers and Geosciences*. 160, 105026.

[18]. Hajsadeghi, S., Asghari, O., Mirmohammadi, M., Afzal, P. and Meshkani, S.A. (2020). Uncertainty-volume fractal model for delineating copper mineralization controllers using geostatistical

simulation in Nohkouhi volcanogenic massive sulfide deposit, Central Iran. *Bulletin of the Mineral Research and Exploration*. 161 (161): 1-11.

[19]. Soltani, F., P. Afzal, and Asghari, O. (2014). Delineation of alteration zones based on Sequential Gaussian Simulation and concentration-volume fractal modeling in the hypogene zone of Sungun copper deposit, NW Iran. *Journal of Geochemical Exploration*, 140, 64-76.

[20]. Soltani, F., Moarefvand, P., Alinia, F. and Afzal, P. (2019). Characterization of rare earth elements by coupling multivariate analysis, factor analysis, and geostatistical simulation; case-study of Gazestan deposit, central Iran. *Journal of Mining and Environment*. 10 (4): 929-945.

## مقایسه عملی الگوریتم‌های شبیه‌سازی متوالی گوسی و مستقیم با استفاده از یک داده سه‌بعدی تخلخل

حمید ثابتی<sup>۲،۱\*</sup> و فرزاد مرادپوری<sup>۳</sup>

۱- گروه مهندسی معدن، دانشگاه صنعتی بیرجند، بیرجند، ایران

۲- مرکز تحقیقات ذخائر طبیعی و محیط زیست، انستیتو فنی عالی، دانشگاه لیسبون، لیسبون، پرتغال

۳- گروه مهندسی معدن، دانشکده فنی و مهندسی، دانشگاه لرستان، خرم‌آباد، ایران

ارسال ۲۰۲۲/۰۵/۰۱، پذیرش ۲۰۲۲/۰۶/۰۱

\* نویسنده مسئول مکاتبات: sabeti@birjandut.ac.ir

## چکیده:

روش‌های شبیه‌سازی زمین‌آماري برای تولید مدل‌های مقید به خواص آماری از متغیرهای فضایی پیوسته، به طور گسترده مورد استفاده قرار گرفته‌اند. دو روش عمده برای این انجام شبیه‌سازی متغیرهای پیوسته وجود دارد، شبیه‌سازی متوالی گوسی و شبیه‌سازی متوالی مستقیم. برتری اصلی روش مستقیم در مقابل روش گوسی این است که تبدیل گوسی داده‌های ورودی در روش مستقیم صورت نمی‌پذیرد. در این پژوهش، این دو روش شبیه‌سازی شرح داده شده و کاربرد آنها بر روی یک داده شبه واقعی تخلخل مورد بررسی دقیق قرار گرفته است. این داده شامل مقادیر تخلخل ۱۶ چاه قائم است که از مکعب داده‌های سه‌بعدی خروجی فرایند وارون‌سازی لرزه‌ای استخراج شده است. داده‌های یک چاه برای ارزیابی نتایج کنار گذاشته شده و دو الگوریتم شبیه‌سازی متوالی گوسی و مستقیم بر روی این داده اعمال شده است. مقایسه هیستوگرام‌ها نشان می‌دهد که بازتولید هیستوگرام در الگوریتم شبیه‌سازی متوالی گوسی اندکی بهتر است اگرچه بازتولید جوامع آماری در هیستوگرام نتایج هر دو الگوریتم به طور یکسان قابل مشاهده است. نتایج الگوریتم شبیه‌سازی متوالی مستقیم دارای میانگین نزدیکتری به میانگین داده‌های ورودی هستند. با در نظر گرفتن یکی از تحقق‌ها از هر کدام از دو روش شبیه‌سازی، خطای جذر میانگین مربعات بین مقدار شبیه‌سازی شده و مقدار واقعی تقریباً یکسان است. از طرف دیگر، با در نظر گرفتن میانگین ۱۰۰ تحقق، این خطا در مورد الگوریتم شبیه‌سازی متوالی مستقیم کمتر از شبیه‌سازی متوالی گوسی است.

**کلمات کلیدی:** زمین‌آمار، شبیه‌سازی متوالی گوسی، شبیه‌سازی متوالی مستقیم، واریوگرام، تخلخل.