Geostatistics for Risk Level Mapping: Synthetic Data Example

Waskito Pranowo[1], Adhitya Ryan Ramadhani[2,]

{waskito.pranowo@universitaspertamina.ac.id[1], adhitya.rr@ universitaspertamina.ac.id [2], arramadhani@mun.ca[2]}

Geophysical Engineering, Universitas Pertamina[1], Mechanical Engineering, Universitas Pertamina [2], Centre for Risk, Integrity and Safety Engineering, Memorial University of Newfoundland[2]

Abstract. Risk level mapping is important for mitigation plans and any disaster-related decision. However, the data in certain areas can be sparse for some reasons. The risk levels are not known or analyzed at some positions. In statistics, these positions are called unsampled locations. Geostatistics can play a role in estimating the risk at unsampled locations. The most common geostatistics method that can be used is kriging. Moreover, kriging can calculate the uncertainty of its estimation. This paper aims to investigate the benefit of kriging in risk level mapping. The synthetic data experiment is conducted to explain how kriging works in risk level mapping. Kriging method is able to estimate the risk level at unsampled locations and take into account uncertainties.

Keywords: Risk level mapping, geostatistics, estimation, kriging, uncertainty.

1 Introduction

In the risk analysis in a certain area, sometimes only several locations are analyzed. In addition to this, many geoscience and engineering data are imprecise due to various limitations and uncertainties. Therefore, the risk level of the area is difficult to be assessed and may be sparse. While identifying risk levels in this specific area is very important. In risk assessment, the ability to predict the occurrence of any disruptive events is crucial and valuable. Thus, a method that can assess risk in unsampled locations is required.

The nearest neighbour [1] is the simplest method to estimate the risk level at unanalyzed (in the next, we called unsampled) locations. However, this method cannot calculate the uncertainty of the estimation. At the same time, uncertainty is a frequently arising issue that needs to be considered. Geostatistics is a family method that can estimate data and calculate its uncertainty. Geostatistical methods are found to be effective in dealing with problems related to the estimation of spatial variables [2].

Geostatistics is a branch of statistics that focuses on spatial estimation of spatially correlated variables for earth science applications and their uncertainty [3], [4]. Some of the geostatistics methods are linear kriging, nonlinear kriging, co-kriging, simulation [3], and multi-point geostatistics [5]. Kriging is a common and widely used method to create a risk map. Although

ICONIC-RS 2022, March 31-April 01, DKI Jakarta, Indonesia

DOI 10.4108/eai.31-3-2022.2320971

its application is mainly in hydrocarbon reservoir characterization, it can be extended to other applications, such as groundwater exploration [6], [7], environmental cases [8], astronomy [9], and remote sensing [10]. In addition to these, Pokhrel et al. (2013) used kriging method to map liquefaction potential in an area. Chica-Olmo et al. (2014) used Indicator Kriging to assess the risk of groundwater nitrate pollution. The proposed method was robust and was beneficial in supporting health risk analysis studies. Deal & Sabatini (2020) utilized kriging to find suitable zones for manual drilling. The probability maps resulting from the proposed method were able to estimate the probability of success for future manual drilling activities. Risk maps are usually a useful tool resulting from the utilization of kriging method and are also found useful to identify high-risk level areas in some studies for various areas [13]–[18]. Thus, kriging method is used in this paper to create a risk map.

This paper aims to investigate the benefit of kriging methods used in risk level mapping. The application of kriging is presented using synthetic data. Risk level maps are produced to identify areas with significant probabilities of severe disaster.

The remainder of this paper is organized as follows. Section 2 presents the theory and methods. Data description used to investigate the application of kriging method is discussed in Section 3. Results and discussion from the estimation are presented in Section 4. The concluding remarks of this paper are presented in Section 5.

2 Theory and method

2.1 Kriging

Kriging is one of the spatial estimation methods, apart from triangular interpolation, nearest neighbour, and inverse-distance. Kriging estimates an unsampled location using the neighbour sampled data. The estimation is based on the distance and the configuration of the neighbour data. Therefore, kriging requires spatial relationship information in the process. The variogram is the most common representation of this spatial relationship, while the other is covariance.

The variogram is calculated by

$$\gamma(\vec{L}) = \frac{1}{2n(\vec{L})} \sum_{i=1}^{n(\vec{L})} \left[x(\vec{u}_i) - x(\vec{u}_i + \vec{L})\right]^2 \tag{1}$$

where $\vec{L}$ is lag, $x(\vec{u}_i)$ is data at $\vec{u}_i$ position, $x(\vec{u}_i + \vec{L})$ is data at $\vec{u}_i + \vec{L}$ position, $n(\vec{L})$ is number of data couple, and $\gamma(\vec{L})$ is variogram value.

The variogram calculates two sampled data (a data couple) relationships. After the calculated variogram is obtained, the next process is variogram modelling. Variogram modelling is a fitting of calculated variogram by a model from mathematical equations. The common models are Gaussian, Exponential, Spherical, and Nugget models [3] (Fig.1). The example of the calculated variogram fitting by a model is shown in Fig. 2.

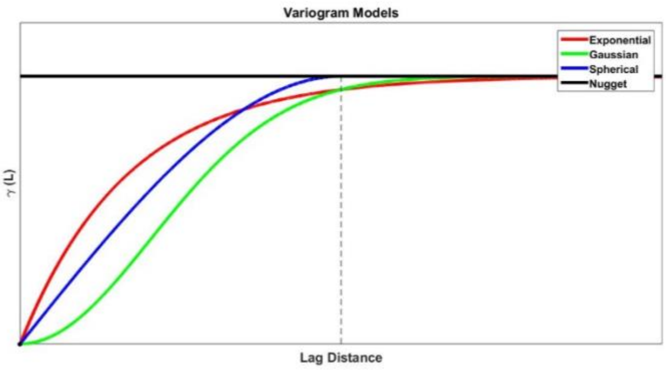

Figure 1. Variogram model: Exponential (red), Gaussian (green), Spherical (blue), and Nugget (black)

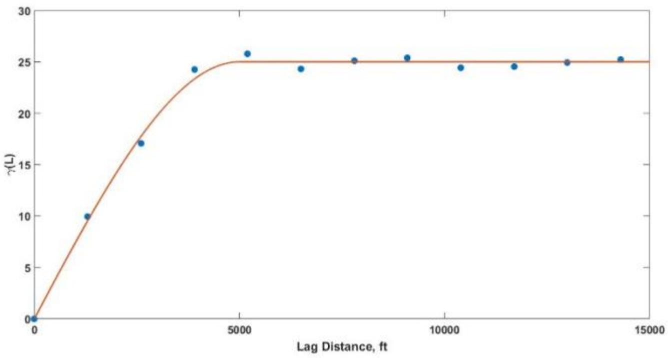

Figure 2. Fitting calculated variogram with a spherical model.

Next, the kriging estimation is calculated by

$$x_0 = \sum_{j=1}^{N} \lambda_j x_j \tag{2}$$

where x_j is the neighbour sampled data, λ_j is the weight of each neighbour, and x_0 the estimated value at unsampled data.

There are several techniques in kriging estimation, such as simple, ordinary, and so on. Eq. 2 is the equation for ordinary kriging. Each method has its assumptions. The weight of each neighbour requires the variogram model.

Based on Eq.2 and 3, the kriging is applied for only quantitative data. Qualitative data, such as risk level, requires transformation to be quantitative data. However, at the end of the kriging process, it requires inverse transformation to take back the qualitative data.

2.2 Probability map

The output of kriging is not only an estimated data map but also an error variance map. The error variance describes the uncertainty of an estimated value due to the configuration of neighbors' positions. The probability of a range at each point estimation is calculated using its variance,

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]} dx \quad (3)$$

where σ is error variance and μ is estimation mean at each point.

The steps of risk mapping using kriging follows:

a. Get the sampled risk level at several points in a region

b. Transform the risk level to be quantitative data

c. Calculate a variogram of (b) (Eq. 1)

d. Fit the calculated variogram (c) with a most appropriate model (Fig. 1)

e. Estimate the risk at unsampled locations using (Eq. 2)

f. Calculate the probability (Eq. 3)

3 Data description

This paper uses synthetic map data of sampled risk levels (Fig.3) which is generated manually. The risk levels are divided into very-high, high, medium, low, and very low levels. The synthetic data has an increasing trend of level from east to west.

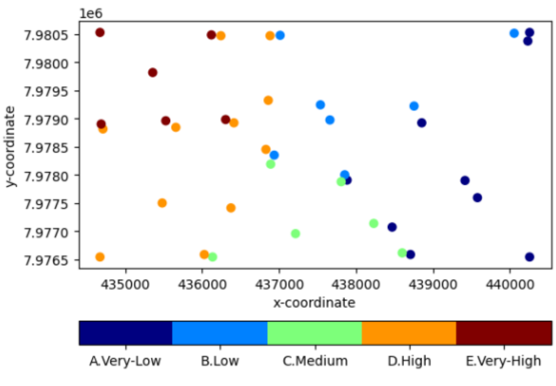

Figure 3. Synthetic data of analyzed risk level at several points (sampled locations).

4 Result and discussion

The full risk level map of risk level is generated using kriging (Fig.4 left). From the estimated map, the risk level increases from east to west. However, the estimation is not entirely true. Fig. 4 (right) shows the error variance of the estimation. The smaller values are given at the positions of sampled data. The further away from the sampled point, the variance value increases. For example, there are high variances in the northern and eastern regions caused by the distance of the position from the sampled points.

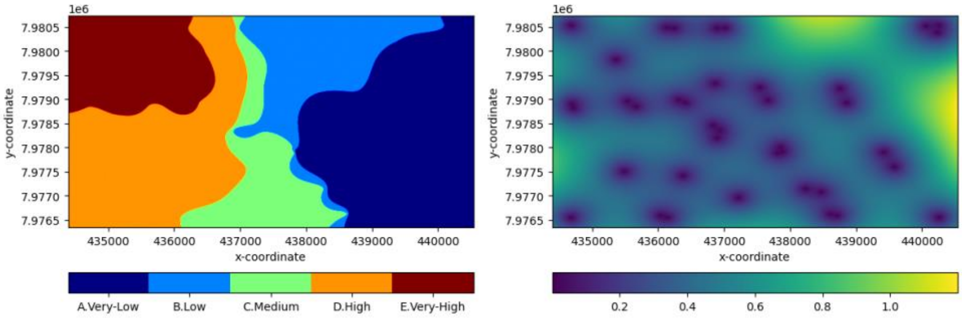

Figure 4. Estimated distribution of risk-level (left) and its error variance (right) map.

Because the estimation is 100% correct, it is better to analyze the probability of each risk level. Figure 5a shows the probability of low risk. The probability map shows that although the kriging shows there is a low-risk level at the middle of the map, the probability is low. It happens because two risk levels, which are very-low and medium levels, flank the low-risk region. On the other hand, a very-high level region has a high probability (Fig.5b). In contrast to the low-level region, the very high region is not surrounded by two different levels. The high probability dots are the positions of sampled data. These positions have 100% probability because they are sampled data, not estimated. Besides the probability of certain risk levels, we can calculate the probability of a range. Figure 6 shows the probability of high to very high risk.

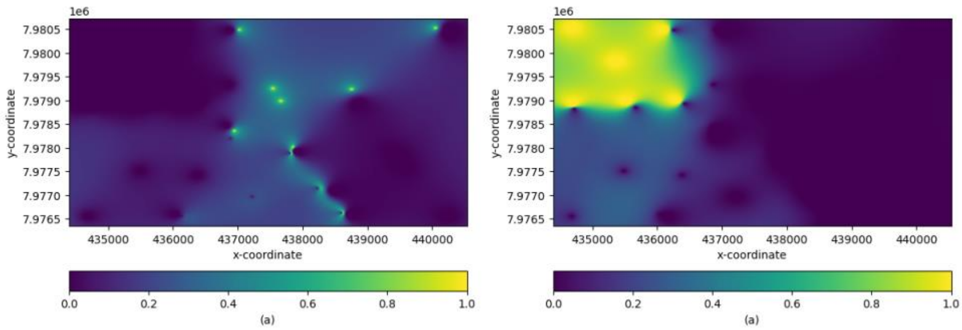

Figure 5. Map of the probability of distribution of low (a) and (b) very high risk-level.

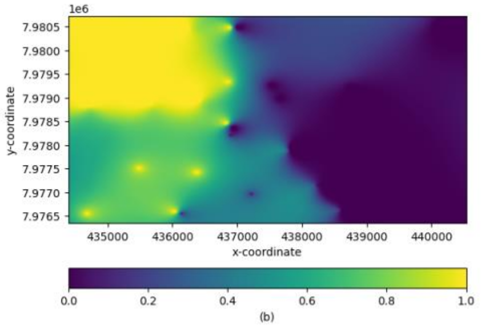

Figure 6. Map of the probability of distribution of high to very high risk.

Conclusion

The risk analysis data in a region may be sparse. It means only several locations are analyzed. However, the kriging method shows being able to estimate the risk level at unsampled locations. Apart from estimation, kriging could calculate the probability of the estimated data. Therefore, for the risk mapping, kriging could be an option. This method can be implmented to map risk levels of an area prone to natural distaster. This risk map is very beneficial in providing information on which specific area that will have the highest consequence if a natural disaster really occurs.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Reference

[1] Cover T, Hart P. Nearest neighbor pattern classification. IEEE Transactions on Information Theory. 1967;13(1):21-27. Doi: 10.1109/TIT.1967.1053964.

[2] Chica-Olmo M, Luque-Espinar JA, Rodriguez-Galiano V, Pardo-Igúzquiza E, Chica-Rivas L. Categorical Indicator Kriging for assessing the risk of groundwater nitrate pollution: The case of Vega de Granada aquifer (SE Spain). Science of The Total Environment. 2014;470–471(1):229-239. Doi: 10.1016/j.scitotenv.2013.09.077.

[3] Kelkar M, Perez G. Applied geostatistics for reservoir characterization. Society of Petroleum Engineers; 2002.

[4] Sheriff RE. 7. G. In: Sherrif RE. Encyclopedic dictionary of applied geophysics. Society of Exploration Geophysicists; 2002. p. 156-170. Doi: 10.1190/1.9781560802969.chg.

[5] Mariethoz G, Caers J. Multiple-point geostatistics. Chichester, UK: John Wiley & Sons, Ltd; 2014. Doi: 10.1002/9781118662953.

[6] Zimmerman DA, Marsily G, Gotway CA, Marietta MG, Axness CL, Beauheim RL, et al. comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. Water Resources Research. 1998;34(6):1373-1413. Doi: 10.1029/98WR00003.

[7] Tonkin MJ, Larson SP. Kriging water levels with a regional-linear and point-logarithmic drift. Ground Water. 2002;40(2):185-193. Doi: 10.1111/j.1745-6584.2002.tb02503.x..

[8] Bayraktar H, Turalioglu FS. A Kriging-based approach for locating a sampling site—in the assessment of air quality. Stochastic Environmental Research and Risk Assessment. 2005;19(4):301-305. Doi: 10.1007/s00477-005-0234-8.

[9] Pastorello N, Forbes DA, Foster C, Brodie JP, Usher CM, Romanowsky AJ, et al. The SLUGGS survey: exploring the metallicity gradients of nearby early-type galaxies to large radii. Monthly Notices of the Royal Astronomical Society. 2014;442(2):1003–1039. Doi: 10.1093/mnras/stu937.

[10] Papritz A, Stein A. Spatial prediction by linear kriging. In: Stein A, Van der Meer F, Gorte B, editors. Spatial statistics for remote sensing. remote sensing and digital image processing. Dordrecht: Springer; 1999. p. 83-113.

[11] Pokhrel RM, Kuwanto J, Tachibana S. A kriging method of interpolation used to map liquefaction potential over alluvial ground. Engineering Geology. 2013;152(1):26-37. Doi: 10.1016/j.enggeo.2012.10.003.

[12] Deal PT, Sabatini DA. Utilizing indicator kriging to identify suitable zones for manual drilling in weathered crystalline basement aquifers. Groundwater for Sustainable Development. 2020;11. Doi: 10.1016/j.gsd.2020.100402.

[13] Gerassis S, Boente C, Albuquerque MT, Ribeiro MM, Abad A, Taboada J. Mapping occupational health risk factors in the primary sector—A novel supervised machine learning and Area-to-Point Poisson kriging approach. Spatial Statistics. 2020;42(4). Doi: 10.1016/j.spasta.2020.100434.

[14] Maio FD, Belotti M, Volpe M, Selva J, Zio E. Parallel density scanned adaptive Kriging to improve local tsunami hazard assessment for coastal infrastructures. Reliability Engineering & System Safety. 2022;222. Doi: 10.1016/j.ress.2022.108441.

[15] Sun XL, Wu YJ, Zhang C, Wang HL. Performance of median kriging with robust estimators of the variogram in outlier identification and spatial prediction for soil pollution at a field scale. Science of The Total Environment. 2019;666:902–914. Doi: 10.1016/j.scitotenv.2019.02.231.

[16] Triantafilis J, Odeh IO, Warr B, Ahmed MF. Mapping of salinity risk in the lower Namoi valley using non-linear kriging methods. Agricultural Water Management.2004;69(3):203-231. Doi: 10.1016/j.agwat.2004.02.010.
[17] Giustini F, Ciotoli G, Rinaldini A, Ruggiero L, Voltaggio M. Mapping the geogenic radon potential and radon risk by using Empirical Bayesian Kriging regression: A case study from a volcanic area of central Italy. Science of The Total Environment. 2019;661:449-464. Doi: 10.1016/j.scitotenv.2019.01.146.
[18] Berens AS, Diem J, Stauber C, Dai D, Foster S, Rothenberg R. The use of gamma-survey measurements to better understand radon potential in urban areas. Science of The Total Environment. 2017;607–608:888–899 (2017). Doi: 10.1016/j.scitotenv.2017.07.022.