EARTH SCIENCES
RESEARCH JOURNAL
Earth Sci. Res. J. Vol. 27, No. 4 (December, 2023): 391 - 402

UNIVERSIDAD NACIONAL DE COLOMBIA

HYDROGEOLOGY

Geostatistical Estimation and Simulation in Dam Hydrogeological and Geotechnical Research: A Comprehensive Review

Shawgar Karami, Homayoon Katibeh*, Mohammadamin Karbala
Department of Mining Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran.
* Corresponding author: katibeh@aut.ac.ir

ABSTRACT

In dam engineering, the accurate assessment of hydrogeological and geotechnical parameters, including water pressure test (WPT), leakage, permeability, transmissibility, fractures' distribution, and rock quality designation (RQD) is fundamental for ensuring the safety, longevity, and performance of dam sites. Over the past few years, geostatistical approaches have emerged as valuable tools for estimating and simulating these significant features, offering the potential to reduce errors and minimize study costs. This research reviews the most significant, valid, and efficient research in this field and comprehensively presents the studies' results. An overview of the hydrogeological features of the dam sites will be presented. Then, the application of geostatistical approaches in each parameter is provided. Also, the strengths and weaknesses of these approaches are studied based on the prevailing conditions of the site. This research proves that geostatistics is an appropriate and efficient tool that can increase the accuracy of studies, reduce errors, and save time and money.

Keywords: Dam hydrogeological features, Geostatistical approaches, Geotechnical research, Hydrogeological conditions, Spatial correlation, Permeability distribution.

Estimación geoestadística y simulación en hidrogeología e investigación geotécnica de represas: una revisión

RESUMEN

En la ingeniería de represas una evaluación exacta de los parámetros hidrogeológicos y geotécnicos, como el análisis de la presión del agua, vertido, permeabilidad, transmisibilidad, distribución de fracturas, y la designación de calidad de roca, es fundamental para garantizar la seguridad, longevidad y desempeño de las áreas de las represas. En los últimos años, los enfoques geoestadísticos se han posicionado como herramientas útiles para la estimación y simulación de estas características determinantes y ofrecen la posibilidad de reducir los errores y minimizar los costos de estudio. En este trabajo se revisan los estudios más significativos, válidos y eficientes en este campo y se presentan los resultados de los estudios. Se presenta además una revisión de las características hidrogeológicas de las áreas de las represas. Luego se analiza la aplicación de los enfoques geoestadísticos de cada parámetro. También se estudian las fortalezas y debilidades de estos enfoques con base en las condiciones prevalentes del sitio. Este trabajo prueba que la geoestadística es una herramienta eficiente que puede incrementar la exactitud de los estudios, reducir los errores y ahorrar tiempo y dinero.

Keywords: Características hidrogeológicas de las represas; enfoques geoestadísticos; investigación geotécnica; condiciones hidrogeológicas; correlación espacial; distribución de permeabilidad.

Record
Manuscript received: 13/08/2022
Accepted for publication: 28/01/2024

How to cite this item:
Karami, S., Katibeh, H., & Karbala, M. (2023). Geostatistical Estimation and Simulation in Dam Hydrogeological and Geotechnical Research: A Comprehensive *Review. Earth Sciences Research Journal, 27(*4), 391 - 402. https://doi.org/10.15446/esrj.v27n4.104250

1. Introduction

Dams are unique structures complicated in their load response and interactive relationship with the hydrogeological and geotechnical characteristics. Understanding the spatial variability of these characteristics is significant in their evaluation, and obtaining valid data for the assessments is essential. However, a detailed evaluation of a dam site's geological, geotechnical, and hydrogeological features requires drilling boreholes in a regular high-density sampling network. It may not be economically justifiable, and, in many cases, the area has inappropriate geological and topographic conditions. Furthermore, in dam projects, the volume of samples to determine rock mass conditions can cover only a minor part of the area. Therefore, it is necessary to search for appropriate interpolators and estimators to improve the design of maps and determine the high uncertainties and intrinsic variability of soil and rock properties with high accuracy.

Earth science phenomena are spatially dependent on changes occurring at different geological scales. All soils and geological formations show random changes in their spatial characteristics. Evaluating these spatially correlated geological and hydrogeological variations at the dam site can be complicated. Undoubtedly, the environment of a dam site is affected by spatial variations due to the nature of different units and formations and the variety of other geological features, including faults and folds, permeability and porosity of various units and formations, and hydrogeological features.

Geostatistics is one of the advanced techniques for studying, interpolating, and evaluating the spatial data distribution in geoscience phenomena. The chief purposes of geostatistics are to describe and interpret the behavior of existing sample data and to use it for predicting potential values at unsampled locations. It is used where the variables are distributed in space, and spatial correlation can be found between them (Journel and Huijbregts, 1976; Journel, 1989). Geostatistics uses measured points' mathematical and statistical properties, including autocorrelation, saving time and costs (Jalali et al., 2016). It can provide a continuous surface when sample points are available in different locations. They use mathematical functions for interpolation that are directly based on measured neighborhood values (Matheron, 1971).

This research reviews the application of different geostatistical methods in hydrogeological and geotechnical studies of dam sites. The most significant reasons to conduct the research and analyze the hydrogeological and geotechnical conditions can be as follows: evaluating and identifying the possible water escape routes and estimating the leakage of foundations and abutments, studying the permeability of various units and formations of a dam site, assessing the site hydrogeology and features affecting it, creating sealing structures in the foundation, body, and abutments to prevent water waste in the dam reservoir, investigating the engineering geological and geotechnical characteristics, and problems related to hydro-stratigraphy of the layers.

First, the estimation and simulation approaches in dam sites' research are described in section 2. Then, in section 3, an overview of the hydrogeological characteristics of the dam sites, including WPT, permeability, transmissibility, hydraulic gradient, RQD, leakage, and fracture distribution, will be presented. Finally, section 4 provides the application of geostatistical approaches in each hydrogeological parameter of the dam site. This research proves that geostatistics is an appropriate and efficient tool that can increase the accuracy of studies, reduce errors, and save time and money.

2. Geostatistical Estimation and Simulation Approaches

Geostatistical concepts begin with the study of statistical processes, followed by spatial regression, kriging interpolation, and finally, quantifying uncertainties and estimation errors (Matheron, 1971). Geostatistics's most significant advantage is its flexibility over other spatial interpolation and averaging approaches. Another pro of this technique is evaluating the estimation error value (Journel and Huijbregts, 1976; Journel, 1989). The geostatistical tools can be considered numerical techniques, describing spatial features and using predominantly random models as time series analysis.

Figure 1 shows the step-by-step flowchart of geostatistical estimation or simulation. The flowchart shows that data distribution and the outlier values should be evaluated besides confirming the spatial correlation. Assessing the statistical distribution nature of raw data in estimating and recognizing the population's statistical features helps use them and more appropriately analyze the results. If the statistical distribution of the data does not follow the Gaussian distribution, linear geostatistical tools may show biased results. Therefore, evaluating the data distribution is necessary before any modeling.

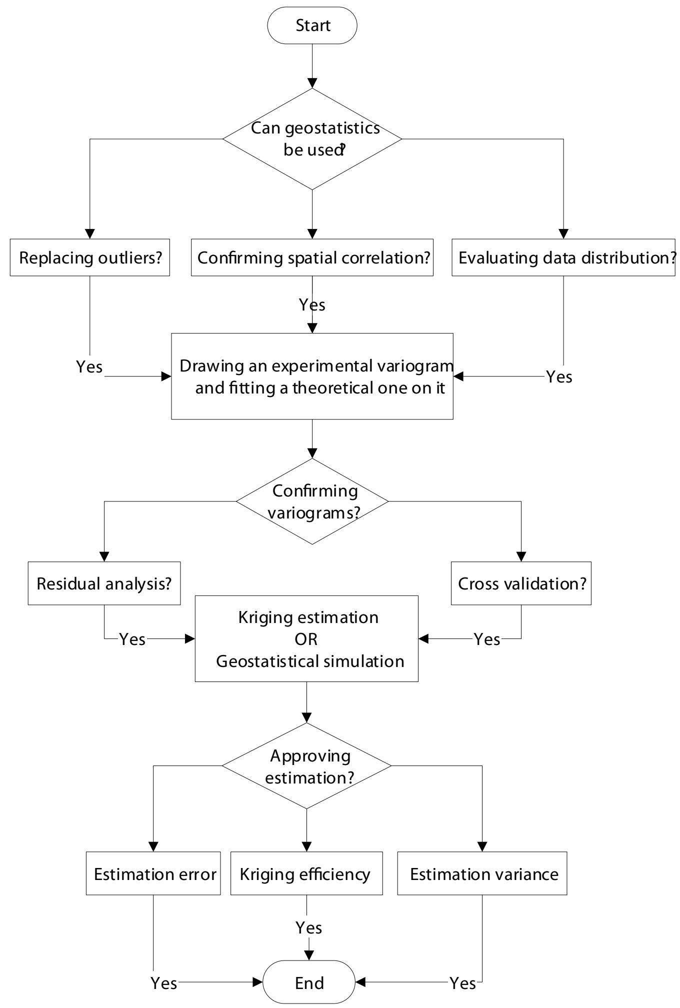

Figure 1. Step-by-step flowchart of geostatistical estimation (Karami et al., 2021)

Also, stationarity is a significant concept in geostatistics, permitting the statistical inference of probability. In the following, geostatistical concepts, including detecting outliers, stationarity hypothesis, variography, geometric and zonal anisotropy, estimation, and simulation, are assessed in the individual subsections in more detail.

2.1 Detecting Outlier Data

Outlier data are inconsistent observations that do not follow the pattern in the actual data set. These inconsistent observations may be due to error or natural variability. Evaluating the sample containing outlier data reveals significant gaps between such observations and actual data and deviations between them. Available statistical tests must confirm the presence of these data because their existence considerably impacts the results of geostatistical approaches (Bárdossy and Kundzewicz, 1990).

Outliers are distributional and are evident as large or small values in a histogram or box and whisker plot (Tukey, 1977), or they are spatial, whereby the values are different from other values in their spatial vicinity. The latter need not be distributional outliers and may be identified from a pixel map of values (Kerry and Oliver, 2007) or by other more elaborate approaches such as internal and external analysis techniques (Gnanadesikan and Kettenring, 1972). Davies and Gather (1993) expressed that an approach to identifying outliers is that they have a different distribution from the remaining observations. These researchers suggested an actual outlier identifier based on Hampel's technique.

Cressie (2015) stated that if the skewness is outside the bounds of ±1, the histogram or box and whisker plot can be evaluated. If outliers cause asymmetry, this is more obvious in a schematic box and whisker plot than a histogram, and the extreme values should be assessed further. If they result from errors in the assembly of data or laboratory analysis, it is better to eliminate them from the data. Bárdossy and Kundzewicz (1990) stated that a detection procedure for outliers could be at the stage of determining the experimental variogram. If the squared difference cloud is assessed for certain distance classes, it can be found that most outliers of the cloud are related to the same observation station.

2.2 Stationarity Hypothesis

The geostatistics is based on random functions (RF), and the measurements are assumed to be a realization of a specific RF for a given position (Matheron, 1971); therefore, some restrictions, called stationarity hypothesis, should be made on the data (Journel and Huigbregts, 1976). A central assumption in geostatistics is the stationarity of the process. However, the spatial variability of many natural phenomena depends on the local geology, which is often nonstationary (Brenning, 2001).

The first aspect of stationarity is choosing the domains to perform geostatistical calculations. Abrupt changes in the regionalized variable are often captured in the domain definition. The second aspect is the location dependence of statistics within the domain. There is only access to some details of the complex geological processes that lead to the studied variable. One can observe the result and make reasoned judgments about the location dependence of statistical parameters. In a stationary procedure, the unconditional joint probability distribution does not change when shifted in time or space (Dias and Deutsch, 2022).

The stationarity order depends on the order of the statistical moments required to be stationary. A variable is stationary if its distribution is invariant under translations. If the first moment, the mean, is invariant under translation the RF, is first-order or weak stationarity. Strict/strong stationarity requires that all distribution moments remain invariant under translations. This condition is rare in the natural world and is difficult to verify from limited experimental data. The geological phenomena is a second-order stationary if only the first two moments, mean (equation (1)) and covariance (equation (2)), are constant (Journel and Huijbregts, 1976; Davis and Sampson, 1986; Armstrong, 1998; Vieira et al., 2010; David, 2012; Dias and Deutsch, 2022).

$$m = m(x) = E\,[Z\,(x)] \qquad (1)$$

$$C(h) = E\,\{[Z(x)\,Z(x+h)\} - m^2 \qquad (2)$$

Z(x) and Z(x+h) represent two values of a given variable in the two different points, (x) and (x+h).

The stationary regionalized variables satisfy the intrinsic hypothesis, but the converse is not always true, as intrinsic variables can be nonstationary. The practical benefit of using an intrinsic regionalized variable is a broader choice of the possible variogram models compared to the cases of second-order stationarity. In practice, a decision on the stationarity of a given regionalized variable is made in conjunction with consideration of the uniformity of a specific variable and the scale at which the variable is stationary (Abzalov, 2016). In the presence of a trend, when the first-order moment varies over the domain, making it nonstationary, it must be removed, and the remaining residuals are treated as stationary (Harding and Deutsch, 2021).

Geostatistical modeling is performed based on the variogram tool after examining the statistical distribution of data, identifying and correcting outlier values, and evaluating stationariy.

2.3 Variography

One of the hypotheses in geostatistics is that in a specific direction for each distance, the difference variance between the two regionalized variables is a constant and depends on their coordinates. A variogram is a tool for quantifying the observed relationship between sample values and their neighborhood (Lin et al., 2001). It describes the data's spatial variability and structure and the difference's variance of a regionalized variable. According to the set of values in the sampled locations, variogram-based spatial estimation of kriging or simulation shows how the variance of the data values at different points changes with distance (h in Figure 2a) (Matheron, 1971).

The experimental variogram, γ (h), expressed in equation (3), is a graphical representation of the mean square variability between two neighborhood points (for example, in estimating the values of the WPT, the two points Lu and Lu+h in Figure 2a) as a function of the distance (h).

$$\gamma(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z(x_i) - Z(x_i + h)]^2 \qquad (3)$$

h is the lag distance that separates the pair of points, Z(x) and Z(x+h) are the regionalized variable values at the point x and a point at the distance x+h, respectively, and N is the number of data.

After determining the experimental variogram, the most suitable theoretical model (Figure 2b) that best describes the sampled experimental data can be chosen (Kresic, 2006). In the earth sciences, samples taken from close distances are more similar than those taken farther apart. Variogram uses significant parameters in modeling variables, including sill, nugget effect, and range. Figure 2 shows a schematic model of a theoretical variogram and its parameters in a blocked dam site. The nugget effect (C_0), sill ($C+C_0$), and range (R) are used in the kriging equation as well as the simulation process; therefore, it is significant to evaluate how well these parameters can produce actual data (Jolly et al., 2005).

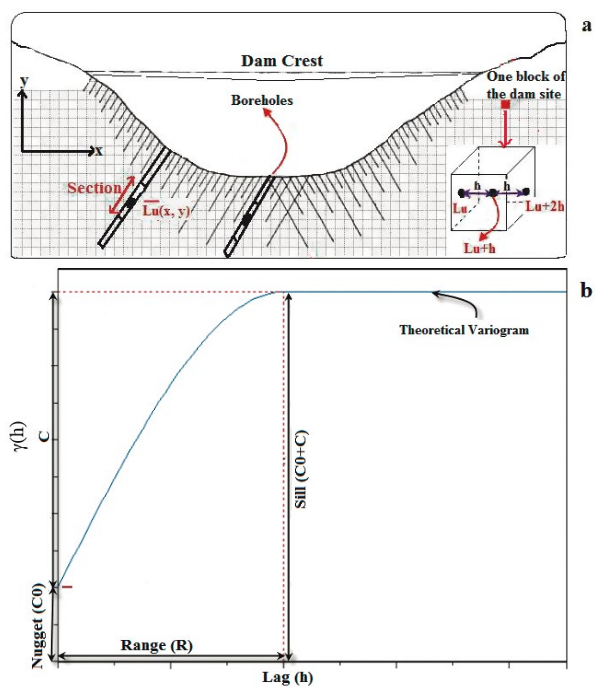

Figure 2. a) Schematic representation of a blocked dam site, showing the concept of variogram with the lag value of h in a block of it. b) Modeling the theoretical variogram demonstrated in the dam site.

3.2.1 Geometric and Zonal Anisotropy

A thorough geostatistical analysis includes a careful study of how the data's second-order variation depends on the relative orientation of data locations. If the second-order dependence between observations at any two locations is a function of only the Euclidean distance between the locations, the second-order variation is isotropic; otherwise, it is anisotropic (Zimmerman, 1993). In spatial statistics, the assumption of isotropy is due to two reasons. (i) Isotropic models are mathematically more straightforward to build than anisotropic ones, and the estimation of their parameters is more feasible, especially when the sample size is small. (ii) Anisotropic models are often restricted to the classical geometric and zonal anisotropy models, which transform the coordinates by a transformation matrix (Chilès and Delfiner, 2012).

Anisotropy models are usually geometric, zonal or separable. These models are composed to provide more complex anisotropies (Allard et al., 2016). Journel and Huijbregts (1976) define zonal anisotropy as any kind of anisotropy other than geometric. Isaaks and Srivastava (1989) define zonal anisotropy as an anisotropy for which the sill varies with direction, but the range does not. The existence and type of anisotropy can be detected by constructing directional variograms (Guedes et al., 2008).

The most common anisotropy among the phenomena with spatial dependence is geometric anisotropy (Journel and Huijbregts, 1976; Isaaks and Srivastava, 1989; Zimmerman, 1993). Geometric anisotropy means the correlation is stronger in one direction than another. There are two parameters defining geometric anisotropy: the direction of greater spatial continuity, expressed in the azimuth system, and the geometric anisotropy factor, being greater than one (Diggle et al., 1998). The finding of geometric anisotropy and its subsequent incorporation into the spatial model is significant in estimating the non-sampled locations (Guedes et al., 2013).

2.3.2 Variography Validation

A notable shortcoming of previous studies is the inadequate validation of variograms, which serve as the foundation for geostatistical estimations and simulations. The accuracy of variogram calculation and plotting is imperative, as any geostatistical procedure founded upon incomplete or inaccurate variograms loses its scientific validity. Before presenting any estimated or simulated model, evaluating the theoretical variogram parameters based on the original data is necessary. It is done in two ways: residual analysis and cross-validation.

The residual analysis represents the estimated value minus the actual value and evaluates the histogram of the residual value for each location. The residual histogram should show a normal distribution with a mean of zero to confirm the correctness of the variogram statistically (Rossi and Deutsch, 2014; Gong et al., 2014; Jalali et al., 2016, 2019). Figure 3 shows the residual analysis to estimate the permeability of a dam site. As it shows, the mean of the residuals is approximately zero.

Cross-validation is performed by removing a sample and estimating its value through the remaining data, and the estimated values are plotted against the actual values. If the estimated values equal the actual data, the plotted points are precisely on the line y = x. This process continues until all points in the database are estimated (Kerry and Oliver, 2007; Lark, 2000; Boukouvala and Ierapetritou, 2012; Jalali et al., 2016; Clark, 1979; Jolly et al., 2005; Jalali et al., 2019; Adhikary et al., 2011). Figure 4 shows an example of cross-validation to estimate the permeability of a dam site.

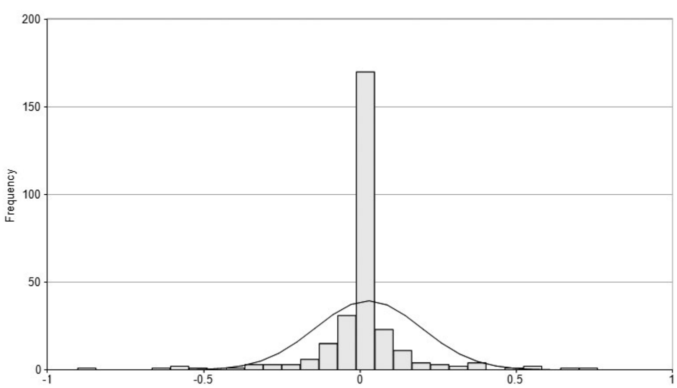

Figure 3. An example of a residual analysis diagram with a mean of zero

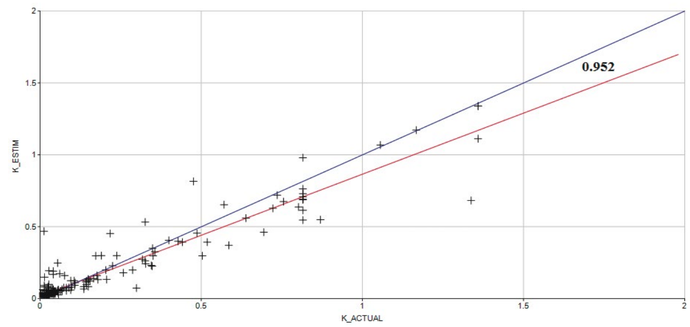

Figure 4. An example of a cross-validation test.

2.4 Kriging Estimate

After modeling and validating the variogram, the estimation process is carried out based on the spatial structure of the environment. Estimates are made at unsampled locations by the kriging interpolation technique. Kriging is one of the most effective geostatistical methods for interpolation and contouring in various fields considering spatial variance, location, and sample distribution (Kresic, 2006). It is an alternative to other point interpolation techniques based on spatial regression of collected samples from a spatial domain. This method has been widely used in mapping geoscience features and has proven to be a powerful interpolation technique in recent decades.

Kriging uses a weighted linear combination to determine unknown values. This method obtains weights for applying any sampled value, which leads to optimum and unbiased estimates. It is done by minimizing the error variance and zeroing the average prediction errors. Regardless of the general statistical properties of the estimates, emphasis is placed on the local accuracy and proximity of the unknown values to the actual values (Webster and Oliver, 2007; Smith and Konrad, 2011). Closer and more samples give more certainty to the estimates. Estimates should reflect the entire target area and the complete range of values, not just certain areas or specific values.

2.4.1 Validating Kriging Estimates

After the kriging estimate, it should be validated using the estimation variance (EV), estimation error (EE), and kriging efficiency (KE). Estimation variance measures the reliability of predictions and is a function of the variogram shape, the sample structure, and the region in which the observation is made, which may be approximated as a point or an area (Journel and Huijbregts, 1976).

The challenge of estimation variance deserves scrutiny, as it should ideally equal or be less than the actual data variance in geostatistical analyses. Estimation variance exceeding the actual variance lacks a logical or scientific basis, compromising the reliability of estimation and simulation outcomes. It is suggested that its variance should be calculated after estimation, and if it cannot be confirmed, the estimate errors should be fixed. The estimation variance of kriging, , is expressed as equation (4).

$$\sigma^2_E = 2\,\gamma(v, V) - \gamma(v, v) - \gamma(V, V) \qquad (4)$$

V is the estimated location, and v is the data point. $\gamma(v, V)$ is the variogram mean when its tail is fixed on the actual data and its head on the nodes to be estimated. $\gamma(v, v)$ is the variogram mean when its tail and head are fixed on actual data. $\gamma(V, V)$ is the variogram mean when each estimated node's head and tail are set on (Journel, 1989).

Also, one of the advantages of geostatistics is the estimation error evaluation. The minimized estimation variance (equation 4) can only be equal to the local estimation error if the error distribution is Gaussian and does not depend on the actual values of the sample (Deutsch and Journel, 1992). The kriging estimation error is expressed as equation (5).

$$EE = Z\,\sigma/Z^{*}\sqrt{N} \qquad (5)$$

σ is the standard deviation, N is the number of samples, and Z* is the estimated value. Z is the confidence level coefficient, equal to 2 at the 95% confidence level and 1.96 at the 90% confidence level.

Krige (1996) proposed kriging efficiency to evaluate the benefits of the kriging method. It is expressed as equation (6) and will be between zero and one.

$$KE = (BV - KV)/BV \quad (6)$$

BV is the theoretical block variance; KV (kriging variance) is the variance between the kriged and actual grades. The best estimators give KV close to zero and KE close to one (Snowden, 2001).

2.5 Geostatistical Simulation

One of the chief challenges in kriging estimates is smoothing the data range. It means kriging underestimates the upper quartile of each data set and overestimates the lower quartile. It makes detecting very low or high values in the estimated data impossible. Smoothing expresses that the created model variability is much less than the actual data set and minimizes the EV (Marinoni, 2003). Geostatistical simulation techniques can solve the smoothing effects of kriging (Ramazi and Jalali, 2015; Jalali et al., 2019).

Simulation techniques provide models that do not show any smoothing. The simulation principle is that each iteration provides alternative realizations. Therefore, repeating hundreds of simulations will likely give hundreds of different realizations for a network point and the whole model. These approaches attempt to provide the best realization of the predicted data by comparing the estimated and actual data. Hence, the simulations do not offer reasonable local estimates but are a good measure of spatial uncertainty (Deutsch and Journel, 1992).

The simulation generates identical probabilistic values for a variable corresponding to existing in-situ measurements. The simulated values have average and variograms similar to the original data and may correspond to them at the measurement points. Instead of producing an optimum prediction, simulation focuses mainly on reproducing observation variations (Sterk and Stein, 1997). Geologists use simulation to visualize fluctuations in significant geological patterns.

The geostatistical simulations can be Validated by assessing the consistency of simulation realizations and generating histograms and variograms of sampled and simulated data. The congruence of non-directional histograms and variograms between the two datasets is a vital indicator of the validity and accuracy of the simulations.

3. Hydrogeology of Dam Sites

This section describes the usual hydrogeological conditions governing the dam sites. Dam engineering is a complex and highly uncertain science, recognizing and reflecting the nature of many significant design inputs. It is a highly specialized activity that relies on many scientific disciplines and balances them with engineering judgment; therefore, it can be considered a challenging field. An accurate geotechnical study of a dam site determines the geological structure, stratigraphy, faulting, cracking, and jointing. It establishes the earth and groundwater conditions adjacent to the dam site, including abutments (Novak et al., 2017).

Many dams worldwide are over 100 meters high, and some are over 200 meters high. The first apparent danger of constructing a dam is that an unnaturally steep hydraulic gradient is created at an unnatural rate across the foundation rock and abutments as the water level rises to this height (Ford and Williams, 2013).

Hydrogeological evaluation is the foundation for valid permeability judgments and sealing measures, focusing on areas that need rock improvement. Therefore, one of the significant hydrogeological problems in the vicinity of the dam reservoir area is the existence of a valley. Suppose any permeable rock zone in the dam reservoir area acts as a hydraulic connection, and the main groundwater level is lower than the desired reservoir level. In that case, it can drain water from the reservoir. This complexity cannot always be detected by geological mapping alone. This is especially true in the case of the groundwater surface position. Thus, piezometers should be installed whenever groundwater fluctuates under the reservoir surface (Ewert, 2012).

Determining groundwater level is one of the chief factors in the geotechnical analysis. In most dams, piezometers are installed to assess groundwater conditions, water levels and seasonal fluctuations, the location of aquifers, and the study of pore water conditions in the dam and its abutments. Also, observation wells are located in different parts of the dam site to study groundwater level fluctuations in different seasons before and after impoundment (Aghda et al., 2019). Furthermore, whether the groundwater level rises steeply under the slopes adjacent to the dam site or moves to the abutments on the river level should be checked. The high or deep position of the groundwater level provides valuable information about the average permeability of the rock mass (Ewert, 2012).

One of the most significant hydrogeological aspects is determining whether there is a difference between more or less permeable zones throughout the dam site. This problem is examined from two aspects: if a zone with low permeability and a thickness of several tens of meters creates a dense barrier, the dam or the impermeable element will be built on this zone. Suppose there is a combination of layers or banks with different permeabilities. In that case, it should be determined whether the results of permeability tests indicate the overall permeability of a more extensive section of rock mass (Ewert, 2012).

Forming water-carrying openings along joints and other discontinuities is a long process. When a specific network of pathways is developed, the direction of groundwater flow cannot easily change, yet it retains its penetration even in an impounded reservoir condition. Groundwater level and fluctuations related to rainfall over an extended period, including at least one dry and one rainy season, should be measured. Also, interpretation of maps and hydrographs leads to the following results: moderate rock permeability, sections that may be excluded from treatment, the presence of only one or more groundwater regimes, and natural groundwater response to rainfall, especially interpretation of groundwater behavior when impounding a dam reservoir (Ewert, 2012).

The dam sites' most significant hydrogeological and geomechanical features include WPT, the permeability of various geological units and formations, transmissibility, hydraulic gradient, leakage value, fracture distribution, and RQD. In the following, these features are introduced and discussed using geostatistics to evaluate and estimate each.

3.1 Water Pressure Test

One of the most common ways to study the hydrogeology of a dam site is to perform Lugeon tests (WPT) on boreholes. WPT results are one of the significant characteristics of rock formations and indicate the environment's ability to transfer groundwater. The WPT measures the rock permeability in separate sections of boreholes in the rock mass by increasing the pressure and measuring the inflow in each pressure stage (Milanovic, 2004; Kresic, 2006).

The WPT results help to estimate the permeability of different zones of the dam foundation. However, due to economic or topographic constraints, this test is performed on a limited number of wells in an area of the dam site. Accordingly, the information obtained from these tests does not represent the whole region because the fluid properties are intrinsically very heterogeneous and unpredictable (Akhondi and Mohammadi, 2014). Therefore, the WPT values are an indicator of the permeability of the dam sites.

Ewert (2005) stated that such a conversion could be done when engineers need the converted permeability coefficient values from the WPT in evaluating the foundation seepage of the dam sites. Many researchers have proposed several equations to convert WPT results to permeability coefficients (Richter and Lillich, 1975; Gilg and Gavard, 1957; Shimizu et al., 1985; Barton and Quadros, 2003; Fransson, 2004; Hvorslev, 1951; Moye, 1967; Ahrens and Barlow, 1951). To convert the WPT results to the permeability coefficient and to find the correlation between the values of these two variables in describing and estimating the permeability of the foundation rock, Aherns and Barlow's (1951) equation (7) is suggested.

$$\begin{aligned} K &= 5.9918 * 10^{-8} * N.\log\left(\frac{L}{2r} + \sqrt{\frac{L^2}{4r^2} + 1}\right) && 1 \leq \frac{L}{r} < 10 \\ K &= 5.9918 * 10^{-8} * N.\log\left(\frac{L}{r}\right) && \frac{L}{r} \geq 10 \end{aligned} \quad (7)$$

K is the permeability coefficient value, N is the Lugeon number, L is the length of the test section in meters, and r is the influence radius of the water flow in meters.

3.2 Permeability

Permeability coefficient, K, or hydraulic conductivity, is the most significant hydrogeological feature in preparing the groundwater model in a dam site. It represents the amount of water movement in voids and cracks and largely determines the groundwater flow and its head distribution (El Idrysy and De Smedt, 2007). The permeability coefficient is calculated from field and laboratory tests. In the dam sites, it is usually based on the WPT results (Ewert, 2005). Due to the complex pattern of discontinuities, it will be almost impossible to determine the permeability of rock mass if the appropriate test method is not used.

Intact or massive rocks are nearly impermeable; jointed rock masses may be permeable depending on their discontinuity characteristics. Rock mass permeability can be determined by considering discontinuity conditions, including resistance, openness, roughness, filling, weathering conditions, distance of discontinuities, and RQD (Kayabasi et al., 2015). The standard classification of permeability is provided in Table 1.

Table 1. Permeability classification based on WPT values and rock mass permeability coefficient.

WPT values (Lu)	permeability coefficient, K(m/s)	Permeability class
<1	$<10^{-9}$	Impermeable
1-5	10^{-9}-10^{-5}	Low permeability
5-25	10^{-5}-10^{-2}	Medium permeability
>25	10^{-2}-1	High permeability

3.3 Transmissibility

Transmissibility is the water's flow rate through a vertical aquifer strip. It controls groundwater flow, provides a general idea of an aquifer's water-producing efficiency and describes the aquifer's capacity to transmit groundwater wholly in its entire saturated thickness. Increased transmissibility can indicate increased hydraulic conductivity, porosity and appropriate interconnected pore spaces. Therefore, this increase results in a high transmissibility value, which approves a high permeability presupposition (George et al., 2017).

Transmissibility is one of the basic hydrogeological properties, and determining its spatial variations is significant in groundwater modeling. Appropriate modeling of preferential flow paths and their transfer behavior requires using transmissibility fields that reproduce the spatial variability patterns observed on Earth. The transmissibility fields sometimes have significant uncertainties, including unknown and complex variations in the study area's observed values of measurable properties (Lin et al., 2001).

3.4 Hydraulic Gradient

The hydraulic gradient is the slope of the water table or potentiometric surface, meaning the variation in water level per unit distance along the direction of maximum head reduction. The hydraulic gradient is the driving force that causes the groundwater to move toward maximum reduction of the total head. Its value is determined by measuring the water level in several wells (Cheremisinoff, 1998). Hydraulic head gradients control the groundwater flow in aquifers, and their determination is a significant part of any groundwater assessment at a dam site.

3.5 RQD

RQD is the percentage of recovered core length with fragments larger than 10 cm relative to the total core length. It is a modified core recovery percentage in which unrecovered core, small pieces of rock, and weathered rock are not considered to reduce rock quality with these characteristics. However, its most significant application is that it is an exploratory tool and acts as a red flag to identify areas with low RQD that require further investigation and may require additional boreholes or other exploration works. Experience has shown that the RQD red flag and subsequent studies often deepen foundation surfaces and complete reorientation or displacement of engineering structures, including dam foundations, tunnel entrances, underground caves, and power plant facilities (Deere et al., 1988).

It is commonly used in core logging and is often the only technique for measuring the degree of jointing along the core drill hole (Palmstrom, 2005). The most significant limitation of RQD is that it provides no information on the core pieces <10 cm; it does not matter whether the discarded pieces are earth-like materials or fresh rock pieces up to 10 cm in length. Like all types of one-dimensional measurements, RQD is directional, but its definition makes it more sensitive to the hole or line direction than others (Choi and Park, 2004).

3.6 Leakage

The stored water in the dams always looks for paths with the least resistance; hence, one of the problems in dam sites is water leakage after impounding. Assessing and predicting the water leakage value can help prevent such occurrences. Many dam sites have reported excessive leakage, especially in karst areas worldwide (Milanović, 1981; Turkmen et al., 2002). Therefore, performing studies and following up on corrective operations is normal to reduce the leakage value. Lack of accurate view of the conditions in the dam site, especially before construction, is one of the main reasons for water leakage (Mohammadi et al., 2007).

The leakage value in a dam site depends on several factors, including the dam height, the valley shape, and the water level in the reservoir. Some indicators for increasing water leakage include increasing the flow rate of downstream springs and the water level in observation wells. Seepage paths are usually along karst conduits, bedding planes, open joint systems, or their intersections (Uromeihy, 2000; Turkmen et al., 2002; Thomas, 1978).

3.7 Fracture Distribution

The general term for a fracture to describe a geological discontinuity includes faults, joints, fissures, cleavage plates, and cracks. Fractures often act as the preferred flow paths for fluids and significantly affect a rock mass's mechanical and hydraulic properties (Adler and Thovert, 1999; Haneberg et al., 1999; Faybishenko et al., 2000). Fractures are one of the most significant features of rocks on micro to megameter scales due to various genetic mechanisms such as cleavage, tensile, and shear. Fractures are studied to understand the history of tectonic and geological structures and describe rock masses' hydraulic and mechanical properties to explore natural resources, rock structures, and storage structures, including dam sites. Accurate imaging of the fracture system and an acceptable 3-D distribution model are essential in all areas. It can help clarify fracture-related phenomena and environments (Koike et al., 2012).

4. Geostatistical Applications in a Dam Sites' Hydrogeology

Applying geostatistical approaches to estimate the hydrogeological characteristics of the dam site plays a significant role in reducing errors and study costs. Via geostatistical tools, the distribution of quantitative values such as permeability for each region point is obtained to analyze hydrogeological or geomechanical conditions accurately. This research attempts to review and evaluate the geostatistics applications in spatial analysis of the dam site's hydrogeological features.

Measuring hydrogeological data in tunnels and underground structures, including mines, quarries, dam sites, and foundation drillings, is essential. Different geostatistical approaches can estimate and simulate the hydrogeological data at the dam sites in heterogeneous porous media. They quantify uncertainty analysis and can be used in systems with high spatial variability in hydrogeological properties (Assari and Mohammadi, 2017). Table 2 summarizes the fields in which geostatistical approaches are used.

Table 2. Summary of geostatistical methods used in the dam site

Studied variables in the dam site	Geostatistical method	Performance of the geostatistical method	Where used the method	Who used the method
WPT	Simple kriging	• Better performance than ordinary kriging	Tangab Dam site, Zagros region, Southwest Iran	Akhondi and Mohammadi (2014)
	Sequential indicator simulation	• Better performance than sequential Gaussian simulations • Suitability in simulating the abnormal distribution of WPT values	Karun Dam site, Iran	Assari and Mohammadi (2017)
Permeability	Multiple indicator kriging	• No dependence on data distribution • Sensitivity to outlier values • It takes a long time to calculate and model the experimental variogram and solve the kriging system separately for each threshold value.	Azad Dam Headrace Tunnel, Kurdistan, Iran	Aalianvari et al. (2018)
	Ordinary kriging	• Lack of accurate point predictions of permeability • Inefficiency in estimating the spatial distribution of permeability • Reflecting the local variations of permeability using dependent variables	Tangab Dam site, Zagros region, Southwest Iran	Akhondi and Mohammadi (2014)
	Cokriging	• Accurate estimate of the spatial distribution of permeability	Trifa Aquifer, Morocco	El Idrysy and De Smedt (2007)
	Sequential Gaussian simulation	• Lack of accurate point predictions of permeability • Better reproduction of feature distribution and local contrasts	Tono area in Japan, overlain by Cretaceous granite	Koike et al. (2015)
			Kikuma granite, Japan	Koike et al. (2012)
	Simulated annealing	• Lack of accurate point predictions of permeability • Better reproduction of feature distribution and local contrasts	Tangab Dam site, Zagros region, Southwest Iran	Akhondi and Mohammadi (2014)
Hydraulic gradient	Ordinary kriging	• The best linear unbiased estimator for estimating the hydraulic gradient	Wolfcamp aquifer, northern Texas	Philip and Kitanidis (1989)
Dam reservoir leakage	Ordinary kriging	• High accuracy in estimating reservoir leakage	The Azad pumped storage power plant, Kurdistan, Iran.	Aalianvari et al. (2013)
Transmissibility	Ordinary kriging	• Provide the best estimate of transmissibility • Inefficiency in case of low spatial variability on a small scale • Generate smoother maps than simulation approaches	Dulliu area in Yun-Lin county, Taiwan	Lin et al. (2000)
			The southeast of Yun-Lin County and the north of Chia-Yih County, Taiwan	Lin et al. (2001)
	Log-normal ordinary kriging	• Inefficiency in the case of small-scale low spatial variability	Dulliu area in Yun-Lin county, Taiwan	Lin et al. (2000)
	Cokriging	• The lowest satisfaction based on the cross-validation results	A highly fractured metamorphic aquifer, Man-Danane region, Ivory Coast, Africa	Razack and Lasm (2006)
	Sequential Gaussian simulation	• Generate the spatial structure of the studied data • No smoothing in a site with high data variability	Dulliu area in Yun-Lin county, Taiwan	Lin et al. (2000)
	Simulated annealing	• Reproduction of statistics and spatial variation of transmissibility • Identify global spatial correlation patterns of transmissibility • Complete display of geological features of the studied area	The southeast of Yun-Lin County and the north of Chia-Yih County, Taiwan	Lin et al. (2001)
RQD	Sequential indicator simulation	• No dependence on the variable normality • Better performance than sequential Gaussian simulation	Karun Dam site, Iran	Assari and Mohammadi (2017)
	Ordinary kriging	• Successful estimate of geomechanical features of tunnels and dam sites	Azad Dam Headrace Tunnel, Kurdistan, Iran	Aalianvari et al. (2018)
Distribution of fractures	Ordinary kriging	• Satisfactory results in estimating fractures' permeability	Kikuma granite, southwest Japan	Koike et al. (2012)
	Sequential Gaussian simulation	• Assess the location of fractures	Tono area in Japan, overlain by Cretaceous granite	Koike et al. (2015)
			Kikuma granite, Japan	Koike et al. (2012)

The following section describes the application of geostatistics in each field in detail.

4.1 Application of Geostatistics in Dam Sites' WPT

The WPT is a significant feature of rock formations and indicates the environment's ability to transfer groundwater. The test values in the spatial range more or less constantly change and can be assumed as a regionalized variable for geostatistical studies (Akhondi and Mohammadi, 2014).

Akhondi and Mohammadi (2014) used simple (equation 8) and ordinary kriging approaches to determine the spatial variability of the karst formation permeability at the dam site. After hydrogeological studies to evaluate the geostatistical simulations, simple kriging with Gaussian variograms was the best technique to estimate the WPT values.

Equation (8) shows the generalized linear regression algorithm in simple kriging (Deutsch and Journel, 1992).

$$[Z^*_{SK}(u) - m(u)] = \sum_{\alpha=1}^{n} \lambda_\alpha(u)[Z(u_\alpha) - m(u_\alpha)] \tag{8}$$

Z(u) is the WPT value at location u, u_α is n data location, and m(u)=E{Z(u)} is the expected value of location-dependent Z(u). λα(u) is the simple kriging weight in place u, and $Z^*_{SK}(u)$ is the simple kriging estimator. In simple kriging, it is assumed that the mean values of the WPT in the study range are constant and known.

4.2 Application of Geostatistics in Dam Sites' Permeability

The permeability measurement is the basis for measuring many hydrogeological parameters in the dam sites. Equation (9) shows the leakage flow in a 2-D plane (Manna et al., 2003).

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial h_{(x,y,t)}}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial h_{(x,y,t)}}{\partial y}\right) + q = \frac{\partial V}{\partial t} \tag{9}$$

K_x is the soil permeability in the x-direction, and K_y is the soil permeability in the y-direction in m/s. For isotropic soils, $K_x = K_y$. q is the inflow/outflow rate to the soil in m^3/s per unit area (m/s), and V is the total volume of water in m^3. t is the time interval in seconds, and $h_{(x, y, t)}$ is the total head at a point with coordinates (x,y) in meters. This equation (9) satisfies the flow continuity, i.e., the difference between inlet and outlet water volume equals the change in water storage volume.

Under steady-flow conditions, the inlet and outlet flow rates are equal, and no volumetric variation occurs; therefore, equation (10) equals zero. Equation (10) presents the leakage for steady-state conditions.

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial h_{(x,y,t)}}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial h_{(x,y,t)}}{\partial y}\right) + q = 0 \tag{10}$$

Khalili Shayan and Amiri-Tokaldany (2015) simplified equation (10) in a stable state for isotropic soils (K_x=K_y) and without any drainage. Equation (11) shows the leakage in the steady state for isotropic soils in the absence of drainage.

$$\left(\frac{\partial^2 h_{(x,y,t)}}{\partial x^2}\right) + \left(\frac{\partial^2 h_{(x,y,t)}}{\partial y^2}\right) = 0 \tag{11}$$

Permeability largely determines groundwater flow and its head distribution. Reliable permeability values are challenging due to the lateral and vertical heterogeneities in water-carrying geological strata (El Idrysy and De Smedt, 2007). Also, permeability data may indicate considerable uncertainty, indicating the complexity of spatial variability (Lin et al., 2001; Tartakovsky, 2013). According to the above equations, K has spatial variability, a characteristic of regionalized variables in geostatistics.

Multiple indicator kriging is one of the simplest and most subtle geostatistical methods developed by Aalianvari et al. (2018). In the indicator kriging, the probability that the permeability value in the estimated block or point is less than a given threshold value is estimated (Öztürk and Nasuf, 2002).

First, the initial permeability data must be converted to indicator values based on the equation (12) conversion function.

$$\mathbf{i(x)} = \begin{Bmatrix} 1 \text{ if } z(x) \le Z_c \\ 0 \text{ if } z(x) > Z_c \end{Bmatrix} \rightarrow \quad \mathbf{i(x)} = \begin{Bmatrix} 1 & \text{if } K \le 1 \\ 0 & \text{if } K > 1 \end{Bmatrix} \tag{12}$$

Z_c is the assumed threshold value. In this case, equation (13) calculates the estimated value of the indicator permeability at each point.

$$\mathbf{i^*(x_0)} = \sum_{j=1}^{n} \lambda_j . \, i(x_j) \tag{13}$$

$i(x_j)$ is the indicator value in the x_i coordinates, and λ_j is the indicator kriging weight for the sample in the x_j coordinates, participating in estimating the point x_0. The estimated permeability value, (x), should vary between a minimum of zero and a maximum of one, indicating the probability that the permeability value in the estimated block is less than the threshold.

The advantages of this technique are non-dependency on data distribution and sensitivity to outlier values. The disadvantage of this technique is the time required to calculate and fit the model to the experimental variogram and solve the kriging system separately for each threshold value (Aalianvari et al., 2018).

Indicator kriging in estimating the permeability of the dam site relies on the following four main features (Gavinhos and Carvalho, 2017):

- It is non-parametric and independent of the distributions' shape.
- It is helpful for highly skewed data such as WPT results as a nonlinear interpolator.
- It is less affected by the high smoothing of the variable.
- It provides a direct probabilistic estimate of the areas needing further treatment in the dam foundation.

El Idrysy and De Smedt (2007) stated that more than the kriging technique is needed to accurately estimate the spatial distribution of permeability. These researchers estimated the regionalized distribution of permeability using kriging and cokriging with a hydraulic gradient slope because, based on Darcy's law (Q = KA (∂h/∂l)), the slope of a hydraulic gradient (∂h/∂l) with permeability (K) is a related directly.

4.3 Application of Geostatistics in Dam Sites' Hydraulic Gradient

In equation (10), the expressions $\frac{\partial h_{(x,y,t)}}{\partial x}$ and $\frac{\partial h_{(x,y,t)}}{\partial y}$ are the hydraulic gradient values at a point with coordinates (x,y). The hydraulic gradient is dimensionless and scattered at points. This property is the feature of the regionalized variables; therefore, it can be considered a geostatistical variable with spatial variability.

Philip and Kitanidis (1989) are the only researchers to use the geostatistical technique of ordinary kriging to estimate the hydraulic gradient. Choosing the best unbiased linear estimator for the hydraulic head made it possible to estimate the hydraulic gradients directly and calculate the mean square estimation error.

4.4 Application of Geostatistics in Dam Sites' Leakage

In equation (10), q is the inlet/outlet flow rate into the soil in m^3/s per unit area (m/s). Controlling leakage from some dam parts is necessary for most long earthen dams. However, the quantity and quality of monitoring data are limited due to the complexity of the actual conditions (Deng et al., 2018). Estimating the water leakage of the dam site is necessary to select the optimum sealing technique (Aalianvari et al., 2013). According to the leakage flow, equation (10), the variable q has spatial variability. Its distribution can be modeled using geostatistical approaches and solve the quantity and quality problem of data to a large extent.

Researchers have always searched for available methods to use the best and most reliable techniques to assess the leakage value of the dam site. Smith and Konrad (2011) consider geostatistical methods an effective tool for describing dam heterogeneities that allow the areas with more severe leakage to be located.

Aalianvari et al. (2013) used a combination of geotechnical and geostatistical approaches to estimate the water leakage of the dam reservoir. Using the ordinary kriging technique, they estimated the potential of water leakage for reservoir walls in places without data.

Equation (14) presents the ordinary kriging algorithm that uses a weighted linear combination of sampled points in a neighborhood around location X_0.

$$Z^*(X_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{14}$$

$Z^*(X_0)$ is the estimated value of leakage q at location X_0, $Z(X_i)$ is the sample leakage at location X_i, and n is the number of leakage samples considered in the estimate. λ_i is the weight assigned to the i^{th} sample value (Webster and Oliver, 2007).

Comparing the results of the analytical and the finite element numerical methods in estimating the reservoir leakage showed the high accuracy of the geostatistical technique.

4.5 Application of Geostatistics in Dam Sites' Transmissibility

Transmissibility (T = K×b) expresses the ability to transfer water from the thickness of the table unit. The ability of an aquifer to transfer and move water (T) depends on two factors: permeability coefficient (K) and aquifer thickness (b). Hydrogeological data on transmissibility may indicate significant uncertainty, including complex and unknown changes in the observed values of measurable properties across the study area, showing complex spatial variability (Lin et al., 2001; Tartakovsky, 2013).

Lin et al. (2000) produced spatial transmissibility maps by sequential Gaussian simulation (SGS) and estimated them using log-normal ordinary kriging and ordinary kriging. The results proved that SGS could create the spatial structure of the data, and kriging estimate maps are smoother than other simulations. In a site with much variability, SGS with many realizations has more advantages than the ordinary and log-normal kriging approaches.

Lin et al. (2001) estimated and simulated transmissibility spatial maps to understand hydrogeological spatial features using ordinary kriging and simulated annealing approaches. Spatial maps of transmissibility estimate and simulation have shown that simulated annealing can reproduce statistics and spatial variation and identify global spatial coherence patterns. Also, ordinary kriging produced smoothed data that cannot record the spatial variability of measured transmissibility data.

Razack and Lasm (2006) estimated the transmissibility of highly fractured hard rock aquifers, a dense and well-connected fracture network with no specific orientation. Transmissibility variographic analysis proved that the variogram of actual data is much more structured than the normalized data. Also, the nugget effect in the variogram of normalized data was much higher than its value in the actual data. This result contradicted all previously published research on structural analysis of transmissibility. Although all previous geostatistical studies have explicitly concluded that a normal distribution improves estimates, this study proves otherwise. Another result of the research was that the cokriging method could not provide the best estimate, unlike previous studies.

The results of these studies prove that conditions governing the environment have a significant influence on hydrogeological parameters, and each hydrogeological environment is unique in its characteristics and spatial behavior.

4.6 Geostatistics Application in RQD and CR

The geotechnical features of RQD and Core Recovery (CR) are significant factors for engineering projects during a dam site's design, construction, and support. One of the most effective approaches to estimating these features is geostatistics. RQD and CR are features of rock formations and can assess cracks, fractures, and discontinuities (Aalianvari et al., 2018).

Variographic analysis of the geotechnical properties of rocks, which are regionalized variables, provides information on spatial variability, spatial correlation, and estimate of variables when analyzed in the geostatistics. Having this data, we can better plan the distance and distribution of boreholes and estimate the variables in places without data. These variables are calculated based on the properties of the rock and the change in properties of a pure regionalized variable (Santos et al., 2018).

In a study by Santos et al. (2018), RQD variograms showed very high variability in geostatistical modeling because the rock body was composed of minor filled-vein fractures with very high variability in different rocks. Hence, the RQD may be weak at intervals with good CR, and there is no clear correlation between them. It is also believed that analyzing regionalized variables of CR and RQD provides better variographic images for larger regions.

Assari and Mohammadi (2017) examined the relationship between the parameters of RQD and WPT. They concluded that their correlation is weak, and the overall trend is negative. This weak correlation is that RQD is a point variable, and the WPT is a regionalized variable with higher support that depends on the rock permeability at a higher volume. The researchers used the non-parametric geostatistical technique of sequential indicator simulation to simulate RQD and WPT. Sequential indicator simulation worked better than SGS to reproduce the spatial coherence of larger values. Another reason for choosing sequential indicator simulation is the suitability of this technique for simulating abnormal distributions of WPT values (Goovaerts, 1997). Sequential indicator simulation is done in several steps:

Modeling of variograms: Several indicator variograms are modeled using standard models, such as spherical models for each threshold.

Indicator kriging: A conditional cumulative distribution function (CCDF) for each cell at the simulation boundary is computed using the original and pre-simulated indicator kriging.

Simulation: A random path is chosen through all cells and for each cell in the simulation boundary. The simulated value of the CCDF is plotted and added to the conditioning dataset to estimate the CCDF of other cells.

Aalianvari et al. (2018) estimated the RQD using ordinary kriging to determine the variable value in unsampled locations and presented its distribution as 3-D maps. They stated that this technique could be used successfully in the geomechanical estimate of tunnels and dam sites, considering reducing costs and saving time.

4.7 Geostatistics in Predicting the Distribution of Fractures

The rock mass structure cannot be viewed directly in three dimensions, so many methods exist for extracting fracture features using different field surveys, including borehole, scanline, regional and analog surveys, digital photogrammetric techniques, and Laser scans. Each data source provides information on fracture characteristics. 2-D studies make it possible to evaluate fracture size, a critical parameter in studying the mechanical and hydraulic properties of rock, while the information provided by borehole measurements is related to the distribution of fracture direction, severity, and the presence of fracture clusters (Hekmatnejad et al., 2019).

Fracture is not just a geometric feature but has several properties, such as opening, filler minerals, roughness, displacement, and other properties that cause problems in their 3-D modeling. Fractures usually form clusters with dominant orientations, and their networks have a hierarchical structure (Koike et al., 2012).

Simulation of rock fracture distribution is significant and usual in various fields of earth sciences. Long and Billaux (1987) proposed geostatistical fracture modeling and used variograms of fracture density changes for fracture centers in 2-D space. Also, Viruete et al. (2001) used the geostatistical method to model, predict and quantify 2-D fracture systems. Koike et al. (2012) used the geostatistical method to simulate the distribution of fractures. They assessed the fractures of the site by SGS and used ordinary kriging to estimate the distribution of original values. Also, a simulated fracture system was available to estimate the permeability of the studied site, which was approximately consistent with the mean hydraulic test results.

Lei et al. (2015) proposed a new approach to upgrade the scale of 2-D fracture network models to preserve the geostatistical and geomechanical properties of a smaller-scale source fracture pattern. They used geostatistical methods for the source pattern to quantitatively interpret its topological complexity. Also, Koike et al. (2015), in a geostatistical study, integrated 3-D models of rock fractures from different hydraulic sources and properties to identify the relationships between fractures and permeability. Using an experimental relationship between permeability and fracture length, the researchers determined the range of fractures affecting hydraulic properties. Also, they integrated the 3-D permeability model resulting from SGS with the fractures of the geostatistical model derived from the borehole data. They identified the fracture sizes and directions affecting the permeable features. In the most recent research, Hekmatnejad et al. (2019) estimated fractures' direction and diameter distribution by the geostatistical method using borehole data.

Summaries and Conclusions

This review study aims to familiarize researchers with the application of geostatistical approaches in the hydrogeological sciences of dam sites. This study does not suggest any superior technique because the geological, hydrogeological and geotechnical conditions governing different dams worldwide are complex and different. Every researcher should be able to identify and use the best technique by analyzing the studied dam site. In other words, the best method for each dam site is unique, and it is impossible to propose only one superior process to researchers.

Dam engineering, characterized by its intricate nature and high levels of uncertainty, draws upon various scientific disciplines. Hydrogeological features hold a prominent position among the pivotal factors in dam design. However, these features introduce considerable uncertainties, primarily stemming from intricate spatial variations. Furthermore, the spatial distribution of pertinent variables assumes a significant role in scientific inquiry. Consequently, the estimation and simulation of these hydrogeological elements have been addressed through diverse methodologies, with geostatistics emerging as a prevalent and favored approach.

Applying geostatistical estimation and simulation within varying dam site contexts constitutes a formidable challenge. It necessitates the careful consideration of all crucial factors governing variations. In this intricate process, establishing logical connections among these parameters precedes the selection of appropriate geostatistical techniques.

Nonetheless, several limitations have been observed in previous studies that warrant further attention. A notable shortcoming is the inadequate validation of variograms. The accuracy of variogram calculation and plotting is significant, as any geostatistical procedure founded upon incomplete or inaccurate variograms loses its scientific validity. Additionally, estimation variance should ideally equal or be less than the actual data variance in geostatistical analyses. Estimation variance exceeding the actual variance lacks a logical or scientific basis. It is suggested that its variance should be calculated after estimation, and if it cannot be confirmed, the estimate errors should be fixed.

An intriguing aspect is the absence of validation efforts in previous hydrogeological studies of dam sites. Future geostatistical investigations in this domain present an opportunity for researchers to validate their findings rigorously. This can be achieved by systematically evaluating estimation error, estimation variance, kriging efficiency, and creating error-related maps. Furthermore, the validation of geostatistical simulations can be facilitated by assessing the consistency of simulation realizations and generating histograms and variograms of sampled and simulated data. The congruence of non-directional histograms and variograms between the two datasets is a vital indicator of the accuracy of the geostatistical simulations.

In the future, advanced validation methodologies can be integrated into geostatistical estimations and simulations for hydrogeological studies in dam engineering. The meticulous validation of results will enhance the credibility and utility of these approaches in guiding dam design and engineering practices. Moreover, the ongoing exploration of new technologies and advancements in geostatistical modeling can further refine our understanding of hydrogeological complexities at dam sites, ultimately contributing to safer and more effective dam engineering practices.

References

Aalianvari, A., Tehrani, M. M., & Soltanimohammadi, S. (2013). Application of geostatistical methods to estimation of water flow from upper reservoir of Azad pumped storage power plant. *Arabian Journal of Geosciences,* 6(7), 2571-2579. https://doi.org/10.1007/s12517-012-0528-3

Aalianvari, A., Soltanimohammadi, S., & Rahemi, Z. (2018). Estimation of geomechanical parameters of tunnel route using geostatistical methods. *Geomechanics and Engineering,* 14(5), 453-458. https://doi.org/10.12989/gae.2018.14.5.453

Abzalov, M. (2016). *Applied mining geology.* Springer International Publishing, Switzerland.

Adhikary, P. P., Dash, C. J., Chandrasekharan, H., & Bej, R. (2011). Indicator and probability kriging methods for delineating Cu, Fe, and Mn contamination in groundwater of Najafgareh Block, Delhi, India. *Environmental Monitoring and Assessment,* 176(1), 663–676. https://doi.org/10.1007/s10661-010-1611-4

Adler, P. M., & Thovert, J. F. (1999). *Fractures and fracture networks.* Springer Science & Business Media.

Aghda, S. M. F., GanjaliPour, K., & Esmaeilzadeh, M. (2019). The Effect of Geological Factors on the Grout Curtain Performance Analysis of Darian Dam Using the Results of Instrumentation Data in the First Impounding. *Journal of the Geological Society of India,* 93(3), 360-368. https://doi.org/10.1007/s12594-019-1185-x

Ahrens, T. P., & Barlow, A. C. (1951). *Report on Permeability tests using drill holes and wells.* United States Bureau of Reclamation Geology, G-97.

Akhondi, M., & Mohammadi, Z. (2014). Preliminary analysis of spatial development of karst using a geostatistical simulation approach. *Bulletin of Engineering Geology and the Environment,* 73(4), 1037-1047. https://doi.org/10.1007/s10064-014-0599-3

Allard, D., Senoussi, R., & Porcu, E. (2016). Anisotropy models for spatial data. *Mathematical Geosciences,* 48, 305-328. https://doi.org/10.1007/s11004-015-9594-x

Armstrong, M. (1998). *Basic linear geostatistics.* Springer Science & Business Media.

Assari, A., & Mohammadi, Z. (2017). Analysis of rock quality designation (RQD) and Lugeon values in a karstic formation using the sequential indicator simulation approach, Karun IV Dam site, Iran. *Bulletin of Engineering Geology and the Environment,* 76(2), 771-782. https://doi.org/10.1007/s10064-016-0898-y

Bárdossy, A., & Kundzewicz, Z. W. (1990). Geostatistical methods for detection of outliers in groundwater quality spatial fields. *Journal of Hydrology,* 115(1-4), 343-359. https://doi.org/10.1016/0022-1694(90)90213-H

Barton, N., & Quadros, E. (2003). *Improved understanding of high-pressure pre-grouting effects for tunnels in jointed rock.* In 10th ISRM Congress, September, ISRM-10CONGRESS.

Brenning, A. (2001). Geostatistics without stationarity assumptions within geographical information systems. Journal: Freiberg Online Geosciences. https://doi.org/10.23689/fidgeo-869

Boukouvala, F., & Ierapetritou, M. G. (2012). Feasibility analysis of black-box processes using an adaptive sampling Kriging-based method. *Computers and Chemical Engineering,* 36, 358-368. https://doi.org/10.1016/j.compchemeng.2011.06.005

Cheremisinoff, N. P. (1998). *Groundwater remediation and treatment technologies.* Elsevier.

Chiles, J. P., & Delfiner, P. (2012). *Geostatistics: modeling spatial uncertainty*. John Wiley & Sons.

Choi, S. Y., & Park, H. D. (2004). Variation of rock quality designation (RQD) with scanline orientation and length: a case study in Korea. *International Journal of Rock Mechanics and Mining Sciences*, 41(2), 207-221. https://doi.org/10.1016/S1365-1609(03)00091-1

Clark, I. (1979). *Practical geostatistics.* Applied Science Publishers, London, 129 pp.

Cressie, N. (2015). *Statistics for spatial data*. John Wiley & Sons.

David, M. (2012). *Geostatistical ore reserve estimation*. Elsevier.

Davies, L., & Gather, U. (1993). The identification of multiple outliers. *Journal of the American Statistical Association,* 88(423), 782-792. https://doi.org/10.1080/01621459.1993.10476339

Davis, J. C., & Sampson, R. J. (1986). *Statistics and data analysis in geology.* New York: Wiley.

Deere, D. (1988). *The rock quality designation (RQD) index in practice.* Rock classification systems for engineering purposes. ASTM International. https://doi.org/10.1520/STP48465S

Deng, G., Cao, K., Chen, R., Zhang, X., Yin, Q., & Zhou, H. (2018). A simple approach to evaluating leakage through thin impervious element of high embankment dams. *Environmental Earth Sciences,* 77(1), 1-11. https://doi.org/10.1007/s12665-017-7195-3

Deutsch, C. V., & Journel, A. G. (1992). *GSLIB: Geostatistical software library and user's guide.* Applied Geostatistics Series. New York, 119(147), 578 pp.

Dias, P. M., & Deutsch, C. V. (2022). *The decision of stationarity.* Geostatistics lessons. 7 pp.

Diggle, P. J., Tawn, J. A., & Moyeed, R. A. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society Series C: Applied Statistics,* 47(3), 299-350. https://doi.org/10.1111/1467-9876.00113

El Idrysy, E. H., & De Smedt, F. (2007). A comparative study of hydraulic conductivity estimations using geostatistics. *Hydrogeology Journal,* 15(3), 459-470. https://doi.org/10.1007/s10040-007-0166-0

Ewert, A. (2005). *Dam Engineering,* 5-65.

Ewert, F. K. (2012). *Rock grouting: with emphasis on dam sites*. Springer Science & Business Media.

Faybishenko, B., Witherspoon, P. A., & Benson, S. M. (2000). *Dynamics of fluids in fractured rock.* Washington DC American Geophysical Union Geophysical Monograph Series, 122 pp. DOI: 10.1029/GM122

Ford, D., & Williams, P. D. (2013). *Karst hydrogeology and geomorphology*. John Wiley & Sons.

Fransson, Å. (2004). *Development and verification of methods to estimate transmissivity distributions and orientation of conductive fractures/features along boreholes (No. SKB-R--04-59)*. Swedish Nuclear Fuel and Waste Management Co.

Gavinhos, V., & Carvalho, J. (2017). Geostatistical Modelling and Simulation Scenarios as Optimizing Tools for Curtain Grouting Design and Construction at a Dam Foundation. *Geostatistics Valencia 2016,* 789-804. https://doi.org/10.1007/978-3-319-46819-8_54

George, N. J., Ekanem, A. M., Ibanga, J. I., & Udosen, N. I. (2017). Hydrodynamic implications of aquifer quality index (AQI) and flow zone indicator (FZI) in groundwater abstraction: a case study of coastal hydro-lithofacies in South-eastern Nigeria. *Journal of Coastal Conservation,* 21, 759-776. https://doi.org/10.1007/s11852-017-0535-3

Gilg, B., & Gavard, M. (1957). *Calcul de la perméabilité par des essais d'eau dans les sondages en allluvions.* Edition de la Société du Bulletin technique de la Suisse romande.

Gong, G., Mattevada, S., & O'Bryant, S. E. (2014). Comparison of the accuracy of kriging and IDW interpolations in estimating groundwater arsenic concentrations in Texas. Environmental Research, 130, 59-69. https://doi.org/10.1016/j.envres.2013.12.005

Gnanadesikan, R., & Kettenring, J. R. (1972). Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics,* 81-124. https://doi.org/10.2307/2528963

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation.* Applied Geostatistics.

Guedes, L. P. C., Uribe-Opazo, M. A., Johann, J. A., & Souza, E. G. D. (2008). Anisotropia no estudo da variabilidade espacial de algumas variáveis químicas do solo. *Revista Brasileira de Ciência do Solo,* 32, 2217-2226. https://doi.org/10.1590/S0100-06832008000600001

Guedes, L. P. C., Uribe-Opazo, M. A., & Junior, P. J. R. (2013). Influence of incorporating geometric anisotropy on the construction of thematic maps of simulated data and chemical attributes of soil. *Chilean Journal of Agricultural Research,* 73(4), 414. https://doi.org/10.4067/S0718-583920130004 00013

Haneberg, W. C., Mozley, P. S., Moore, J. C., & Goodwin, L. B. (1999). *Faults and subsurface fluid flow in the shallow crust*. Washington DC American Geophysical Union Geophysical Monograph Series, 113. DOI:10.1029/GM113

Harding, B., & Deutsch, C. V. (2021). *Trend modeling and modeling with a trend.* Geostatistics Lessons.

Hekmatnejad, A., Emery, X., & Elmo, D. (2019). A geostatistical approach to estimating the parameters of a 3D Cox-Boolean discrete fracture network from 1D and 2D sampling observations. *International Journal of Rock Mechanics and Mining Sciences*, 113, 183-190. https://doi.org/10.1016/j.ijrmms.2018.11.003

Hvorslev, M. J. (1951). *Time lag and soil permeability in groundwater observations (No. 36).* Waterways Experiment Station, Corps of Engineers, US Army.

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. Oxford University Press, 561 pp.

Jalali, M., Karami, S., & Marj, A. F. (2016). Geostatistical evaluation of spatial variation related to groundwater quality database: case study for Arak plain aquifer, Iran. *Environmental Modeling and Assessment,* 21(6), 707-719. https://doi.org/10.1007/s10666-016-9506-6

Jalali, M., Karami, S., & Marj, A. F. (2019). On the problem of the spatial distribution delineation of the groundwater quality indicators via multivariate statistical and geostatistical approaches. *Environmental Monitoring and Assessment,* 191(2), 1-18. https://doi.org/10.1007/s10661-019-7432-1

Jolly, W. M., Graham, J. M., Michaelis, A., Nemani, R., & Running, S. W. (2005). A flexible, integrated system for generating meteorological surfaces derived from point sources across multiple geographic scales. *Journal of Environmental Modelling and Software,* 20(7), 873-882. https://doi.org/10.1016/j.envsoft.2004.05.003

Journel, A. G., & Huijbregts, C. J. (1976). Mining geostatistics.

Journel, A. G. (1989). *Fundamentals of geostatistics in five lessons.* American Geophysical Union, Washington DC.

Karami, S., Jalali, M., Karami, A., Katibeh, H., & Fatehi Marj, A. (2021). Evaluating and modeling the groundwater in Hamedan plain aquifer, Iran, using the linear geostatistical estimation, sequential Gaussian simulation, and turning band simulation approaches. *Modeling Earth Systems and Environment,* 1-22. https://doi.org/10.1007/s40808-021-01295-1

Kayabasi, A., Yesiloglu-Gultekin, N., & Gokceoglu, C. (2015). Use of nonlinear prediction tools to assess rock mass permeability using various discontinuity parameters. *Engineering Geology,* 185, 1-9. https://doi.org/10.1016/j.enggeo.2014.12.007

Kerry, R., & Oliver, M. (2007). Determining the effect of asymmetric data on the variogram. II. Outliers. Computers & Geosciences, 33(10), 1233-1260. https://doi.org/10.1016/j.cageo.2007.05.009

Khalili Shayan, H., & Amiri-Tokaldany, E. (2015). Effects of blanket, drains, and cutoff wall on reducing uplift pressure, seepage, and exit gradient under hydraulic structures. *International Journal of Civil Engineering,* 13(4), 486-500. DOI: 0.22068/IJCE.13.4.486

Koike, K., Liu, C., & Sanga, T. (2012). Incorporation of fracture directions into 3D geostatistical methods for a rock fracture system. *Environmental Earth Sciences,* 66(5), 1403-1414. https://doi.org/10.1007/s12665-011-1350-z

Koike, K., Kubo, T., Liu, C., Masoud, A., Amano, K., Kurihara, A., Matsuoka, T., & Lanyon, B. (2015). 3D geostatistical modeling of fracture system in a granitic massif to characterize hydraulic properties and fracture distribution. Tectonophysics, 660, 1-16. https://doi.org/10.1016/j.tecto.2015.06.008

Kresic, N. (2006). *Hydrogeology and groundwater modeling*. CRC press.

Krige, D. G. (1996). A practical analysis of the effects of spatial structure and of data available and accessed, on conditional biases in ordinary kriging. *Journal of Geostatistics Wollongong,* 96, 799-810.

Lark, R. M. (2000). A comparison of some robust estimators of the variogram for use in soil survey. *European Journal of Soil Science,* 51(1), 137-157. https://doi.org/10.1046/j.1365-2389.2000.00280.x<

Lei, Q., Latham, J. P., Tsang, C. F., Xiang, J., & Lang, P. (2015). A new approach to upscaling fracture network models while preserving geostatistical and geomechanical characteristics. *Journal of Geophysical Research: Solid Earth,* 120(7), 4784-4807. https://doi.org/10.1002/2014JB011736

Lin, Y. P., Lee, C. C., & Tan, Y. C. (2000). Geostatistical approach for identification of transmissivity structure at Dulliu area in Taiwan. *Environmental Geology,* 40(1), 111-120. https://doi.org/10.1007/s002540000150

Lin, Y. P., Tan, Y. C., & Rouhani, S. (2001). Identifying spatial characteristics of transmissivity using simulated annealing and kriging methods. *Environmental Geology,* 41(1-2), 200-208. https://doi.org/10.1007/s002540100383

Long, J. C., & Billaux, D. M. (1987). From field data to fracture network modeling: an example incorporating spatial structure. *Water Resources Research,* 23(7), 1201-1216. https://doi.org/10.1029/WR023i007p01201

Manna, M. C., Bhattacharya, A. K., Choudhury, S., & Maji, S. C. (2003). Groundwater flow beneath a sheetpile analyzed using six-noded triangular finite elements. *Journal of the Institution of Engineers, India. Civil Engineering Division,* 84(AOU), 121-129.

Marinoni, O. (2003). Improving geological models using a combined ordinary–indicator Kriging approach. *Journal of Engineering Geology,* 69(1-2), 37-45. https://doi.org/10.1016/S0013-7952(02)00246-6

Matheron, G. (1971). The theory of regionalized variables and its applications. *Les Cahiers du Centre de Morphologie Mathématique*, 5, 211 pp.

Milanović, P. T. (1981). *Karst hydrogeology.* Water Resources Publications.

Milanovic, P. (2004). *Water resources engineering in karst.* CRC press.

Mohammadi, Z., Raeisi, E., & Bakalowicz, M. (2007). Method of leakage study at the karst dam site. A case study: Khersan 3 Dam, Iran. *Environmental Geology,* 52(6), 1053-1065. https://doi.org/10.1007/s00254-006-0545-1

Moye, D. G. (1967). *Diamond drilling for foundation exploration.* Inst Engrs Civil Eng Trans, Australia.

Novak, P., Moffat, A. I. B., Nalluri, C., & Narayanan, R. (2017). *Hydraulic structures.* CRC Press.

Öztürk, C., & Nasuf, E. (2002). Geostatistical assessment of rock zones for tunneling. Tunnelling and Underground Space Technology, 17(3), 275-285. https://doi.org/10.1016/S0886-7798(02)00023-8

Palmstrom, A. (2005). Measurements of and correlations between block size and rock quality designation (RQD). *Tunnelling and Underground Space Technology,* 20(4), 362-377. https://doi.org/10.1016/j.tust.2005.01.005

Philip, R. D., & Kitanidis, P. K. (1989). Geostatistical estimation of hydraulic head gradients. Groundwater, 27(6), 855-865. https://doi.org/10.1111/j.1745-6584.1989.tb01049.x

Ramazi, H., & Jalali, M. (2015). Contribution of geophysical inversion theory and geostatistical simulation to determine geoelectrical anomalies. *Studia Geophysica et Geodaetica,* 59(1), 97-112. https://doi.org/10.1007/s11200-013-0772-3

Razack, M., & Lasm, T. (2006). Geostatistical estimation of the transmissivity in a highly fractured metamorphic and crystalline aquifer (Man-Danane Region, Western Ivory Coast). Journal of Hydrology, 325(1-4), 164-178. https://doi.org/10.1016/j.jhydrol.2005.10.014

Richter, W., & Lillich, W. (1975). Abriß der Hydrogeologie.

Rossi, M. E., & Deutsch, C. V. (2014). *Mineral Resource Estimation.* Springer Science & Business Media.

Santos, E., Gopinath, T., & Lima, A. (2018). Geostatistical analysis and interpretation of the geotechnical properties of rock massif, Ceara State, Brazil. *Mine Planning and Equipment Selection 2000*, 227-231.

Shimizu, S., Jojima, S. & Niida, Y. (1985). Design and execution of foundation grouting for multipurpose dams in Japan.

Smith, M., & Konrad, J. M. (2011). Assessing hydraulic conductivities of a compacted dam core using geostatistical analysis of construction control data. *Canadian Geotechnical Journal, 48*(9), 1314-1327. https://doi.org/10.1139/t11-038

Sterk, G., & Stein, A. (1997). Mapping wind-blown mass transport by modeling variability in space and time. *Soil Science Society of America Journal, 61*(1), 232-239. https://doi.org/10.2136/sssaj1997.03615995006100010032x

Snowden, D. V. (2001). *Practical interpretation of mineral resource and ore reserve classification guidelines*. Mineral Resource and Ore Reserve Estimation-The AusIMM Guide to Good Practice, 643-652.

Tartakovsky, D. M. (2013). Assessment and management of risk in subsurface hydrology: A review and perspective. *Advances in Water Resources,* 51, 247-260. https://doi.org/10.1016/j.advwatres.2012.04.007

Thomas, H. H. (1978). *The engineering of large dams.* [Dissertation University of Tasmania, Australia.]

Tukey, J. W. (1977). Exploratory data analysis, 131-160.

Turkmen, S., Özgüler, E., Taga, H., & Karaogullarindan, T. (2002). Seepage problems in the karstic limestone foundation of the Kalecik Dam (south Turkey). *Engineering Geology,* 63(3-4), 247-257. https://doi.org/10.1016/S0013-7952(01)00085-0

Uromeihy, A. (2000). The Lar Dam; an example of infrastructural development in a geologically active karstic region. *Journal of Asian Earth Sciences,* 18(1), 25-31. https://doi.org/10.1016/S1367-9120(99)00026-7

Vieira, S. R., Carvalho, J. R. P. D., Ceddia, M. B., & González, A. P. (2010). Detrending nonstationary data for geostatistical applications. *Bragantia,* 69, 01-08. https://doi.org/10.1590/S0006-87052010000500002

Viruete, J. E., Carbonell, R., Jurado, M. J., Martı, D., & Pérez-Estaún, A. (2001). Two-dimensional geostatistical modeling and prediction of the fracture system in the Albala Granitic Pluton, SW Iberian Massif, Spain. *Journal of Structural Geology,* 23(12), 2011-2023. https://doi.org/10.1016/S0191-8141(01)00026-8

Webster, R., & Oliver, M. A. (2007). *Geostatistics for environmental scientists*. John Wiley & Sons.

Zimmerman, D. L. (1993). Another look at anisotropy in geostatistics. *Mathematical Geology,* 25, 453-470. https://doi.org/10.1007/BF00894779