



An efficient karst fracture seepage path construction algorithm based on a generalized disk model

Xiaoping Rui^{1,2} · Tian Yu²

Received: 3 November 2018 / Accepted: 8 March 2019 / Published online: 4 April 2019
© Saudi Society for Geosciences 2019

Abstract

The identification and reconstruction of karst fracture networks have always been an important and difficult research topic in the field of karst groundwater resources and geological environment protection. Based on the original surface fracture data obtained in Zhangfang, Fangshan district in southwest of Beijing, and the surface fracture records processed through the Kriging interpolation method, we propose a method for the construction of a flow path based on a disk model. We focus on identifying the flow path in a three-dimensional fracture structure and underground fracture data obtained by the Monte Carlo method. The flow path along the fracture generalized by the disk model is simulated using a directed graph data structure and is stored in an adjacency matrix. To meet the demands of fast intersection operation with large-scale fracture data and to improve the operating efficiency of a karst fracture simulation system, this paper first divides the study space into smaller blocks and then uses a three-dimensional R-tree index algorithm to reduce the time required to traverse each sampling point. Finally, based on the sampled surface fracture data from Zhangfang, this paper studies the distribution model for a three-dimensional fracture space with the assistance of a remote sensing geological survey, fine geological survey of key karst areas, and sampling analysis. The algorithm for building the path of karst fractures was also simulated. It provides a visualization analysis approach to study the karst development mechanism and numerical simulation of karst water fractures in Zhangfang.

Keywords Karst area · Fracture · Disk model · Seepage path

Introduction

Karst fractures, the groundwater conduits, play an important role in groundwater flow. Therefore, the study on the spatial distribution of fractures is essential to understand groundwater migration (Berkowitz 2002). The karst fracture network model constructed by the integration of both surface fracture data and underground fracture data is the basis for simulating the migration of fissure water and pollutants in the fracture media. The discontinuity, inhomogeneity, and anisotropy in the development of karst fractures result in greatly different

conditions of groundwater flow, which poses huge challenges to research groundwater seepage along fractures (Batu 2009; Brkic et al. 2018; Ford and Williams 2007). This paper aims at simulating crack seepage in karst areas and establishing a three-dimensional fracture network model to provide a solid theoretical basis for computer simulations of groundwater flow along with karst fractures.

Establishing an underground fracture seepage network model in a karst area is key to understanding the intensity and development mechanism of karst fractures (Karimzade et al. 2017), as well as underground transport of pollutants (Gellasch et al. 2013). Moreover, establishing a model can help decision-makers identify the development and distribution of fractures within the study area. The dispersion rate and diffusion edge, contaminant plume, and effects of groundwater flow on environment can be calculated and estimated by combining simulations with external factors like hydrogeological conditions in study field (He 1997; Chen 2011).

In the process of constructing large-scale fracture networks by computer, fractures characterized by connectivity which are important components of the seepage path must be found

Editorial handling: Chen Mingjie

✉ Xiaoping Rui
ruixpsz@163.com

¹ School of Earth Sciences and Engineering, Hohai University, Nanjing 211000, China

² College of Resource and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

before determining seepage paths (Hyman et al. 2016; Medici et al. 2016). Huamei Liu provided the theoretical basis for fracture intersection determination using analytic geometry in which plane equations for two fractures were established. According to the knowledge on analytic geometry, if the two disks intersect, the sum of the distance between their centers to the intersecting line must be less than the sum of their radii, and their distances from the center to the intersection must be less than each radius; this can be used to determine whether two crevices intersect (Liu and Wang 2010). However, some fractures are isolated, meaning that they do not intersect with any other fractures. Further judgment is required every time these fractures are involved in the evaluation, which costs more time in calculation and reduces efficiency. To solve this problem, an R-tree index algorithm is used to exclude the isolated fractures and evaluate whether they truly intersect. Meanwhile, the consumption of computer memory begins to increase in the process of constructing an R-tree index, but it remarkably saves computing time. Zhiyu Li proposed and optimized an algorithm termed “bounding box” for rapid judgments of intersections, which also improved the efficiency of the system. The final judgment results were stored in an adjacency matrix or adjacency list and illustrated in a form of directed graphs, thus paving the way for groundwater seepage simulations in a karst area (Li et al. 2014).

To demonstrate data graphically and intuitively, three-dimensional visualization technology seen as a bridge to link multiple disciplines is a useful tool for analyzing, understanding, and duplicating data. A large amount of data can be used by it to check the continuity and authenticity of the material and find anomalies (Wang 2016). Different concrete or abstract fracture network models can be vividly presented in real time using 3D visualization technology to simulate a three-dimensional fracture network. Eliminating abnormal values and analyzing meaningful abnormal values allow one to better understand their results.

Generalized disk model for karst fractures

To intuitively present the karst fracture networks, the in situ data about fractures such as the coordinates, length, trace length, and separation distance, as well as the occurrence including dips and dip angles obtained from field measurements must be displayed in a three-dimensional form in visualization. Based on the acquired data mentioned above, fractures demonstrated on a spatial level by an appropriate and easily simulated model chosen for the study area provide a prerequisite for recognition of the flow path along the crevices.

The fracture surface morphology must be taken into account when it comes to the three-dimensional simulation of joint fracture, three-dimensional stability analysis of rock mass mechanics, underground fluid flow field,

and material migration path characteristics, to name a few. In the field, the shape of the fracture plane varies and it is usually difficult to directly observe the specific geometry of an intact fracture surface. Therefore, computer simulation of the fracture data is required; otherwise, it is almost impossible to completely display a fracture surface through computer visualizations. First of all, it is necessary to simplify the structural plane to a regular shape such as a circle, an ellipse, or a polygon. Then, the relatively accurate structural parameters could be obtained through the method of probability statistics and spatial analysis.

The studies on fracture models began in the 1960s. A deterministic orthogonal model was first proposed, but this model is imprecise due to the limited number of random parameters and the assumption that the fracture is unbounded (Snow 1965). Bounded fracture models were later proposed, for instance, the three-dimensional seepage model of a disk fracture network by Long (Long et al. 1985), polygonal fracture network model by Dershowitz, Poisson plane model, and Mosaic board model by Einstein (Dershowitz and Einstein 1988). At present, researchers have been attaching importance to the uncertainty of the fracture surface and using probability modeling to simulate the fracture properties, such as its shape, length, and trace length, to name a few. As a result, the Monte Carlo method is widely used in karst fracture models gradually, bringing the method from pure mathematical analysis to practical application.

Studies show that fractures have a tendency to form a circular shape, and the real shape of a fracture is probably a rough oval (Baecher 1977). There is also evidence showing that the space length of a fracture along the strike direction is basically similar to that along dip direction (Robertson 1970). The disk model has its own unique advantages regarding construction of the fracture network in the study area (Zhangfang, Fangshan district in Beijing). Firstly, the fractures in this area are close to a disk shape. Secondly, the disk model can clearly display the fracture parameters. In other words, the center of the disk represents the location of the fracture in space and the radius of the disk indicates the size of the fracture. The length of the tangent line between the disk and the surface is the trace length, and the occurrence of the fracture can also be represented by the underground distribution of disks. Finally, the disk model is easy to simulate with a computer, thus reducing complexity.

Figure 1 shows the structural planes of fracture represented by disks, where the center of the disk is the location of the fracture; the radius of the disk is equal to that of the fracture; and the distribution direction in the underground space represents the occurrence.

At present, the disk fracture network model is most widely used to study karst fractures. Although fractures generally diverge from a circular shape when they expand to the boundary (e.g., ground surface, lithologic interface, etc.), this does

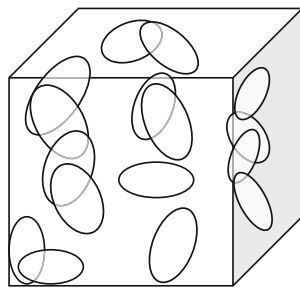


Fig. 1 Disk model

not have a serious impact on the application of the disk model because the model puts emphasis on the entire fracture, ignoring the local deformation.

Many assumptions must be made in advance for a computer simulation due to the limitations to computer operation and display, measuring techniques and conditions, and differences in the development of the fracture structure in space. According to previous research results on karst fractures in Zhangfang, Fangshan district in Beijing (Dong et al. 2014; Qiao et al. 2014), the disk model was used to carry out a generalized simulation of the structural plane of fractures based on the merits of the disk model compared to the other fracture models. That is to say, fractures in the rock mass are assumed to be thin disk-shaped, and its growth level and spatial distribution are represented by the disk radius and occurrence.

Each fracture in the disk model is regarded as a thin space disk. The required data to three-dimensionally visualize fractures is stored in a database, which mainly include the center coordinates $O(x, y, z)$ of the fracture disk, disk occurrence (α, β ; where α means direction and β dip angle), disk radius R , and fracture density ρ . These parameters are shown in Fig. 2. Dip and dip angle are the primary elements of the occurrence of fractures. The dip indicates the inclination of the fracture surface and the dip angle indicates the degree of inclination. The parameters (α, β) need to be transformed into polar coordinates (θ, φ) when drawing a fracture disk. The corresponding relationship is shown in Eq. (1).

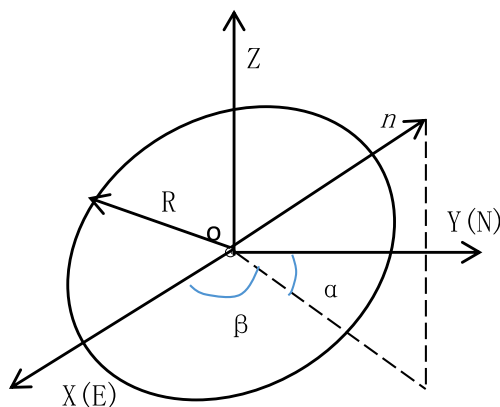


Fig. 2 All parameters in the disk model (n is the normal vector)

A simulation of a three-dimensional fracture network produces many intersecting thin circular disks with different sizes and orientations on three-dimensional space level. As it is impossible to directly measure all of the intact fracture planes, some properties of fractures are determined by probability distributions chosen, which include logarithmic normal distribution, uniform distribution, Fisher distribution, exponential distribution, and gamma distribution. After knowing the types of possible probability distributions for geometric parameters of the fracture system, random sampling by the Monte Carlo method is used to obtain the geometrical parameters to form a concrete fracture network. Consequently, a specific fracture network which has statistical hydraulic equivalence with the actual rock mass is generated in the study area.

Theoretically speaking, the combination of structural surface simulation and computer graphic technology can visualize the geological structure. Although many scholars have conducted a large amount of researches on 3D modeling of rock mass structures, these models are still not systematic or mature because the structural surface exists in the interior of the rock mass, indicating that measurements on the geometric parameters of the joint can only be taken in the area where bedrock is exposed. Therefore, local geometry parameters of joint were collected and used to construct an image of the joint network.

Based on the original data, Kriging interpolation was used to obtain surface fracture data for the entire karst area prior to the process and reconstruction by 3D simulation because in reality, the entire study field could not be sufficiently covered by sampling. As for underground data, the Monte Carlo method was used to simulate about 10,000 pieces of fracture data 500 m from surface down to underground. Afterwards, a seepage network can be built with a water flow path recognition algorithm.

$$\begin{aligned} \phi &= \frac{\pi}{2} - \alpha, \alpha \leq \frac{\pi}{2} \\ \phi &= 2.5\pi - \alpha, \alpha \geq \frac{\pi}{2} \\ \theta &= \beta \end{aligned} \tag{1}$$

Optimization of fracture intersection judgment

Intersection judgment between two fractures

In the study of fluid flow in fractured rock masses, the mutual positional relationship of the fractures determines the continuity of fluid flow. The amount of seepage calculation exponentially increases with the number of fracture plane increases. As the number of fractures in nature is huge, when the number reaches hundred thousand levels, the corresponding seepage

calculation is very time-consuming. However, the disjoint fractures do not contribute to the seepage and only the intersecting fractures draw our attention, meaning that the efficiency of the seepage calculation can be improved by retaining the intersecting cracks and eliminating the isolated fractures. Thus, fracture intersection judgment is the basis of seepage calculations.

To simplify the question, the intersection algorithm of two fracture disks is considered in this paper, Fig. 3 shows the sketch map of two intersecting fractures. In order to judge whether disks 1 and 2 are intersected in the space, assuming the geometric parameters of disk 1 and 2 as follows: the center coordinates of disk 1 is (X_1, Y_1, Z_1) , dip angle α_1 , dip β_1 , radius r_1 and the center coordinates of disk 2 is (X_2, Y_2, Z_2) , dip angle α_2 , dip β_2 , radius r_2 . The plane of the fracture disk can be described as eq. (2).

$$aX + bY + cZ + d = 0 \tag{2}$$

where

$$a = \sin \alpha \times \sin \beta; \quad b = \sin \alpha \times \cos \beta; \quad c = \cos \alpha; \quad d = -(aX + bY + cZ).$$

So, the plane of two fracture disks in Fig. 3 can be expressed as formula (3).

$$\begin{cases} a_1X + b_1Y + c_1Z + d_1 = 0 \\ a_2X + b_2Y + c_2Z + d_2 = 0 \end{cases} \tag{3}$$

The angle θ between two disk planes can be expressed by formula (4).

$$\cos \theta = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{(a_1^2 + b_1^2 + c_1^2)(a_2^2 + b_2^2 + c_2^2)}} \tag{4}$$

When $\theta \neq 0$, the intersection of the plane can be calculated.

The distance from the center point of the disk to intersection line can be obtained by formula (5).

$$D_i = \sqrt{\frac{((y_i - y_0)p - (z_i - z_0)n)^2 + ((z_i - z_0)m - (x_i - x_0)p)^2 + ((x_i - x_0)n - (y_i - y_0)m)^2}{m^2 + n^2 + p^2}} \tag{5}$$

where x_0, y_0, z_0 is any point in the intersection line, m, n, p is the direction vector of the intersection line, which can be represented as formula (6).

$$\begin{cases} m = b_1c_2 - b_2c_1 \\ n = a_2c_1 - a_1c_2 \\ p = a_1b_2 - a_2b_1 \end{cases} \tag{6}$$

According to the knowledge on analytic geometry, the two disks are not intersected if $D_1 + D_2 > r_1 + r_2$. Otherwise, the relationship between distance from the center point of two disks to the intersection line D_1, D_2 and their radius r_1, r_2 should be judged separately. The two disks are intersected only if $D_1 < r_1$ and $D_2 < r_2$.

Highly efficient fracture intersection judgment method

Partition algorithm can further improve the seepage calculation efficiency. At present, the basic mechanism behind the commonly used partition algorithm is to reduce the size of the map and break it into parts. The purpose of a sub-block in this case is to reduce the number of fractures in every operation and speed up computation through batch processing, i.e., the unconnected graphs are artificially divided in space to reduce the value of n , which is the number of fractures in a sub-block. A study area is divided into several blocks, and the number of center points contained in a block is the size of a block. Although the computation time increases due to partitioning and splicing of the graph, the complexity of the original problem can be greatly reduced. However, because of the different structure of fractures, it is not possible to determine which block the intersection between fractures located at the boundary of two blocks belongs to during an actual calculation. The space can be divided into three-dimensional space, or it can be divided along one direction. Different block schemes have different impacts on the results of the algorithm, as shown in Fig. 4.

The partitioning relies on parallel computation, i.e., the entire model is broken into sub-blocks and the largest calculation time among all blocks is taken as the time complexity. Moreover, a blocking method is usually combined with the other optimization algorithms like axial bounding box algorithm which will be introduced in detail in ‘‘Simplification of path traversal’’, because intersections between fracture disks also need determining in each individual block. The time complexity of the algorithm is $O(n^2)$ if the method of exhaustion is used in a sub-block, and the time complexity is $O(n)$ if the axial bounding box algorithm is used in a sub-block. If the method of exhaustion and the axial bounding box algorithm

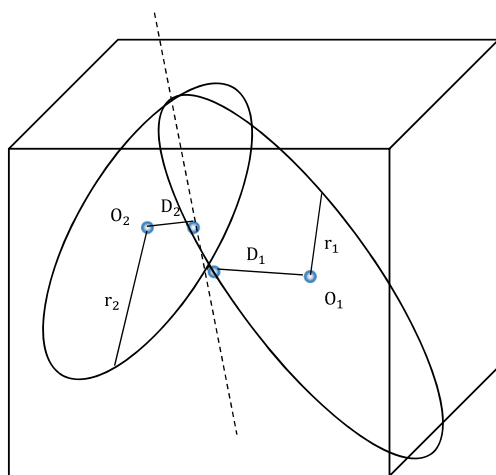
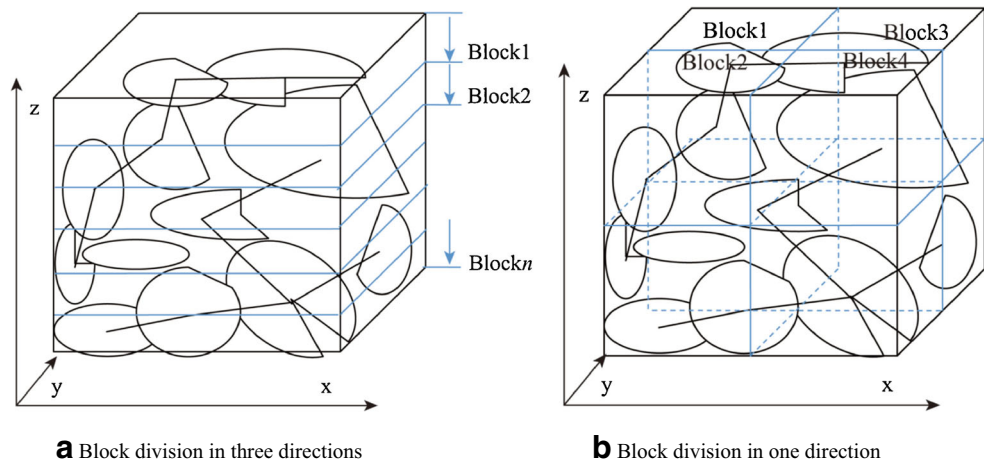


Fig. 3 Sketch map of two intersecting fractures

Fig. 4 Intersection evaluation in different block schemes. **a** Block division in three directions. **b** Block division in one direction



are used in a sub-block, then the time complexity of the entire algorithm is $\max(O(n^2), O(n))$.

Another algorithm is the R-tree algorithm. R-tree is a type of highly balanced tree which is a natural expansion of a B-tree in K -dimensions. It uses a minimum bounding rectangle (MBR) to approximately represent an object in space. The object occupying a certain volume in space can be directly indexed with the R-tree algorithm based on the object's MBR. The R-tree index algorithm can effectively exclude isolated fractures and quickly determine the set of fractures that may intersect to further evaluate intersections. In principle, the R-tree index algorithm is a type of combination of the axial bounding box algorithm and the block algorithm.

Figure 5a, b shows schematic of the fracture disk bounding box distribution in space and the corresponding R-tree graph. Each node in the R-tree index structure is a fracture disk corresponding to an MBR. All non-leaf nodes are MBRs of its child node. The R-tree index generally traverses down from the root node layer by layer and can quickly exclude features that do not meet the screening requirements and reduce unnecessary calculation. The time complexity of the R-tree index algorithm is $O(\log_2 n)$, which is greatly reduced compared with the blocking and axial bounding box algorithms. Especially when the amount of fracture data is huge, the R-tree indexing algorithm will greatly accelerate execution of the program.

Seepage identification of water flow along the fractures

After acquiring surface fracture data by Kriging interpolation and the underground fracture data with the Monte Carlo method, the fast intersecting algorithm was used to judge whether the two fracture surfaces intersect and a karst fracture seepage simulation based on the fast

intersecting algorithm was applied to dynamically simulate groundwater flow direction and directly describe the distribution of underground fractures. In the data structure, a large-scale fracture network can be expanded to an unconnected graph in which each crevice point (center of circle) is the vertex of the graph, and the seepage path is the edge of the graph. Considering the groundwater flow direction, a directed graph can be used to simulate the entire underground fracture network.

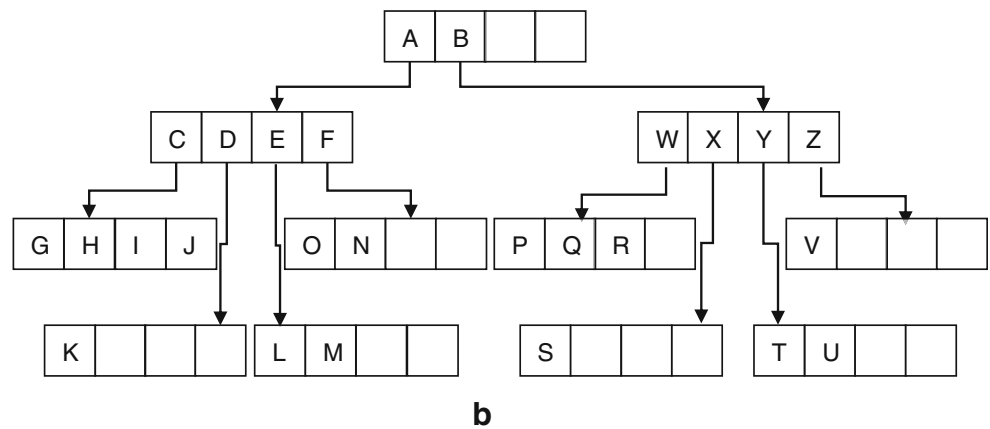
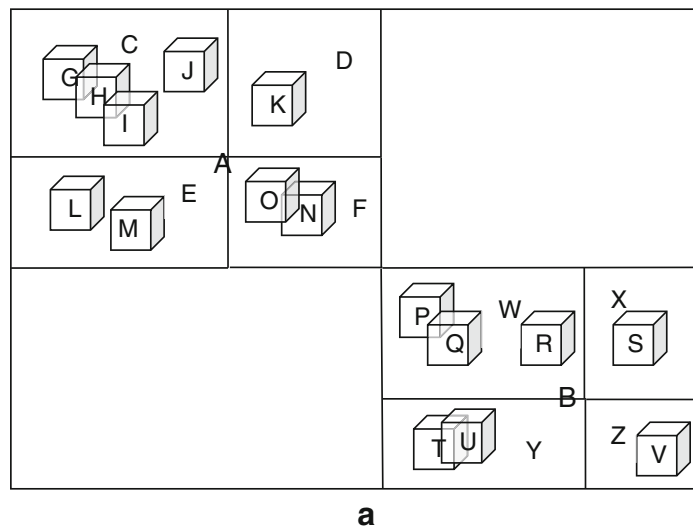
Directed graphs to identify and store fractured networks

The arrow between two connecting fractures is the water flow path, as shown in Fig. 6. In the data structure, the directed graph is stored and traversed with an adjacency matrix or adjacency list. The connecting path between the fractures is shown in Fig. 7, and the lines between adjacent disks in the diagram indicate that adjacent disks are connected.

The adjacency matrix represents adjacency between vertices. Supposing $G = (v, e)$ is a graph, where $v = \{v_1, v_2, \dots, v_n\}$. The connectivity between nodes can be stored in a two-dimensional array, which is the adjacency matrix. The length of each row (or each column) of the adjacency matrix is the number of vertices in the graph. Figure 8 shows a directed graph. According to the definition of the adjacency matrix, an adjacency matrix is created for this directed graph using Eq. (7).

For two connecting vertices in the directed graph, the vertex where the arrow starts is the starting point, and the vertex pointed by the arrow is the ending point. Each position in the adjacency matrix indicates whether or not two vertices at that position are connected. For example, A, B, C, D, and E in Fig. 8 denote vertices; a, b, c, d, e, f, g, and h denote paths. There is a path between C and D, so 1 is placed at CD in the corresponding adjacency matrix. In contrast, there is no path between B and A, so 0 is stored at BA in the adjacency matrix. The adjacency matrix is stored in a linear table structure.

Fig. 5 **a** Fracture disk bounding box distribution in space. **b** R-tree corresponding to **a**



$$\begin{matrix}
 A & B & C & D & E \\
 \left. \begin{matrix}
 A \\
 B \\
 C \\
 D \\
 E
 \end{matrix} \right\} \begin{pmatrix}
 0 & 1 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 \\
 0 & 1 & 0 & 1 & 0 \\
 0 & 1 & 0 & 0 & 0 \\
 1 & 0 & 1 & 0 & 0
 \end{pmatrix}
 \end{matrix}$$

(7)

The seepage path for fractures can also be stored in a form of an adjacency list, which is a single linked list. Although the application of an adjacency list can save storage space and is more convenient for basic graph operations such as adding and deleting vertices, the adjacency matrix is used to store the directed graph because the edge table relationship is

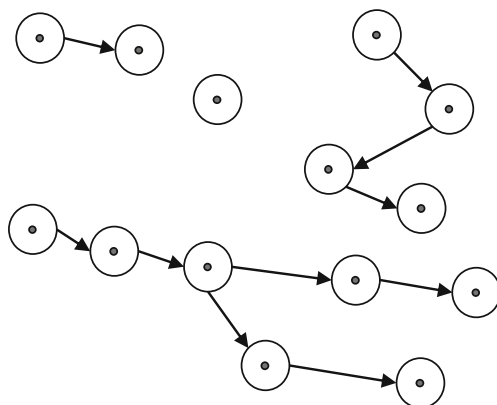


Fig. 6 Underground fracture network generalized as a directed graph

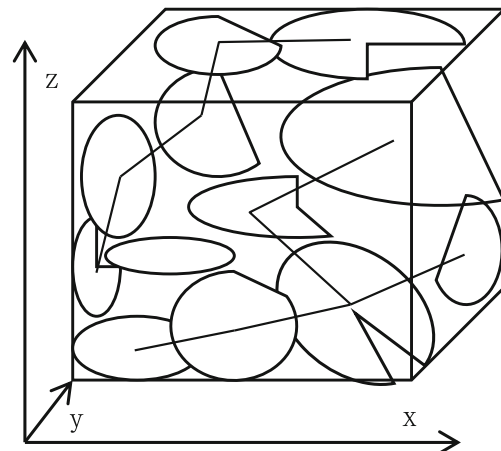


Fig. 7 Centers of intersecting disks acting as path nodes

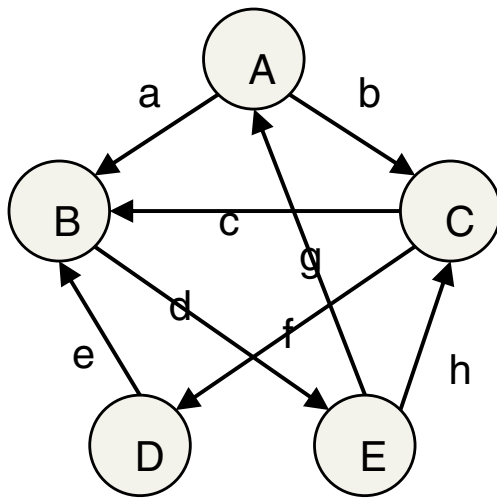


Fig. 8 Directed graph

involved in the adjacency list structure, which is not known in the system and can be a barrier to the usefulness of the subsequent algorithm.

According to the fast intersection algorithm for fractures, the intersected fractures are used to build paths between vertices in a directed graph. And based on the knowledge on hydromechanics, the fluid flows from high hydraulic head to low hydraulic head, determining the flow direction in networks. Intersection of fractures is evaluated and the results of judgment are stored in the adjacency matrix. If two fractures intersect, the hydraulic heads of the fractures are compared and the number “1” is placed at the position where vertex corresponding to fractures with higher hydraulic head and fractures with a lower one intersect in the adjacency matrix. After building the adjacency matrix, a seepage path is formed by connecting the two fractures with a value of 1 in the adjacency matrix.

Simplification of path traversal

During simulation of the seepage paths, there are more likely to have tens of thousands of fracture data records sampled in the field, not to speak of the random fracture data generated with Kriging interpolation and Monte Carlo stochastic simulation. Evaluating whether these fractures intersect with each other requires a considerable amount of time and consumes all available computer memory in an ordinary computer. Therefore, simplifying the calculation process can greatly improve the efficiency of the computer.

The fracture disk is intricate. There are many types of intersecting morphologies of fracture disks, including the existence of isolated fractures which do not intersect with other fractures. Isolated fractures are involved in the intersection evaluation process prior to seepage path identification algorithms. That is to say, if the number of fracture data is assumed to be n , there will be $n-1$ invalid operations for each isolated

fracture, which means that the temporal and spatial complexity of the entire fracture recognition algorithm will increase by a power of 2 if the number of isolated fractures becomes large. In conclusion, the main measure for simplifying the seepage path and reduce the calculation of the intersection of fracture disks is to exclude isolated fractures as much as possible.

The method of exhaustion is the most direct intersection evaluation algorithm. Assuming that the number of fractures is n , the number of judgments to be performed is n^2 , and the time complexity is $O(n^2)$, which means that if the number of fractures is huge, the processing time of the computer can be tens of hours, or even days. Figure 9a shows the method of exhaustion for evaluating intersections. As a result, the most commonly used algorithm in 3D geology appears, which is referred to as the axial bounding box algorithm. The axial bounding box algorithm is considered a rough intersection evaluation algorithm. Each fracture disk has a minimum axial bounding box, i.e., their smallest external cube. If two fracture disks intersect, their minimum axial bounding boxes must intersect. The axial bounding box intersection evaluation algorithm is shown in Fig. 9b. The minimum axial bounding box algorithm can be used as an optimization algorithm to judge intersection because it is easier to determine the intersection of bounding boxes than to directly determine the intersection of fracture disks, thus reducing the calculation time. The time complexity of bounding box is $O(n)$.

Simulation of seepage water flow

In the graph structure, the center of fracture is represented as the vertex of the graph, and the seepage path is the edge of the graph. As shown in Fig. 10, the lines between the two intersecting fractures are the seepage paths.

Take the length of each fracture segment $l_j (j = 1, 2, \dots, m)$ in the fracture network as the weight of the seepage flow in the fracture network, and the m -order diagonal matrix D represents the adjacency matrix of the weighted fracture network.

$$D = \begin{bmatrix} l_1 & 0 & \dots & 0 \\ 0 & l_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & l_m \end{bmatrix}$$

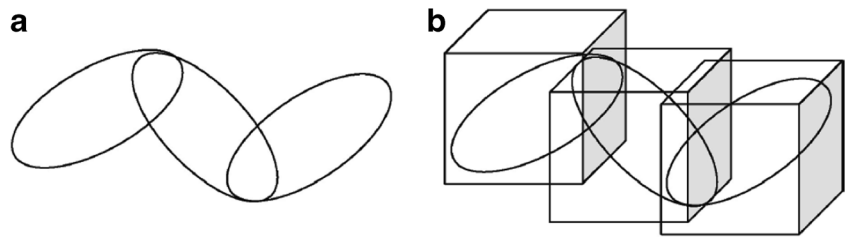
Based on the fracture network model, the seepage velocity of each fracture segment is

$$v = K * D^{-1} * M^T * \Delta H \tag{8}$$

where K represents the hydraulic conductivity, ΔH is the hydraulic-head difference between two points, and M is the adjacency matrix shown in eq. (7).

In addition to the hydraulic conductivity K , the seepage simulation also requires the use of water flow q , flow velocity v , permeability k , and permeability coefficient tensor \vec{k} .

Fig. 9 Evaluating intersections using axis-aligned bounding boxes



The hydraulic conductivity is an important hydrogeological parameter. In isotropic media, the value of K which is a scalar and independent of the direction of the percolation is defined as the unit flow rate per unit hydraulic gradient, indicating how easy the fluid passes through the porous media. The K value depends not only on the properties of the rock (i.e., particle size, texture, particle arrangement, filling condition, and development degree, etc.), but also on the physical properties (i.e., capacity, viscosity, etc.) of the osmotic liquid. Theoretical analysis shows that the size of the void plays a major role in the K value. Permeability is the penetration rate of a liquid in media with a dynamic viscosity coefficient of 1 and the pressure gradient of 1. However, in anisotropic media where the direction of the percolation is arbitrary, the K value is related to the direction of seepage. Since the hydraulic gradient disagrees with the seepage direction, a permeability coefficient tensor is introduced.

The methods to obtain these parameters mainly include in situ hydraulic experiments, fracture geometry measurement, and the numerical inversion of seepage in discrete fracture networks. This article uses the second method.

The hydraulic conductivity of fractures with width b can be obtained by

$$K_f = \frac{\rho g}{12\mu} b^2 (\cos a)^2 = \frac{g}{12\nu_w} b^2 (\cos a)^2 \tag{9}$$

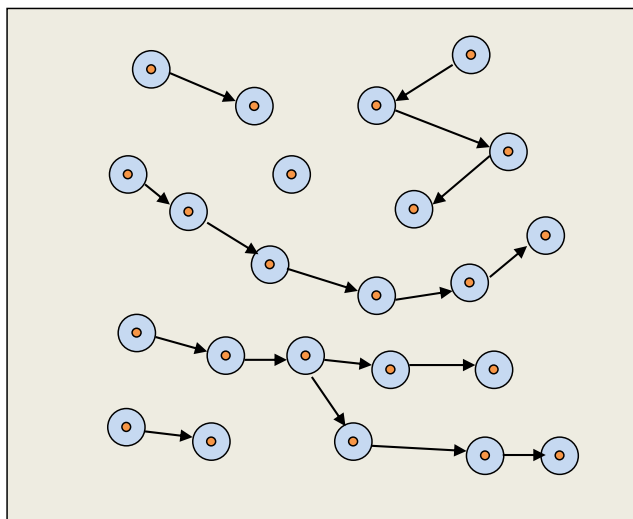


Fig. 10 Generalized underground fracture network as a directed graph

And the flow volume is

$$q_f = v_f * b = \frac{g b^3}{12\nu_w} J \tag{10}$$

Flow velocity

$$v = KJ \tag{11}$$

Permeability coefficient

$$k_{ff} = \frac{b^2}{12} (\cos a)^2 \tag{12}$$

From above equations, the relation between K and permeability coefficient (k) is

$$K = \frac{\rho g}{\mu} k = \frac{g}{\nu_w} k \tag{13}$$

In this paper, the Cartesian coordinate system is used, and the permeability coefficient tensor \vec{k} is set up based on the flow movement in a single fracture:

$$\begin{aligned} \vec{K} &= \begin{bmatrix} k_{xx} & k_{xy} & k_{xz} \\ k_{yx} & k_{yy} & k_{yz} \\ k_{zx} & k_{zy} & k_{zz} \end{bmatrix} \\ &= K_e \begin{bmatrix} 1 - (\cos\beta\sin\gamma)^2 & -\sin\beta\cos\beta(\sin\gamma)^2 & -\cos\beta\sin\gamma\cos\gamma \\ -\sin\beta\cos\beta(\sin\gamma)^2 & 1 - (\sin\beta\sin\gamma)^2 & -\sin\beta\sin\gamma\cos\gamma \\ -\cos\beta\sin\gamma\cos\gamma & -\sin\beta\sin\gamma\cos\gamma & 1 - \cos^2\gamma \end{bmatrix} \end{aligned} \tag{14}$$

For the above equations, μ is the dynamic viscosity coefficient, J is the hydraulic gradient, ν_w is the viscosity coefficient for water, g is the gravity acceleration, a is the angle of fracture and water flow direction, b is the fracture width, ρ is density, β is the dip, and γ is the dip angle.

Case study of the algorithm

The karst area in Zhangfang, Fangshan district in Beijing, is a typical representative of northern karst, and fracture data was systematically collected to provide a reliable basis for accurately constructing the fracture network model and evaluating karst development. A total number of 2069 sets of joint data were obtained

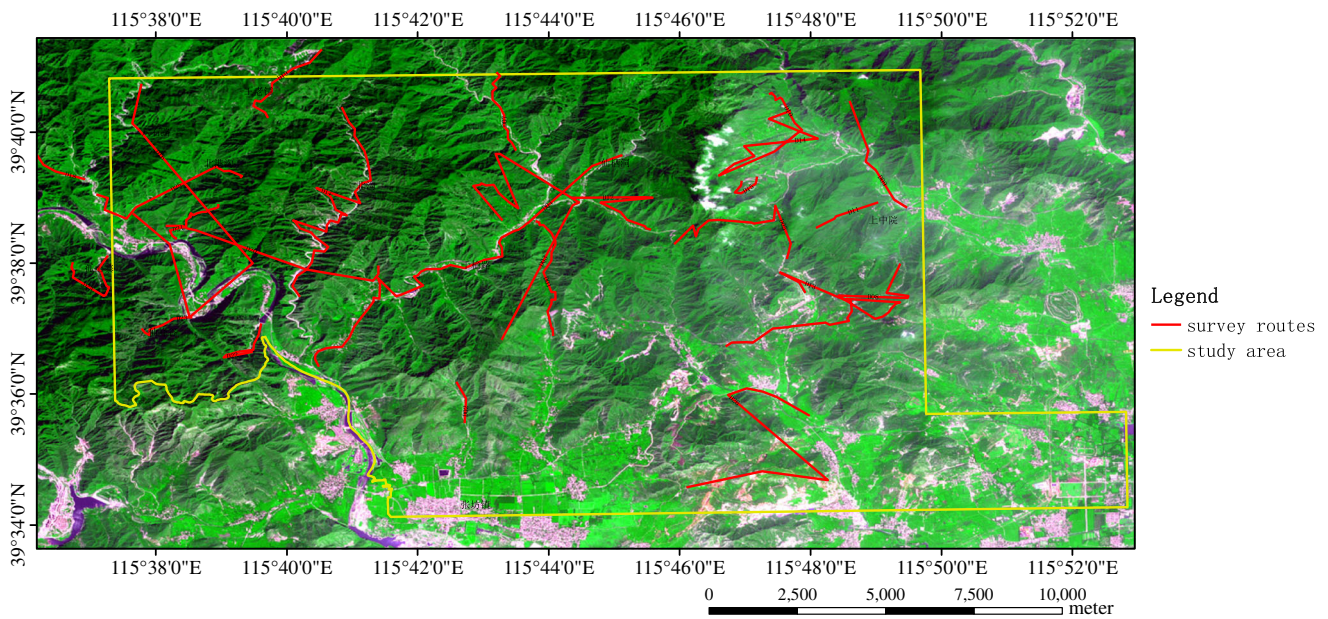


Fig. 11 Simulation map of the study area (with remote sensing base map)

during field survey. Most fractures in the karst area are classified as joints, and the acquired joint data are assumed to give a comprehensive description of the properties of the fracture. These data include spatial attributes such as joint name, rock characteristics, dip, dip angle, frequency, silt width, extension length, latitude, and longitude. Surface fracture data covering the entire study area were obtained by interpolation. By comparison, data generated from Kriging interpolation highly coincide with the measured data. After that, the Monte Carlo method was used to extend the fracture data from surface into the ground to form a complete spatial distribution of the fracture in the karst area.

Zhangfang and its periphery area of Fangshan district in Beijing were chosen as the study area. The Zhangfang karst-water distributed area, as one of the seven karst distribution areas in Beijing, is a relatively independent hydrogeological unit. The study area is shown in Fig. 11 is about 20,000 m * 12,000 m and the karst fractures distribute from surface to 500 m underground. The processed remote sensing image from this area was used as the base map, where the yellow line denotes the fault zone, the blue part on the left being the lamprophyre vein, and the purple part on the right being the monzogranite vein. The surface fracture data was obtained mainly by fine outcrop exploration in the study area. The red part in Fig. 11 is the sampling routes, where the parameters of fractures such as the direction, dip, dip angle, width, length, and development density were sampled every 5 m. In total, more than 8000 sets of fracture data were measured.

Interpolation of surface fractures

After sampling, the fractures on the surface can be interpolated by a quadrat based on Kriging method. In terms of the idea of quadrat, this paper proposed a fracture quadrat model dividing the study area into grid cells with specific size, as shown in Fig. 12. Fracture data statistics were carried out for each single grid cell. The cell was set up to null if no fractures were observed, otherwise the fracture data in the cell was counted. The statistical method was consistent with the method of drawing the fracture-density-rose diagram: The sampled data of fracture occurrence and density were grouped by spatial azimuth interval (such as 5° or 10°), and statistical information (i.e., the number of fractures, average dip, average width, and length) of each group was summarized. In this way, each grid cell containing the fracture data corresponds to a rose chart and this grid cell can be used as a known point to interpolate the other grid cells in the unknown area.

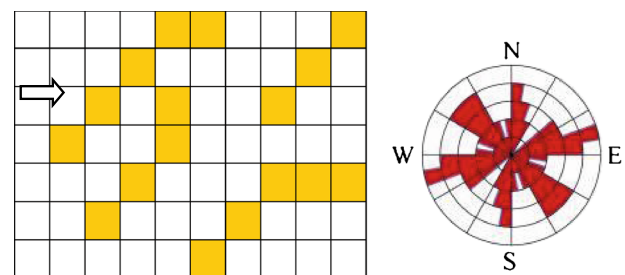


Fig. 12 Quadrat model of fractures data and its statistics

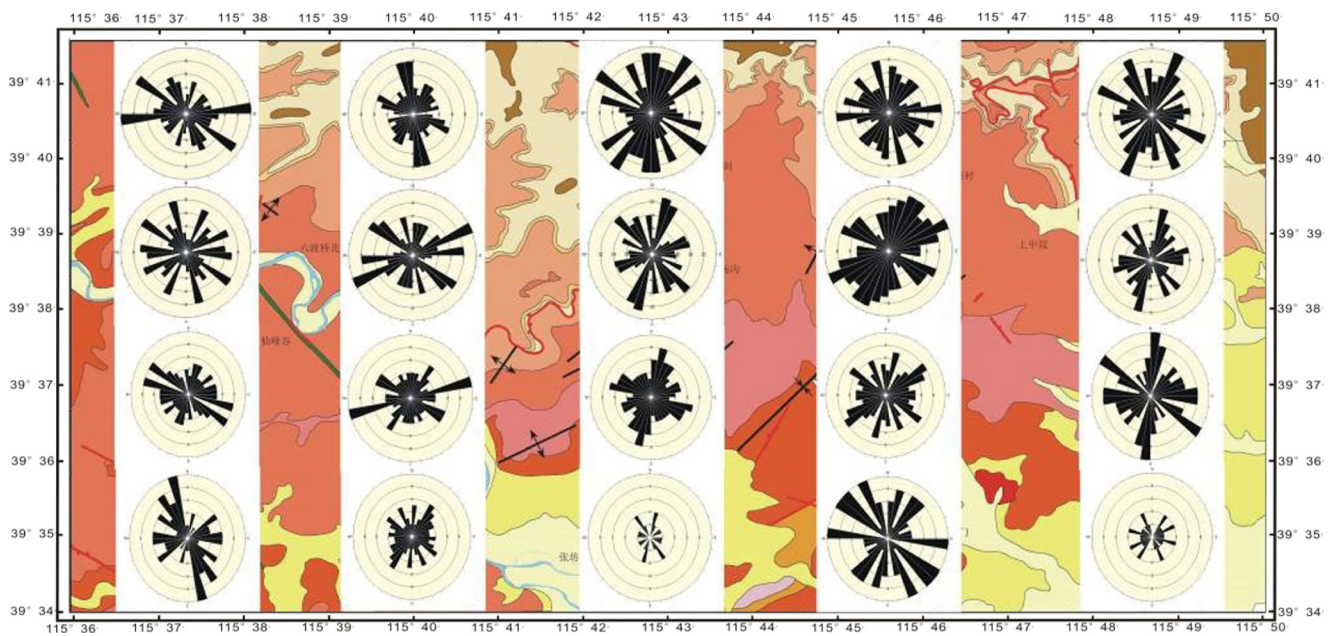


Fig. 13 Spatial distribution of surface joints

In the interpolation, each direction group in the fracture-rose-chart can be interpolated individually and sequentially, and statistic information mentioned above was used. Thus, the fracture-rose-charts for each of the other grid cells can be obtained by Kriging interpolation. The interpolated rose diagram is shown in Fig. 13.

Simulation of underground fractures

This paper assumes that uncertainty exists in the karst underground space. In the modeling process, only the original measured data of the outcrops near the surface and the interpolated surface fracture data were utilized. There is evidence showing

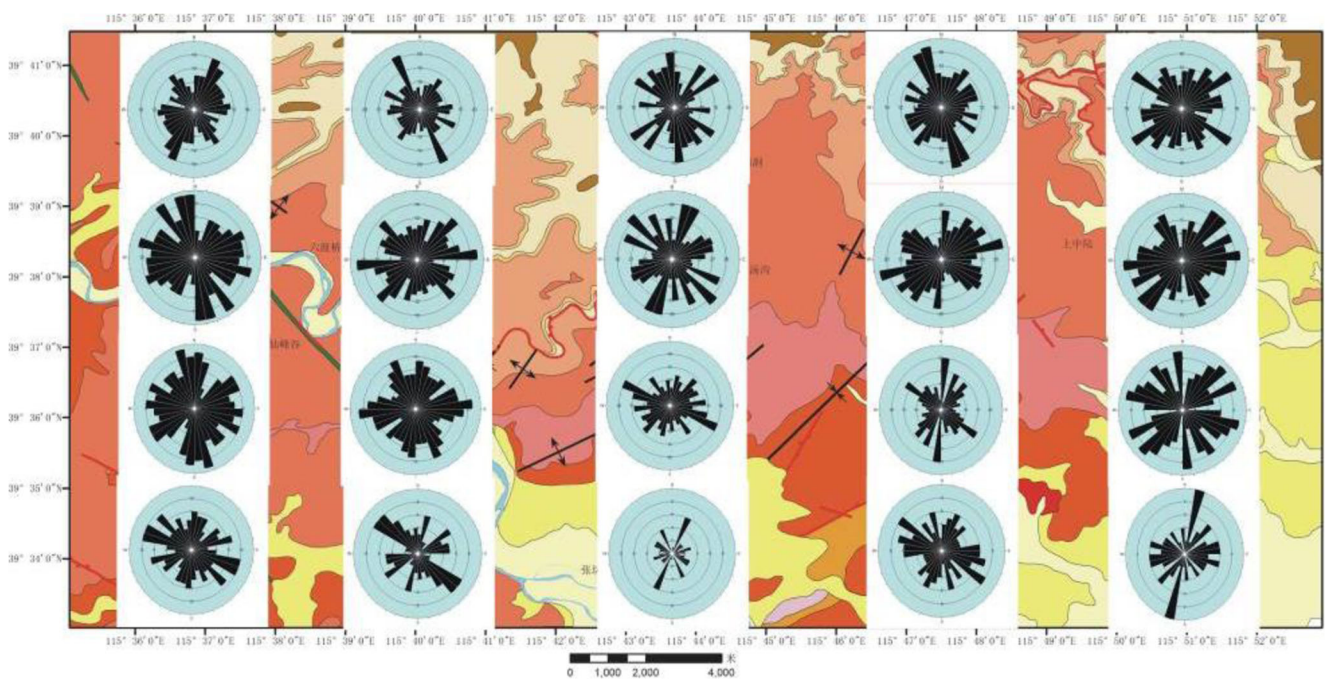
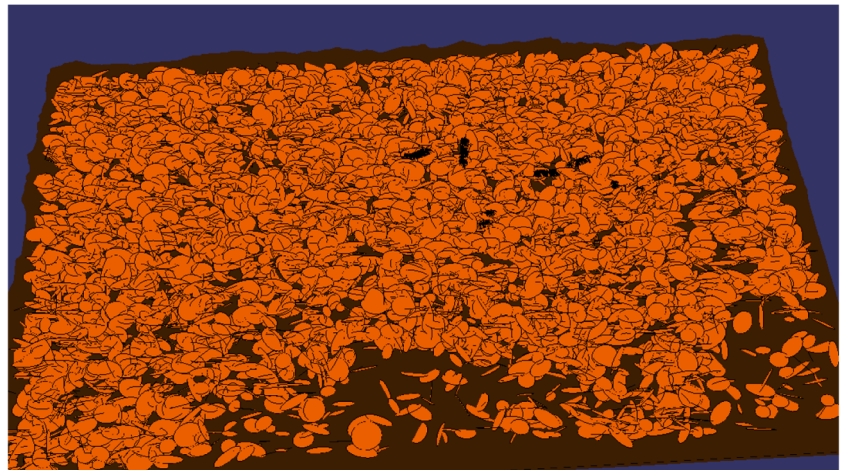


Fig. 14 Underground fractured joint spatial distribution

Fig. 15 The simulation of the fracture in the whole study area



that there is a spatial correlation between fractures in the underground space. The fact that the fractures are not independent, but have a certain spatial relation with each other provides a possibility for speculating underground fractures from surface data. This paper took the feature of dip angle as an example to illustrate the property simulation for the subsurface fracture. In this paper, the known geoscience knowledge, statistical methods and engineering experience can be regarded as the prior knowledge to speculate underground fracture data. The combination of prior knowledge with the measured data was achieved by the Bayesian theory and the problem of sampling fractures with spatial variability was solved by the Markov chain and Monte Carlo method.

The distribution of the dip angle of the underground fractures satisfies the normal distribution and the probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty \quad (15)$$

where μ and σ are the expectation and the standard deviation of the dip, respectively, of which a rough estimate of the range can be given by the priori knowledge. For example, the expectation and standard deviation of the dip can be roughly estimated by the fracture-rose diagram.

Fig. 16 Local fracture map



There are many combinations of μ and σ , and each combination is possible. According to the maximum entropy theory, the more the uncertainty is retained in the probability model, the better the model will be. Given the a priori knowledge and the field data, from the total probability theorem, the probability of dip is

$$P(dip|prior, test) = \int_{\mu, \sigma} P(dip|\mu, \sigma)P(\mu, \sigma|prior, test)d\mu d\sigma \quad (16)$$

where $P(dip|\mu, \sigma)$ is the probability density distribution of the dip given a set of μ and σ . Since it conforms to the normal distribution, it can be expressed by the formula (1). $P(\mu, \sigma|prior, test)$ indicates the probability density distribution of μ and σ under the conditions of prior knowledge and measured data.

In statistics, the joint probability distribution of an event is equal to the product of the conditional probability and the marginal probability. Assume event A, B, event A represents the ground-truth probability distribution of model C, and event B represents the probability distribution obtained from measured data, then there is:

$$P(A, B) = P(A|B)*P(B) = P(B|A)*P(A) \quad (17)$$

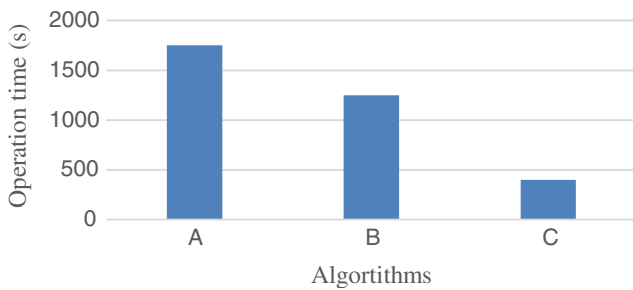


Fig. 17 Result on efficiency of different algorithms

Thus:

$$P(A|B) = P(B|A) * P(A) / P(B) \tag{18}$$

Since event B is the probability distribution of event C obtained through field observation, the marginal probability $P(B)$ of event B is a fixed value, then there is

$$P(A|B) \propto P(B|A) * P(A) \tag{19}$$

where $P(A|B)$ is the posterior probability distribution of the model C, that is, the original probability distribution of the model C given the observation value; $P(A)$ is the prior distribution of model C; $P(B|A)$ is the likelihood distribution.

Thus:

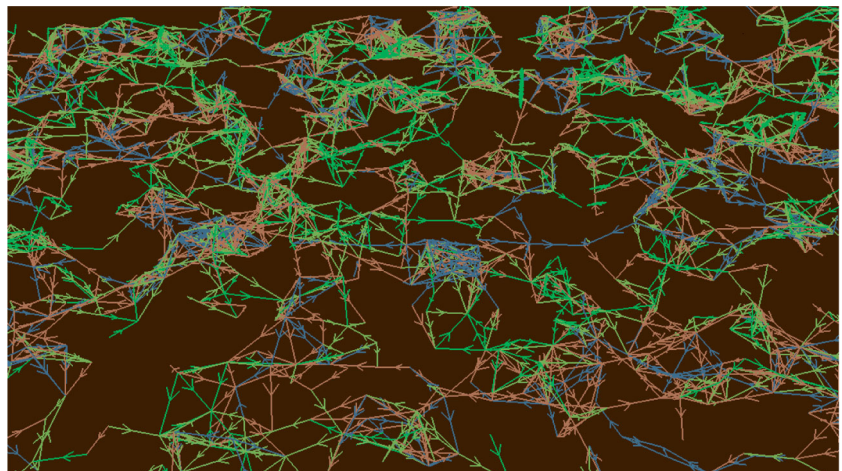
$$P(\mu, \sigma | prior, test) = KP(prior, test | \mu, \sigma) P(\mu, \sigma) \tag{20}$$

where $K = \frac{1}{\int_{\mu, \sigma} P(prior, test | \mu, \sigma) P(\mu, \sigma) d\mu d\sigma}$ is the constant.

$P(\mu, \sigma)$ is the distribution probability without any measured data, that is, the prior probability distribution, or the joint probability distribution of μ and σ . This paper assumes that μ and σ conform to uniform joint distribution, and the probability distribution is

$$P(\mu, \sigma) = \frac{1}{(\mu_{max} - \mu_{min})(\sigma_{max} - \sigma_{min})} \quad \mu \in [\mu_{min}, \mu_{max}], \sigma \in [\sigma_{min}, \sigma_{max}] \tag{21}$$

Fig. 18 Dynamic groundwater flow path in the fracture network



Therefore, $P(\mu, \sigma)$ can be calculated by obtaining the range of μ and σ based on priori knowledge.

$P(prior, test | \mu, \sigma)$ is the likelihood probability which can be obtained by maximum likelihood estimation. The maximum likelihood method argues that if only one test is performed and the result A occurs, the condition is considered to be favorable for the result A. This paper assumes that each fracture sample is independent, and its likelihood probability is consistent with:

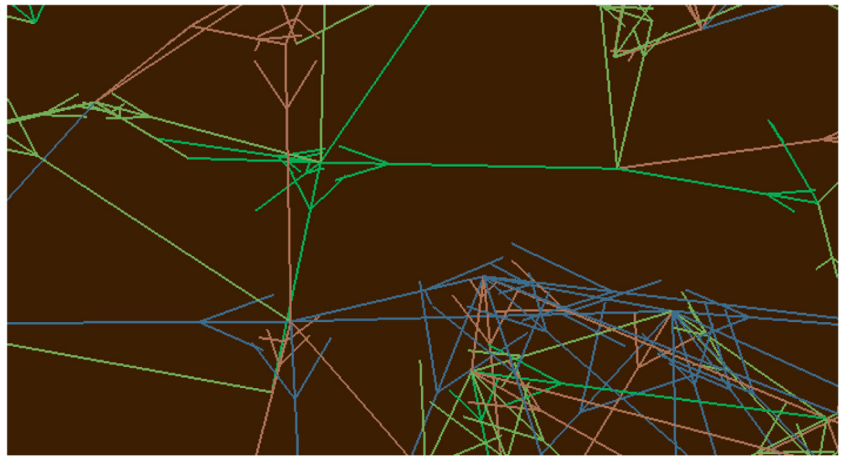
$$L(\theta) = L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta) \quad \theta \in \Theta \tag{22}$$

where θ is the parameter to be estimated, Θ is the range of possible values of θ , x_1, x_2, \dots, x_n are sample values of events X_1, X_2, \dots, X_n , $\prod_{i=1}^n f(x_i; \theta)$ is the joint probability distribution when events X_1, X_2, \dots, X_n takes value x_1, x_2, \dots, x_n , $f(x_i; \theta)$ is the probability density function, i.e., $P(dip | \mu, \sigma)$ in this case. Finally, $P(dip | prior, test)$ is specified.

To simulate the unknown properties of underground fractures from the measured fracture data, a random sampling based on the posterior probability distribution of the underground fracture is required. Since the process of generating random numbers by computer, which depends on certain algorithms, is not a random process in the actual sense, the generated random number is a pseudo-random number. And the Monte Carlo method is an important method to generate such a pseudo-random number.

In geostatistics, there is a correlation between fractures of which the spatial distribution exhibits inhomogeneity and complexity and conforms to certain spatial variability characteristics. That is to say, the properties between the fractures are mutually related. Monte Carlo random sampling based on Markov chain provides an effective way to deal with the complexity in sampling. The core idea is to construct a Markov chain in the initial state space, of which the stable distribution

Fig. 19 Local map of the dynamic groundwater flow path in the fracture network



is the probability distribution of the target, and then transfer the current state to the next state by the state transition function (the next state only related to the current state). After a period of sampling, the probability distribution is approached to the distribution of target in the real situation, and the sample from this probability distribution is finally selected (Zhao 2014).

The fracture data is sampled based on the Gibbs sampling algorithm. Assuming that the k -dimensional sample space $x = (x_1(0), x_2(0), \dots, x_n(0))$, and the probability distribution of its initial state is $\pi(x_1(0), x_2(0), \dots, x_n(0))$. Then, according to the conditional transition probability $P(x_{k+1} | x_1(0), x_2(0), \dots, x_k(0))$, the $(k+1)$ th value is obtained, which is denoted as $x_{k+1}(1)$. The $(k+2)$ th value, which is denoted as $x_{k+2}(1)$, can also be obtained by the conditional transition probability $P(x_{k+2} | x_1(0), x_2(0), \dots, x_k(0), x_{k+1}(1))$ and repeating this process. According to the law of large numbers, when the number of samples is large enough, the probability distribution of the sample approximates the target distribution under real conditions (Li 2008).

The spatial distribution map of the speculated underground joints is shown in Fig. 14.

Results

Open Scene Graph (OSG) was used to construct the “karst fracture simulation system”. Terrain data was loaded as the base map in “Karst fracture simulation system of Zhangfang, Fangshan district in Beijing” as shown in Figs. 15 and 16. The number of fractures generated with the Monte Carlo simulation was around 65,000.

Moreover, the efficiency of the algorithm proposed in this paper, R-tree index algorithm combined with the optimization of block algorithm, is compared with the exhaust algorithm and the R-tree index algorithm in terms of operation time. Figure 17 shows the result of different algorithms, where A represents exhaust algorithm; B represents R-tree index algorithm; C represents the optimized algorithm combining the R-

tree index algorithm with block algorithm. The efficiency of C is higher than other two algorithms obviously.

A dynamic flow direction diagram can be obtained by evaluating fracture intersections and the seepage flow path (Figs. 18 and 19). The arrows in the figure are dynamically updated by the callback function in OSG. The position of the arrow indicates the position of the fluid, and the direction of the arrow indicates the fluid flow direction.

Conclusions

Seepage simulations are an important aspect in the study of three-dimensional underground fractures and their permeability. Choosing a suitable fracture model is a prerequisite for accurately simulating the groundwater seepage field under different geological conditions.

From the seepage simulation results, the following conclusions can be drawn:

(1) In this paper, the disk model was used to simulate the fracture planes due to its simplicity and efficient evaluation of fracture intersections. However, this does not mean this model is better than the others. Nowadays, researches on the construction of a fracture plane have been more and more complicated. But in computer simulation, the convenience of the disk model makes the disk model more applicable to large study area and large amount of data.

(2) The fast intersection evaluation algorithm in this paper is based on the disk model and the associated optimization algorithm aims at rapidly evaluating intersections. The “karst fracture simulation system” was implemented and tested. The number of isolated fractures is large and most fractures were randomly simulated with the Monte Carlo algorithm. Therefore, the removal of isolated fractures and reduction of unnecessary intersections can greatly improve the efficiency of the system. When the boundary problem is neglected, optimization of the block algorithm can also assist with implementation of the R-tree algorithm.

Funding information This study was supported by the Beijing Natural Science Foundation (Grant No. 8172046), the National Natural Science Foundation of China (Grant No. 41771478), and the Hebei Province Natural Science Fund (Grant No. D2016210008).

References

- Baecher GB (1977) Statistical description of rock properties and sampling. In Proc. of the U.S. symposium on rock mechanics
- Batu V (2009) Fracture control of ground water flow and water chemistry in a rock aquitard. *Ground Water* 47:488–4U1
- Berkowitz B (2002) Characterizing flow and transport in fractured geological media: a review. *Adv Water Resour* 25:861–884
- Brkic Z, Kuhta M, Hunjak T (2018) Groundwater flow mechanism in the well-developed karst aquifer system in the western Croatia: insights from spring discharge and water isotopes. *Catena* 161:14–26
- Chen CX (2011) Groundwater hydraulics. China University of geosciences press, Wuhan
- Dershowitz WS, Einstein HH (1988) Characterizing rock joint geometry with joint system models. *Rock Mech Rock Eng* 21:21–51
- Dong Y, Ju YW, Zhang YX (2014) Characteristics of joints in karst area and their influence on karstification in Zhangfang in Fangshan region of Beijing. *J Univ Chin Acad Sci* 31:783–790
- Ford D, Williams P (2007) Karst hydrogeology and geomorphology. United Kingdom. Wiley, Chichester
- Gelasch CA, Bradbury KR, Hart DJ, Bahr JM (2013) Characterization of fracture connectivity in a siliciclastic bedrock aquifer near a public supply well (Wisconsin, USA). *Hydrogeol J* 21:383–399
- He SH (1997) Research on a model of seepage flow of fracture networks and modelling for the coupled hydro-mechanical process in fractured rock masses. In *Computer methods and advances in Geomechanics, Vol 2*, 1137–1142
- Hyman JD, Aldrich G, Viswanathan H, Makedonska N, Karra S (2016) Fracture size and transmissivity correlations: implications for transport simulations in sparse three-dimensional discrete fracture networks following a truncated power law distribution of fracture size. *Water Resour Res* 52:6472–6489
- Karimzade E, Sharifzadeh M, Zarei HR, Shahriar K, Seifabad MC (2017) Prediction of water inflow into underground excavations in fractured rocks using a 3D discrete fracture network (DFN) model. *Arab J Geosci*, 10
- Li HB (2008) Gibbs sampling in Markov chain Monte Carlo-based algorithm. Wu han: Hubei University
- Li ZY, Wang MY, Zhao JH (2014) Optimizing fracture intersection analysis procedure in 3D fracture network seepage simulation. *J China Coal Soc* 39:2286–2292
- Liu HM, Wang MY (2010) The computer processing and visualization system for 3-D fracture networks. In *Proceedings of 2010 3rd Ieee international conference on computer science and information technology*, 139–144
- Long JCS, Gilmour P, Witherspoon PA (1985) A model for steady fluid-flow in random 3-dimensional networks of disc-shaped fractures. *Water Resour Res* 21:1105–1115
- Medici G, West LJ, Mountney NP (2016) Characterizing flow pathways in a sandstone aquifer: tectonic vs sedimentary heterogeneities. *J Contam Hydrol* 194:36–58
- Qiao XJ, Hou QL, Ju YW (2014) Research about the control of geological structure on karst groundwater system in Zhangfang, Beijing. *Carsologicasinica* 33:184–191
- Robertson AM (1970) The interpretation of geological factors for use in slope theory
- Snow DT (1965) A parallel plate model of fractured permeable media
- Wang H (2016) Research on three dimensional visualization technologies. In *Proceedings of the 2016 4th international conference on advanced materials and information technology processing*, 261–265
- Zhao TY (2014) Markov chain Monte Carlo-based stability analysis for Pingzishan landslide. 31–32. Cheng du: Southwest jiao tong University