Multiple-Point Statistical Simulation of rock fracture network as a key control on the hydrogeology and salinity: a case study from the Qarabagh area, West Azarbayjan Province, Iran

M. Jahan Mohammadi\textsuperscript{a}, E. Ranjineh Khojasteh\textsuperscript{a}*, M. Faridazad\textsuperscript{a}

\textsuperscript{a} Faculty of Mining Engineering, Sahand University of Technology, Tabriz, Iran

Received: 5 December 2019; Received in revised form: 2 March 2020; Accepted: 16 March 2020

Abstract

Modeling and characterization of the geometry and distribution of Rock Fracture Networks (RFNs) are essential in applications such as hydrogeological or environmental evaluations. It is widely accepted that RFNs are potentially associated with the hydrogeological (thus salinity) characteristics of the surrounding environments. Despite the complexity and inaccessibility of RFNs, stochastic methods provide a functional framework to predict their characteristics in the subsurface. An efficient tool for modeling RFNs is the Discrete Fracture Network (DFN) which also includes a number of geostatistical techniques that consider spatial variability structure. The advantages of these techniques are: realistic results, ease of application, and uncertainty assessments. Multiple-point geostatistics/statistics (MPS) is a modern and effective geostatistical tool for realistically simulating RFNs. In the present study, we modeled the RFNs in a location near the Qarabagh area, in the western Urmia Lake; in this regard, we used the Single Normal Equation Simulation (SNESIM) algorithm of the MPS geostatistical method using Training Images (TIs) instead of variograms. The required datasets and information for this modeling was provided using the field measurements of the fracture orientations and dips, as well as the outcrop photographs. The outcomes of these models can be used in predicting the salinity distribution in the surrounding area. Therefore, through the SNESIM algorithm, TIs obtained from the outcrop photographs, and direct measurements, 100 RFN realizations were generated at each station. These realizations were then averaged to predict the locations with higher and lower fracture probabilities and to assess the general trend of the fracture distributions.

Keywords: Discontinuities modeling; Multi-point statistics/geostatistics; Training images; Salinity; Qarabagh area; Urmia Lake

1. Introduction

Investigation of Rock Fracture Networks (RFNs) is an essential step in environmental and hydrogeological issues. Moreover, several studies over various environments have reported the significant effects of RFNs on salinity distribution models. For instance, Loefman (1997) constructed numerical flow and transport models for the Äspö research site in Sweden. Their research was based on the geochemical field data using a coupled groundwater-salt transport model which also considers RFNs as an effective factor. A similar study was reported in Skagius (2010) about a Swedish repository site. Karvonen (2013) investigated surface hydrological characterization together with the salt transportation modeling in the Olkiluoto nuclear power plant site in western Finland; they considered RFNs as a key element in this model of transport. RFNs have roles in similar problems related to the distribution of the salinity in various environments; therefore, their modeling and prediction could be an effective approach to evaluating saline environments such as Urmia Lake. Unfortunately, the gradual death of Urmia Lake over the recent decades, increased salinity, and growth of aerosol distribution risk have noticeably endangered the neighboring ecosystems and human life; as a result, the area of the present study is currently adjacent to the salt sheets instead of the water.
On the other hand, joints and fracture systems in rock masses have shown a large degree of complexity in their geometry and other characteristics (such as filling materials). However, the amount and quality of information and parameters obtained from surveying such areas are usually noticeably limited. Even in drilling some boreholes, the obtained images are restricted by their diameter and depth, generating a large degree of uncertainty. In order to reduce such uncertainties, the best solution is to model these discontinuities (Jing and Stephansson, 2007; Zeeb et al., 2013).

To simulate discontinuities and fractures in rocks, several techniques have been developed; these methods can geometrically and geotechnically be divided into three categories: deterministic, stochastic (statistical), and geostatistical. In the first method, the geometry of rock mass is defined only by a limited number of main discontinuities (Tatomir, 2012). In the second approach, the geometric parameters of the joint sets are generated and simulated using their proven distribution functions. A Boolean (or object-based) method can be utilized here to construct the RFN geometries (Dowd et al., 2007; Xu et al., 2010). Lepillier et al. (2020) presented a workflow to incorporate the outcrop measurements from linear scanline surveys into the RFN modeling system. Li et al. (2020) also proposed a new approach for the two-dimensional prediction of RFNs and evaluated their connectivity and effects on rock permeability. The third technique uses geostatistical tools for modeling and predicting RFNs. Basically, in these methods, the parameters defining joint sets are estimated by geostatistical methods. These parameters have a spatial dependency or autocorrelation such as dip, dip direction, dimensions, and joint openings. Traditional geostatistical prediction of RFNs are usually only in few steps of Poisson RFNs modeling (Dowd et al., 2007). For example, Billaux et al. (1987) performed the geostatistical modeling of RFNs; they used a fracture density variogram to model the fracture centers in a two-dimensional space (Koike et al., 2012). More advanced applications of geostatistical tools for modeling RFNs were introduced and discussed using simulated annealing in the framework of geostatistical criteria to more completely account for the spatial continuity models (Assteerawatt, 2007; Tatomin, 2012).

Zeeb et al. (2013) also evaluated the sampling methods for characterizing RFNs from the outcrop measurements; they suggested an approximate number of 225 measurements to acquire a sufficient accuracy in stochastic or other RFN models. Further similar attempts were reported in relevant references (Tóth 2018 and Morgan, 2019).

The multiple-point statistics/geostatistics (MPS) was further employed in the geostatistical prediction of the RFNs (Dowd et al., 2007). This concept was first proposed by Guardiano and Srivastava (1993). This method accounts for the correlations between three or more locations/points at a time. Hence, it is capable of reproducing the connectivity of many locations and complex curvilinear geological structures. The main components of MPS are ‘Training Images (TIs), algorithms, and hard and soft data (Arpat 2005). Boucher (2008 and 2009) implemented super-resolution mapping with MPS. Ge and Bai (2011) extracted linear objects from remotely sensed imagery using MPS. The use of MPS in gap filling applications of remote sensing (Mariethoz et al., 2012) and facies modeling considering more complex geometries were also reported (Strebelle and Journel, 2001). Bruna et al. (2019) used outcrop photographs to generate the TIs of the MPS method in order to characterize the subsurface reservoir rocks.

In this field, several MPS algorithms have been developed, each with their own specific strengths and weaknesses. SNESIM, SIMPAT, IMPALA, and FILTERSIM are among the important MPS algorithms. In this study, SNESIM algorithm (Strebelle and Journel, 2000) was used to simulate RFNs within the study area in a two-dimensional space.

In the study area (located in Western Azerbaijan Province, Iran), discontinuities were surveyed, and their unknown parts were predicted using the SNESIM algorithm of the MPS method. The MPS algorithms better reproduce the RFN geometries, especially their connectivity, which is a determinative factor in transport modeling.

2. Materials and Methods

2.1. Study area and field measurements

As mentioned above, the study area is situated northwest of Urmia Lake (55 km to the north of Urmia) close to the Qarabagh village of the Salmas County, Western Azarbayjan Province, Iran.

To investigate the exposed RFNs, we selected two existing road trenches where the fractures were outcropped. Figure 1 depicts the location map of the two mentioned road trenches and a satellite image of the study area.

In trenches 1 and 2, orientations (azimuths), dip angles, trace lengths, and apertures (openings) were measured for 272 and 224 joints and
fractures, respectively (Figure 2). Table 1 summarizes the recordings resulting from the measured fracture characteristics along a scanline. We used the scanline method to record the existing fractures and joints in the study area. In this method, one or more lines aligned towards appropriate directions are considered using the tools such as tape measures. The places where fractures cross the tape measure are then recorded along with their directions, lengths, openings, and factors of that kind. For scanline measurement purposes of the fractures in this study, we used a metallic tape measure which was moved to different places in the recording outcrop.

The extracted information concerning these fracture directions at each station were later used to categorize the recorded fractures. Moreover, several almost planar photographs (from the top) were taken at each measurement station; these images were later used in generating Training Images (TIs) and checking the recorded fractures and their connectivity. In creating TIs, the field measurement was employed together with the taken photographs; this was done to complement and probably correct the recordings from either of them.

2.2. Geostatistical simulations using the Multiple-Point Statistics (MPS) algorithm

Established and formulated by Matheron (1963), geostatistics relies on the regionalized variables concept. This concept requires the existence of a spatial continuity (or variability) structure. These spatial continuity (or variability) structures can be calculated and modeled using several geostatistical tools such as covariograms (or variograms); these tools can be later used to predict the variables of interest at unknown (unmeasured) locations with a limited number of measured points. As a most popular geostatistical estimator, kriging is known as the Best Linear Unbiased Estimator (B.L.U.E.) because it makes unbiased predictions (without a systematic error) using a linear combination of the known values and has minimum estimation variance.

Fig. 1. Location of measurement sites (trenches 1 and 2) on the satellite image of the study area
Fig. 2. The location map of the fracture recording stations in the two measurement trenches: trench 1 (the top figure) and trench 2 (the bottom figure).

Table 1. The measured characteristics of RFNs at a recording station

<table>
<thead>
<tr>
<th>Name</th>
<th>Strike</th>
<th>Dip</th>
<th>Dip direction</th>
<th>Spacing (cm)</th>
<th>Opening (mm)</th>
<th>Fracture trace length (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JS1</td>
<td>35</td>
<td>56</td>
<td>305</td>
<td>44</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>JS2</td>
<td>60</td>
<td>65</td>
<td>330</td>
<td>55</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td>JS3</td>
<td>54</td>
<td>80</td>
<td>324</td>
<td>76</td>
<td>1</td>
<td>50</td>
</tr>
<tr>
<td>JS4</td>
<td>100</td>
<td>73</td>
<td>10</td>
<td>79</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>JS5</td>
<td>105</td>
<td>68</td>
<td>15</td>
<td>110</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td>JS6</td>
<td>63</td>
<td>50</td>
<td>333</td>
<td>149</td>
<td>3</td>
<td>70</td>
</tr>
<tr>
<td>JS7</td>
<td>37</td>
<td>78</td>
<td>307</td>
<td>174</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td>JS8</td>
<td>45</td>
<td>63</td>
<td>315</td>
<td>220</td>
<td>5</td>
<td>500</td>
</tr>
<tr>
<td>JS9</td>
<td>60</td>
<td>61</td>
<td>330</td>
<td>240</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>JS10</td>
<td>110</td>
<td>75</td>
<td>20</td>
<td>265</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>JS11</td>
<td>95</td>
<td>56</td>
<td>5</td>
<td>274</td>
<td>0</td>
<td>71</td>
</tr>
<tr>
<td>JS12</td>
<td>55</td>
<td>51</td>
<td>325</td>
<td>371</td>
<td>0</td>
<td>105</td>
</tr>
<tr>
<td>JS13</td>
<td>40</td>
<td>80</td>
<td>310</td>
<td>398</td>
<td>0</td>
<td>62</td>
</tr>
<tr>
<td>JS14</td>
<td>115</td>
<td>55</td>
<td>25</td>
<td>520</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td>JS15</td>
<td>98</td>
<td>62</td>
<td>8</td>
<td>541</td>
<td>0</td>
<td>110</td>
</tr>
<tr>
<td>JS16</td>
<td>63</td>
<td>50</td>
<td>333</td>
<td>567</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>JS17</td>
<td>38</td>
<td>62</td>
<td>308</td>
<td>570</td>
<td>3</td>
<td>130</td>
</tr>
<tr>
<td>JS18</td>
<td>60</td>
<td>70</td>
<td>330</td>
<td>585</td>
<td>0</td>
<td>123</td>
</tr>
<tr>
<td>JS19</td>
<td>110</td>
<td>64</td>
<td>20</td>
<td>605</td>
<td>13</td>
<td>100</td>
</tr>
<tr>
<td>JS20</td>
<td>118</td>
<td>53</td>
<td>28</td>
<td>680</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td>JS21</td>
<td>35</td>
<td>82</td>
<td>305</td>
<td>745</td>
<td>1</td>
<td>84</td>
</tr>
</tbody>
</table>
Geostatistical methods also provide estimations with their estimation variances, enabling users to evaluate the uncertainties involved in the estimation procedure. These advantages make geostatistical tools one of the most efficient and reliable methods for predicting the unknown values in a wide range of applications.

However, the use of conventional spatial continuity (or variability) measures like covariograms (or variograms) offers certain shortcomings. The problem originates from the fact that all the above-mentioned spatial continuity (or variability) tools use a two-point framework in calculating and modeling these structures. Nevertheless, these two-point statistics are unable to sufficiently capture the variability nature of the complex geological features such as curvilinear structures or RFNs (Caers and Zhang, 2004).

In this connection, Multiple-Point Statistics/Geostatistics (MPS) are applied to more complicated features such as fracture traces, as mentioned earlier. It also facilitates the integration of subjective information and soft data in the prediction models.

In the study area, MPS was implemented to predict the RFNs using the SNESIM algorithm (Strébelle and Journel, 2000) and field measurements.

The MPS algorithm essentially predicts the conditional local mass density functions of the existing categories in each simulation location using a stored data-base which can be expressed as a multiple-point histogram. This histogram is created by scanning a TI (believed to resemble the target reality) within a predefined template containing three or usually more cells/points and storing the observed patterns within the stated template. Accordingly, the frequency of the patterns observed in the TI is calculated and stored within this database. In the next step, a simulation grid is defined to perform the simulation. Afterwards, the algorithm moves through a random path and calculates the mass density functions at each simulation grid-node, conditioned to the observed measurements.

To do so, the first step is to create the required TIs.

2.3. Training images (TIs)

TIs are databases related to geological patterns that resemble the reality; thus, they reflect the variability structure of the system under study. Therefore, if the plausible geological patterns of the model are inferred and drawn, the appropriate TIs can be constructed. TIs can be achieved from different sources, including interpreted photographs of outcrops, a sketch drawn and properly digitized by an expert, an unconditional simulation using object-based methods, or the result of a process-imitating simulation. A TI merely conveys the geometrical characteristics of the target reality or the complex spatial relationships among multiple subsurface features. Hence, the TI is not conditioned to any data point and does not hold any information depending on the location (Honarkhah and Caers, 2010).

In order to produce the required TIs, the photographs taken at each station were used. A 30 cm arrow aligned towards the north was placed on each photographing site to facilitate the recognition of the geographical directions and the length scales on the photographs. Moreover, the recorded fracture information in these stations were implemented as ancillary information to correct and accomplish the photograph data. An important feature considered at this stage was the connectivity of the fractures that should be correctly reflected in the TIs. Next, the deduced TIs were drawn based on their inferred azimuths and fracture trace lengths.

Moreover, the average opening of the available joints in the region was considered in the drawn fracture line thickness; the whole TIs were saved in jpg format. Figure 3 is an example of the steps of sketching such a TI. After that, using a MATLAB code, the TIs were converted into a readable format in SGeMS software.

2.4. The SNESIM algorithm

Strébelle and Journel (2000) developed a MPS algorithm called Single Normal Equation Simulation (SNESIM) which later became one of the most popular MPS procedures in a wide range of applications. The name of the algorithm emphasizes the use of only a single normal equation while modeling the probability of a category at a specific simulation grid node. The SNESIM algorithm steps can be simply expressed as below:
1. Define a multiple-point template $T_j$ to scan the TI.
2. Scan the TI using the multiple-point template $T_j$ and store the observed proportional frequencies of patterns in a database called search tree.
3. Assign the existing conditioning data to their closest simulation grid-node.
4. Define a random path visiting all the simulation grid-nodes only once.
5. At each location $u$:
   a. Determine the present conditioning data event $\text{cond}_i(u)$ inside the template $T_j$.
   b. Calculate the conditional probability distribution function (cdf: $\text{prob}(I(u) = k | \text{cond}_i(u))$
based on the proportions of the patterns stored in the search tree.

c. Randomly draw a value from this cdf and assign the simulated value to the location $u$. Consider the simulated value as conditioning data in the following steps.

6. Repeat step 5 until all nodes in the grid are simulated (Madriz, 2009; Remy et al., 2009).

The search trees in SNESIM are designed to quickly calculate the probability distribution functions, facilitate the access of the algorithm to the patterns, and reduce the calculation time during the simulation procedure.

![Image](image_url)

Fig. 3. The steps (a) to (c) for creating a TI by connecting the fracture trace lines to each other on a taken photograph, represented for one of the measurement stations.

2.5. **Hard data**

The term “hard data” refers to a data value that should exactly be reproduced by the modeling method at its measured location (Honarkhah and Caers, 2010). Therefore, hard data conditioning is beyond a simple insertion of the data into the model and freezing it (Arpat, 2005).

The hard data is highly necessary in simulation. It is made based on the properties recorded in the field, namely the location, orientation, and trace (Fig. 4). To prepare the hard data-sets based on the recorded location of fractures on the scanline, azimuth, and trace of fractures, they were drawn in AutoCAD Map3D. They were drawn with the desired thickness after calling them to the ArcGIS software. For the ease of running a simulation, we considered a thickness equal to the average opening of the fractures.

After this step, the images were converted into a readable format in MATLAB using the existing codes on SGeMS (Hansen, 2004).

3. **Results and Discussion**

3.1. **Overview**

As explained in Section 2, the TIs of RFNs were primarily created at each measurement station; their conditioning (hard) data were then determined by drawing the field measured location, trace line lengths, and directions through their scanlines. Next, the MPS simulations were generated using SNEMSIM algorithm. Finally, the average map of the RFNs in the vicinity of each station was calculated and illustrated. The following sections explain the details of these outcomes.

3.2. **Creating TIs**

As previously mentioned, creating appropriate TIs is an important step in predicting RFNs within a geostatistical framework of the MPS method.
The details of creating TIs were delineated in the previous section (Section 2-4).

The TIs were drawn for each recording station using the photographs combined with the direct measurements. Figure 4 illustrates a TI created through the explained routine. It is worth mentioning that the average fracture opening value inferred from the field measurements here was 2.5 mm; this value was considered as the average thickness of the fracture lines in the TIs. These TIs were then used as an input (together with the hard data) of SNESIM program to generate the required simulations.

3.3. Sketching the hard data

The fractures recorded at each station, with specific locations and orientations, were drawn according to the direct field measurements along the scanlines; they were inserted as conditioning (hard) data into the SNESIM algorithm before conducting the simulations.

Figure 5 shows the sketch of the recorded RFNs as hard data at station 5.

These images were digitized and fed into the SNESIM program of the SGeMS software (Remy et al., 2009) along with the TIs to generate the required simulations.

3.4. MPS simulation results using the SNESIM algorithm

The TIs, hard (conditioning) data from the two previous stages, and other simulation parameters (such as the extent of simulation domains, search template dimensions, and number of realizations to be generated) were fed into the SNESIM simulation code to create 100 realizations at each recording station. Each of these realizations indicated a plausible scenario of what could occur in reality in each station or the areas nearby.

The simulation domains were considered as a 200cmx200cm block to cover enough area for representing the characteristics of RFNs.

Figure 6 represents two simulation results (two realizations) for the fifth recording station. The simulation was repeated to generate 100 realizations, and the average of simulations at each location was finally calculated (Figure 7). As stated in the Introduction, the results of these realizations can be utilized as input for a fluid flow simulator or for the predictions of the salinity distributions. As expected and shown by the figures, the simulations based on the MPS algorithms reproduce the connectivity of RFNs very well. This feature is an important advantage of the MPS techniques, particularly compared to other RFNs modeling methods. The general trend associated with the distributions of RFNs could also be assessed through the averaged (E-type) map of the multiple realizations.
5. Conclusion

The purpose of this study was to evaluate the potential of the MPS method in making reliable predictions of the RFNs in an area in the northwestern coast of Urmia Lake. This kind of prediction can be useful in providing the required characteristics related to the transport modeling, especially for hydrogeological and salinity distribution purposes.

Hence, the RFN characteristics were measured in the study area, and several almost planar photographs were taken at each recording station. The direct measurements later served as conditioning data for MPS simulations. Furthermore, the obtained photographs and the direct measurements were used in creating the TIs. Generating TIs is a highly important and sensitive part of MPS simulations. Therefore, all the information and geological data had to be used to create the realistic TIs. Having TIs and conditioning data, the simulations were performed through the SNESIM algorithm. The realizations were represented, and their averages were calculated at each simulation site.

It is obvious that the MPS outcomes displayed a very realistic connection between the fractures and provided a reliable model of RFNs in the study area. The dominant number of field observations represented a general trend in the NE-SW direction. The high density of fractures could be due to the existence of large and small faults near the study area.

The results of this study suggested the applicability of MPS algorithms in predicting the RFNs using field observations from direct measurements and photography in similar cases.

Additionally, the following suggestions could be made for a follow-up research in the study area:

1) Using Markov chain Monte Carlo method for RFNs modeling based on borehole data in the study area provided the drillings are possible (for detailed studies) for comparing their efficiencies with the MPS methods.

2) Implementing and developing photogrammetric methods to provide more detailed while quick and accurate information for constructing even more accurate RFN models.
3) Using the simulation results in flow and transport models to evaluate the environmental risks existing in the study area or similar cases.

Acknowledgements

The authors would like to thank Dr. Vahed Aghazadeh, the former financial/official deputy of the Sahand University of Technology, for partially covering the field trip logistic costs during this research. We are also grateful to Ahmad Barzegar and Behzad Esmaeili, the graduates of the Sahand University of Technology, for assisting us in the field measurements. We would heartily appreciate Mr. Ali Khezri, from the city of Salmas, who devotedly supported the group in providing various requirements and facilities during the field measurements.

References


Modeling Dept., University of Stuttgart, Stuttgart, Germany.