

Received April 16, 2020, accepted May 3, 2020, date of publication May 7, 2020, date of current version May 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2993050

A Triangulated Irregular Network Constrained Ordinary Kriging Method for Three-Dimensional Modeling of Faulted Geological Surfaces

QINGREN JIA¹, WENWEN LI², AND DEFU CHE^{1,3}

¹College of Resources and Civil Engineering, Northeastern University, Shenyang 110819, China

²School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287, USA

³Key Laboratory of Ministry of Education on Safe Mining of Deep Metal Mines, Northeastern University, Shenyang 110819, China

Corresponding author: Defu Che (chedefu_neu@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 41871310, in part by the Fundamental Research Funds for the Central Universities under Grant N17241004, in part by the China Scholarship Council (award to Jia for 1 year's study at the Arizona State University), and in part by the Qianjiaying Coal Mine.

ABSTRACT Three-dimensional (3D) modeling of geological surfaces, such as coal seams and strata horizons, from sparsely sampled data collected in the field, is a crucial task in geological modeling. Interpolation is a common approach for this task to construct continuous geological surface models. However, this problem becomes challenging considering the impact of the faults on geological surfaces. Existing methods tend to solve this problem through three steps, including interpolating stratum and fault surface, applying a fault modeling method to modify the geological surface, and optimizing the modified surface to pass sample points fallen into the fault displacement zone. This paper presents a more concise method to generate a faulted geological surface, in which 1) a constrained Delaunay triangulated irregular network (CD-TIN) is constructed to facilitate the neighborhood search process of the ordinary kriging (OK) interpolation, 2) the CD-TIN is also directly constrained by horizon cut-off lines formed from theoretical fault displacement profiles, and 3) subsequently, neighbors of the location to be estimated are selected effectively in the CD-TIN considering the fault topology. The proposed method significantly improves the time efficiency of the OK interpolation by utilizing the CD-TIN and incorporates fault effects directly into the interpolation process by inserting fault horizontal cut-off lines into CD-TIN. Moreover, by integrating the fault effects directly into the interpolation process, the surface modeling process is accomplished in a single stage instead of two separate stages of interpolation first and then modifying the surface in the fault area. By this strategy, the proposed method significantly improves the time efficiency of the OK interpolation algorithm and achieves more accurate modeling of the faulted geological surface. Experiments were designed to compare the performance of our method with several commonly used approaches, and the results indicate that the proposed TIN-constrained OK method achieves better accuracy and efficiency in modeling faulted geological surfaces than other methods. This method could also be used in geospatial interpolation studies, such as meteorological data interpolation.

INDEX TERMS 3D geological modeling, fault modeling, interpolation, ordinary kriging, neighborhood search.

I. INTRODUCTION

In subsurface studies, three-dimensional (3D) structural modeling helps geologists understand the physical world [1]–[3] and has been widely utilized in applications such as resource

The associate editor coordinating the review of this manuscript and approving it for publication was X. Huang.

management, disaster control, and tectonic evolution. Faults lead to the distributed damage of geological layers, and constructing faults in structural models play a vital role in mitigating risks in mining design. However, many works concentrate on the modeling of the physical properties of faults [4], [5], whereas the distributed damage of the surrounding geological layers draws less attention.

Constructing a structural model with layers and faults typically involves two steps. First, the surfaces of geological layers, i.e. the surfaces that represent the boundary between two rock layers, is modeled from the collected data. Second, the surface is cut and displaced by faults using a fault modeling method.

Traditionally, the modeling of surfaces relies on defining geological boundaries and linking cross sections to generate 3D surfaces manually [6], [7]. Automatic modeling methods based on extensive interpolators are then applied. These methods interpolate sample points extracted from horizontal observations of the geological surface to reconstruct a surface model, and they can be categorized into explicit and implicit approaches. Explicit methods, such as the inverse distance weighting interpolation method, and geostatistical methods (mainly kriging techniques; [8]–[11]) are utilized to explicitly calculate a continuous 3D surface by a function $z = f(x)$ that interpolates properties $Z = \{z_i, i = 0, 1, \dots, n\}$ of 2D sample points $P = \{x_i(x_i, y_i), i = 0, 1, \dots, n\}$ to get 3D points (x, y, z) on the surface. Alternatively, in the implicit surface reconstruction method, a geological surface is considered as an isovalue surface of a 3D scalar field $f(x, y, z)$ interpolated from the collected data. Commonly used implicit methods include discrete smooth interpolation [12], [13], and a set of radial basis functions (RBF) [14], [15]. RBF techniques, which have an extensive application in computer-aided design, interpolate a scalar field $f = f(x, y, z)$ based on the observations and then extract and visualize the zero isovalue surface ($f = 0$) of the scalar field. The Hermite RBF (HRBF) method [16] leveraging the data to be interpolated with second-order information (normal vectors). A compactly supported RBF method is given to decrease the interpolation time by constraining the neighborhood radius. Reference [17] and [18] developed a multiscale CSRBF approach using an Octree-based spatial index and normal vectors, advances in both time efficiency and interpolation accuracy.

Faults are then introduced to intersect and adjust the surface. Traditional fault modeling methods insert topological discontinuities by cutting and moving the smooth surface in a 3D Cartesian coordinate system [18], [19]. Then, implicit approaches are introduced to interpolate the fault surface into an implicit surface of a scalar field, then the regional model is modified in this scalar field [20]–[22]. In most of these methods, a fault is presented as topological discontinuity in a geological layer. First, fault geometries are computed by interpolating fault observations or estimated from seismic images based on semblance (e.g., [24], [25]). Second, fault surfaces are intersected with geological surfaces. Third, to keep the kinematic consistency of the surface, local movements of the intersection lines, namely, the fault cutoff lines, needs to be performed and propagated to the whole surface by manual [2], [26]. This fault modeling process increases the time required to build a valid model [27], so mathematical fault displacement modeling methods are suggested [28], [29], in which the slips on and around the fault geometry

are modeled by interpolation [30] or 3D parametric fault representation techniques [20], [31].

Although these methods are widely used in geological applications, an interpolated surface, which is initially constrained by all sample points and then cut and moved by faults, may no longer pass some of these controlled points. Fitting observations to get a geological surface that honors both horizontal and fault observations become difficult [32]. Optimization methods, such as numerical optimization and particle swarm optimization, can be used to move the faulted surface to honor the constraint of observations. The former finds an “optimal” set of parameters for trishear algorithms of fault modeling [33], [34], and the latter chooses kinematic parameters by minimizing a cost function consisting of the misfit between the data observed and the built surface model [32]. However, the complexity of the modeling workflow grows, as numerical calculation and kinematic interpretation is introduced by these methods, and, therefore, they should be implemented cautiously.

In contrast to the two-phase modeling process of these previous methods, we proposed a new workflow to reconstruct a geological surface with the presence of faults in a single processing step. As previous researches, this method also presents fault by topological discontinuity. However, it combines the fault modeling and interpolation processes and produces a geological surface honoring the constraint of faults and original sampling points. Geological modeling tasks are set to benefit from this method at two points: 1) no calculation required for the displacement of geological layer, and 2) the optimization methods do not need to be introduced, which will significantly improve the time efficiency of geological modeling. This method is achieved by leveraging a triangulated irregular network (TIN)-constrained ordinary kriging (OK) method. The OK method is used as the main interpolation method because of its effectiveness in geological applications and the simplicity of its mathematical form. The constrained Delaunay triangulations (CDT) algorithm [35] is performed to construct a TIN using the sample points and fault cutoff lines to constrain its neighborhood search process for locating neighbor points used in the interpolation. A workflow is then developed to construct the 3D models of faulted geological surfaces. The experimental results show that the proposed interpolation method improves both the accuracy and efficiency of the faulted geological surface model compared to the OK, HRBF, and multiscale CSRBF methods.

The remainder of this paper proceeds as follows. Section 2 highlights the related works in the literature. In Section 3, the framework of the 3D modeling of the faulted surface is described in detail. Section 4 describes the experimental results. Section 5 concludes the paper and discusses further research directions.

II. RELATED WORK

Kriging techniques provide the best linear unbiased estimation for un-sampled points as well as the ability to quantify

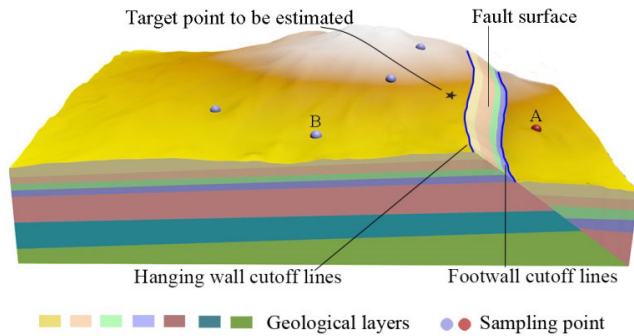


FIGURE 1. Sample points (red and purple) located on a faulted geological surface.

the uncertainty of estimates by error variance [36], [37], and they are frequently adopted to build geological layers [10].

To implement kriging, the neighborhood search method is needed to find nearby points of location for estimation. The distance-based neighborhood method, namely, identifying sample points based on their distance to the target location, is commonly implemented to accelerate the interpolation process [38], [39]. To improve the spatial balance of selected data, references [39] and [8] divided the space into quadrants around the target location and selected data within each (sub)quadrant based on distance. Other criterion for defining a neighborhood include K-nearest neighbors (KNN), which selects k closest points to a target point using a spatial data structure constructed from the whole sample dataset. The introduced spatial data structures in kriging studies include the k -dimensional tree [40]–[42] and the fast KNN search [43]. These structures provide storage and a fast index for the collected sample points and can avoid brute force computation of distances between the target point and points in the whole dataset.

However, these neighborhood search methods may not be optimal insofar as (i) they ignore the spatial topology among data, and only a geometric standpoint is applied to measure the closeness and redundancy of data [44]. The spatial topology, which defines containment, connectivity, and adjacency for spatial features, can be helpful in supporting better data selection. For example, a sample point (say point A in Fig. 1) that is geographically close to the target point for which the value is to be interpolated (star point in Fig. 1) but is located on the opposite side of the fault surface should be not be considered as a neighbor and should be excluded from the subsequent interpolation calculation. (ii) The brute force computation of distances between interpolated points and all sample points in the collected dataset is a very time-consuming procedure, particularly when the interpolation is conducted using big data [40].

Since a fault in a geological layer can be represented as a linear feature (for an unbounded fault terminating at the boundary of the modeling domain) or a polygonal feature (for a finite fault with a limited length in the modeling domain), another strategy, based on line intersection, is adopted according to [19]. For each sample point, a line segment is drawn to

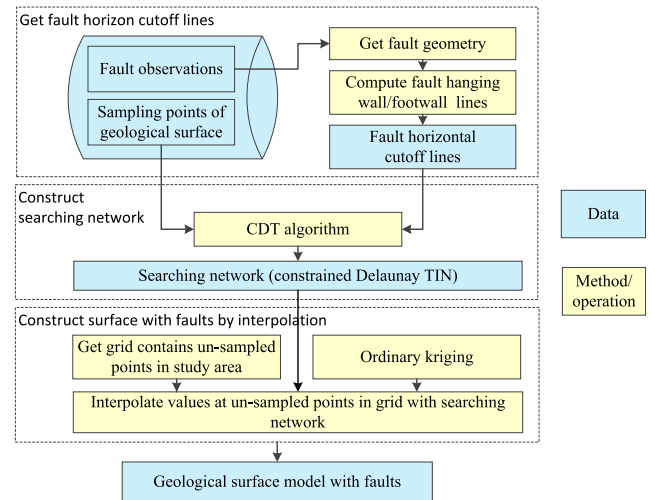


FIGURE 2. Workflow for 3D modeling of faulted geological surface.

connect it with the location for which its value is to be estimated, and it is verified whether this line segment intersects with the fault surface. Those verified sample points are then filtered from the result of the neighborhood search in each estimation. However, efficiency is a major concern in this approach, as the intersection operation in each estimation can be time consuming because the size of sample dataset, fault numbers, and estimation locations increases.

To address the aforementioned problems in the neighborhood search of the kriging method, our proposed TIN-constrained neighborhood search method uses a TIN to restrain the neighborhood search process. TIN comprises a series of vertices, which are the locations of sampled observations of the geological layer here, with 3D associated coordinates connected by edges to form a triangular tessellation. The CDT algorithm used in triangulation can force certain required segments (linear features) into the triangulation. Thus, by inserting a polygon consisting of fault cutoff lines into a TIN using the CDT algorithm and removing the internal triangles of this polygon, a fault can be represented as a hole bounded by fault cutoff lines in the TIN.

With the constraints of neighborhood defined in the TIN, the fault effect can be automatically incorporated during the interpolation process. Moreover, the spatial structure of the TIN stores adjacent triangles for every triangle in the TIN. As all vertices of triangles are sample points, the neighbors in triangles adjacent to the triangle containing the target point to be estimated can be reached in $O(1)$ time, which significantly reduces the amount of computation needed for neighborhood search and accelerates the entire interpolation process.

III. WORKFLOW FOR THREE-DIMENSIONAL (3D) MODELING OF FAULTED GEOLOGICAL SURFACE

A. METHOD OVERVIEW

There are four steps involved in the proposed 3D modeling of faulted geological surface workflow (Fig. 2). These are (i) preprocessing the collected data, i.e. the extraction of sample points, fault data, and boundary, (ii) computing

TABLE 1. Fault parameter used to build fault local coordinate system.

Parameter	Description
$\mathbf{O} (o_x, o_y, o_z)$	Center location of fault
v_s	Strike direction, specifies the orientation of the fault and measured clockwise from north
v_f	Normal direction, vertical to the fault surface at center point
v_d	Slip direction
d	Slip magnitude, tells us how far the rocks moved

fault geometries to obtain horizontal cutoff lines for each fault, (iii) interpolating sampled points with the constraint of fault horizontal cutoff lines, and (iv) building the geological surface within the geological boundary. Each step is listed below and described in detail later in this paper:

1) PREPROCESS THE COLLECTED DATA

Sample points of the geological surface can be extracted from archival geological data, such as drilling data, and geological maps, such as cross sections. Fault observations, including locations and displacement parameters, and boundary of the geological layer are also extracted from geological maps in this step.

2) GET HORIZONTAL CUTOFF LINES FOR EACH FAULT

The fault horizontal cutoff lines, defined as the hanging wall line and the footwall line for a given fault and its intersecting layer, are generated in this step by constructing a fault local coordinate system and calculating displacement along the fault surface.

3) IMPLEMENT INTERPOLATION

The TIN-constrained OK method is implemented in this step, as detailed in Section 3.3, with an input of the extracted geological sample points and computed fault horizontal cutoff lines and an output of a set of points located on the geological surface.

4) BUILD THE SURFACE WITHIN THE GEOLOGICAL BOUNDARY

After combining boundary and fault horizontal cutoff lines with estimated points within the boundary, faulted geological surface are then constructed by using the CDT algorithm based on these data.

B. GET FAULT HORIZONTAL CUTOFF LINES

In this work, we adopt three steps to calculate the horizontal cutoff lines for each fault: (i) compute a fault surface, (ii) build a local coordinate system for each fault, and (iii) calculate the horizontal cutoff lines.

In addition to the observed points on the fault surface, the parameters listed in Table 1 are also used to establish a fault surface and a local coordinate system, including the

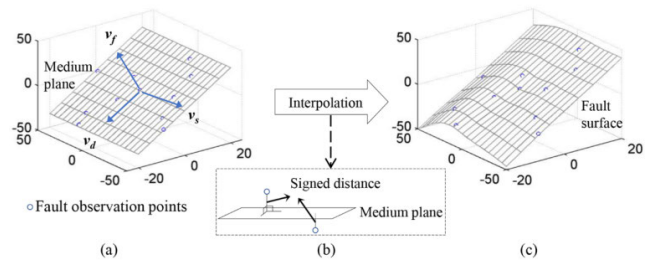


FIGURE 3. (a) Input fault observations and regular grid computed from medium plane aligned with v_s and v_d . (b) Signed distances between fault observation points and medium plane. (c) Fault geometry interpolated on the regular grid.

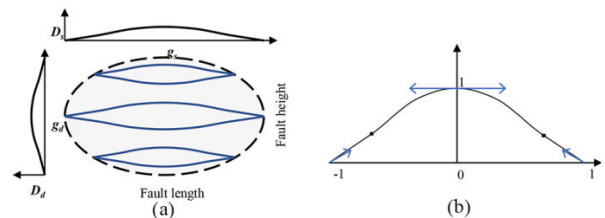


FIGURE 4. The slip associated with a theoretical isolated normal fault. (a) Slip on the fault surface, described using displacement profile D_s and D_d . (b) Theoretical model for D_s and D_d [28]. The values of the derivatives are controlled at the extremities and center of the profiles (arrows) and additional control points can be introduced in order to tune the curves (squares).

center position $\mathbf{O} (o_x, o_y, o_z)$ on the surface and the unit vectors $\{v_s, v_f, v_d\}$ [32]. By creating a fault medium plane (Fig. 3a), which is a reference plane aligned to v_s and v_d , signed distances can be defined as the distances between points and the plane (Fig. 3b). The geometry of a fault can then be computed (Fig. 3c) by interpolating these signed distances, and the fault local coordinate system \mathbf{G}_F can be established. The coordinates of a point $\mathbf{X} = \{x, y, z\}$ within $\mathbf{G}_F = \{g_s, g_f, g_d\}$ are computed using

$$\begin{pmatrix} g_s \\ g_f \\ g_d \end{pmatrix} = [v_s \quad v_f \quad v_d] \begin{pmatrix} x - o_x \\ y - o_y \\ z - o_z \end{pmatrix} + \begin{pmatrix} 0 \\ d(g_s, g_d) \\ 0 \end{pmatrix}$$

To acquire the fault cutoff lines, profiles indicating the relationship between the fault extent and amplitude are expected, which can be utilized to calculate the displacement for each point on the fault cutoff lines with fault characteristics. Basically, the amplitude of the fault displacement should be at its maximum at point \mathbf{O} and decreases to zero over a distance. Some other profiles are also available [19]. A parametric throw profile [20], using Hermite splines to present the attenuation followed by along-dip D_s and along-strike D_d (Fig. 4a), also can be used. Other profiles could be used instead, and more complex profiles could be obtained by mixing several simple profiles [45].

Fault horizontal cutoff lines can then be calculated along the fault surface in the local system and transformed to the coordinate system (usually Cartesian) where the sample

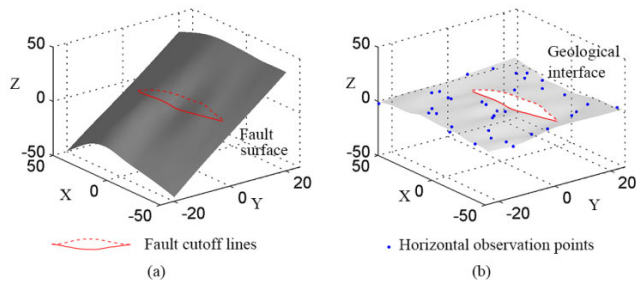


FIGURE 5. Data available to be interpolated. (a) Fault cutoff lines extracted from fault surface. (b) Sample points of geological surface and fault cutoff lines shown in a 3D space.

points are represented (Fig. 5a). Fault cutoff lines and horizontal observation points are then interpolated to reconstruct the geological surface. Fig. 5b illustrates these data and a possible geological surface implementation.

C. TRIANGULATED IRREGULAR NETWORK (TIN)-CONSTRAINED ORDINARY KRIGING (OK) METHOD

1) OK METHOD

Suppose $\{y(\mathbf{x}) \in A\}$ denotes the response variable of spatial position \mathbf{x} . Knowing that the value of an attribute at each sampled point of $X = \{\mathbf{x}_i, i = 1, 2, \dots, n\}$ is $y(\mathbf{x}_i)$, the estimated value $y(\mathbf{x}_0)$ at unsampled position \mathbf{x}_0 is the weighted sum of the values of X

$$\hat{y}(\mathbf{x}_0) = \sum_{i=1}^n \lambda_i y(\mathbf{x}_i), \tag{1}$$

where λ_i is the weight parameters corresponding to each sample point determined by variogram, which defines the variance of the difference between values at two locations, such as an exponential model

$$\gamma(h) = C_0 + C(1 - e^{-h/a}),$$

in which h is the distance between two sample points, a is the range, C_0 denote the nugget effect and C denote the partial sill. The last three parameters are fitted from an empirical variogram:

$$\hat{\gamma}(h + \delta h) = \frac{1}{2N(h + \delta h)} \sum_{i < j} (y(x_i) - y(x_j))^2, \tag{2}$$

$i, j = 1, 2, \dots, n$

where δ is a discretized metric called ‘‘lag’’ or ‘‘bandwidth tolerance,’’ and $N(h + \delta h)$ is the number of paired sample points.

Then, to estimate $\hat{y}(\mathbf{x}_0)$ at unsampled position \mathbf{x}_0 , weights in Eq. (1) can be calculated by

$$\begin{bmatrix} \gamma_{11} & \dots & \gamma_{1k} & 1 \\ \dots & \dots & \dots & \dots \\ \gamma_{k1} & \dots & \gamma_{kk} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_k \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma_{01} \\ \dots \\ \gamma_{0k} \\ 1 \end{bmatrix}, \tag{2}$$

where γ_{ij} denotes the semivariance value between \mathbf{x}_i and \mathbf{x}_j , and k is the number of sample points used in the estimation.

After weights are calculated, $\hat{y}(\mathbf{x}_0)$ can be calculated from Eq. (1).

However, as the volume of sample in the dataset grows, the dimensions of the matrices constructed by the dataset in Eq. (2) increases, which leads to a rapid increase in the time cost. The aforementioned distance-based or quadrant-distance-based neighborhood search methods (hereinafter, distance-based methods) are employed to limit the count k of sample points that participated in the estimation by selecting nearby neighbors.

The neighborhood search process follows Tobler’s first law of geography, which states, ‘‘Everything is related to everything else, but near things are more related than distant things’’ [46]. When the neighbor count is limited, the time required to solve Eq. (2) can be theoretically constant in each estimation process, which helps when applying OK interpolation to a large dataset.

2) TIN-BASED NEIGHBORHOOD SEARCH METHOD

We explored the TIN-based neighborhood search method to tackle two challenges—the utilizing of spatial topology among data and the time-consuming nature of distance computations mentioned in Section 2—by constraining and accelerating the neighborhood search process of the kriging method. In our method, the neighborhood search process would benefit from the structure of the TIN, not only to applying fault’s effect on neighborhood selection but also to accelerate this process by accessing adjacent triangles in the TIN.

a: CONSTRUCT TIN-BASED NEIGHBORHOOD SEARCH NETWORK

A TIN-based neighborhood search network, constructed by a CDT algorithm using sample points and fault cutoff lines, is defined to store the spatial relationship and topology among these data. A Poly2Tri library (<http://code.google.com/p/poly2tri/>) is used to the implement CDT algorithm, and the steps to establish the search network are as follows:

–First, project sample points and fault cutoff lines of the surface to a plane. This is because the Poly2Tri library was developed for a 2D data set.

–Second, construct a triangulation network within the geological boundary by applying the CDT algorithm to these data.

The hanging wall and footwall of a reverse fault overlaps when the data is projected to 2D space (Fig. 6a). However, the CDT algorithm cannot generate overlapping areas—for example, Fig. 6b illustrates a TIN constructed by the CDT algorithm using the sample points and cutoff lines of the fault shown in Fig. 6a. Here, a non-overlapping discontinuous area, bounded by fault cutoff lines, is created, and this appears to be not in line with Fig. 6a. So, to model a reverse fault in the TIN and keep a correct topology for sample points, lines linked to a point in the hanging wall line (brown) and a corresponding point in the footwall line (green) are reconnected (Fig. 6c).

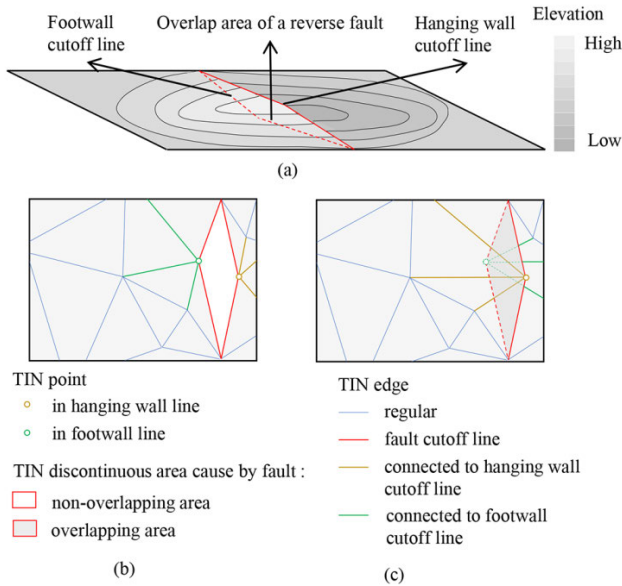


FIGURE 6. Insert a reverse fault into a TIN-based search network. (a) A reverse fault. (b) Fault incorrectly constructed in TIN by CDT algorithm. (c) After reconnecting lines linked to fault cutoff lines, reverse fault is constructed correctly.

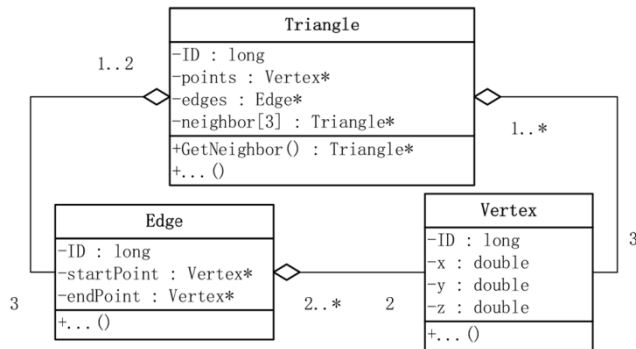


FIGURE 7. The UML class diagram for the TIN data structure.

b: NEIGHBORHOOD SEARCH

Since each vertex in the TIN-based search network denotes a sample point or a node on the fault cutoff lines, the neighborhood search process can be executed by utilizing the structure of TIN.

Fig. 7 shows the basic geometry components and their relationship in a constrained Delaunay TIN. All elements are categorized into three classes—*Vertex*, *Edge*, and *Triangle*. An *Edge* object serves as one of the three edges comprised in a *Triangle* object and connects two sample points storing their locations in *Vertex* objects. Each *Triangle* object also keeps its adjacent triangles, which can be accessed by the member function of *GetNeighbors()*.

Then, to choose neighbor points in the search network to estimate the value at target location x , two steps are expected. The first step is to locate the triangle containing x , and this is achieved by a two-phase procedure of selecting triangles whose enclosing rectangle contains x , and locating the triangle T_0 containing x from these triangles.

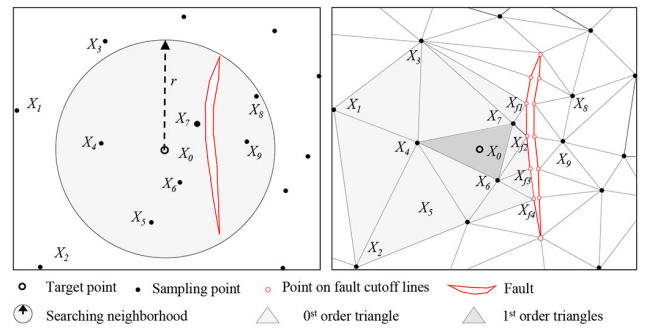


FIGURE 8. Neighborhood search method (a) within a certain distance r and (b) along TIN-based search network.

The order of the triangle, which measures the connectivity from x to each triangle in the TIN, is defined in the following step. The 0^{th} order triangle denotes T_0 confirmed in the first step, and the t^{th} order triangles ($t > 0$) are given the definition of triangles that shares vertices or edges with $(t-1)^{th}$ order triangles. Then, by specifying the order of triangle t , all the vertices in $0 \sim t$ order triangles will be selected and used in the estimation.

To define the difference between the TIN-based neighborhood method and the distance-based neighborhood method, we should consider the following scenario. Suppose we select neighbor points from dataset $X = \{X_i(x_i, Z_i(x_i)), i = 1, 2, \dots, n\}$ to estimate the value $Z(x_0)$ at point X_0 . Using the distance-based neighborhood search, all sample points are sorted according to their distance to X_0 , and the six nearest points, i.e. $X_4 \sim X_9$, were chosen (Fig. 8a). Even when the local surface is divided into two sides by fault f , x_8 and x_9 are apparently located at the opposite side of X_0 of the fault.

As a comparison, in our TIN-based neighborhood search method, all the vertices of $0 \sim t$ order triangles were selected in this process and subsequently interpolated. In Fig. 8, suppose $t = 1$, vertices in 0^{th} order (sample points X_4, X_6 , and X_7) and 1^{st} order triangles (sample points X_1, X_2, X_3 , and X_5 and fault cutoff line points $X_{f1} \sim X_{f4}$) were selected (Fig. 8b). All of these neighborhood points are located in a continuously changing region where X_0 is set and are obviously more meaningful for estimating the value at X_0 .

Compared with the other KNN methods mentioned in Section 2, only the TIN-based neighborhood search concerns the constraints of faults. By inserting faults in the TIN and using the TIN as a search network, this method can select sample points that are not only spatially close but also highly topologically correlated to the target point of estimation, which ensures that the prediction of the OK method is produced from a continuous domain.

IV. EXPERIMENTS

A. MODELING OF SIMULATED DATASET

To evaluate the performance of the proposed TIN-constrained OK method, a surface model of a geological surface S was simulated from a process $z = s(x, y)$, where $x, y \in [-50, 50]$

TABLE 2. Parameters of the simulated fault.

Center point	Strike direction	Dip angle (°)	Slip direction	Slip magnitude (m)
(0, 0, 0)	(0, 1, 0)	60	$(1/2, 0, -\sqrt{3}/2)$	10

m and $z = 0$ m. A finite normal fault f was then simulated to intersect, cut, and deform S , with parameters listed in Table 2.

A data set $P = \{p_i(x_i, y_i, z_i), i = 1, 2, \dots, 100\}$ including 100 points was randomly sampled from the faulted surface, and then normal vectors of the surface corresponding to each sample point were calculated. The simulated surface, fault, and sampled data set are shown in Fig. 8a.

In our experiments, OK interpolation methods using two distance-based neighborhood search methods were employed to compare the performance in both accuracy and efficiency with TIN-constrained OK. Two implicit methods using second-order information, including HRBF, and multiscale CSRBF methods, were also added to the comparison. By leveraging the gradient constraints, the latter two methods provide better interpolation results than the original RBF method.

For readability, the used methods are abbreviated as TIN-OK, D-OK (OK method using distance-based neighborhood), Q-OK (OK method using quadrant-distance-based neighborhood), HRBF, and multi-CSRBF (multiscale CSRBF).

1) INTERPOLATION RESULTS OF SIMULATED DATASET

In the first scenario, four interpolators were applied to the simulated dataset to estimate z values for unsampled locations to form surface models drawn in Fig. 9 b, c, d, and e, respectively. All surfaces are rendered within the same color range from light purple to red (corresponding to the value range $[-5\sqrt{3}, 5\sqrt{3}]$ m of z), and the dark areas shown in the figures indicate that the estimated values exceed this value range.

From Fig. 9, the values predicted from the D-OK and Q-OK methods in the fault displacement zone are narrowed because they chose sample points from both sides of fault when estimating values of locations near fault. The surfaces modeled from HRBF and multi-CSRBF change sharply near the fault when additional normal data are used (Fig. 9e). In line with expectations, our proposed TIN-OK performs better than the other methods with a surface model (Fig. 9b) that very similar to the simulated “ground truth” surface (Fig. 9a).

A further experiment was conducted to compare the performance of the TIN-OK and D-OK method on fault modeling. For our proposed TIN-OK method, the fault can be constructed with the constraint of fault horizontal cutoff lines in TIN; for the D-OK method, to construct fault with OK interpolated surface, a fault modeling method using a fault

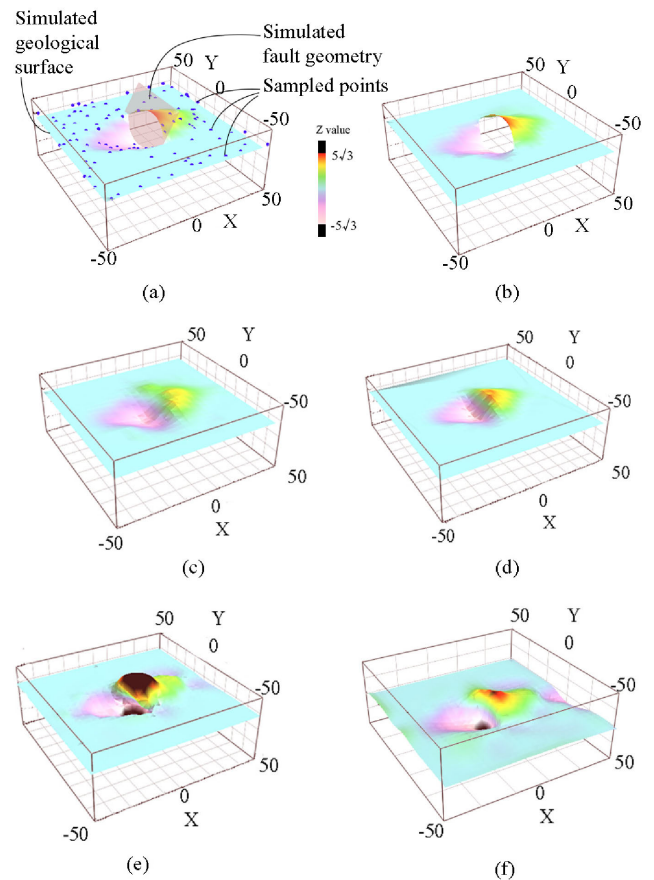


FIGURE 9. (a) Simulated geological surface and fault, and interpolation results from (b) TIN-OK, (c) D-OK, (d) Q-OK, (e) HRBF, and (f) Multi-CSRBF methods.

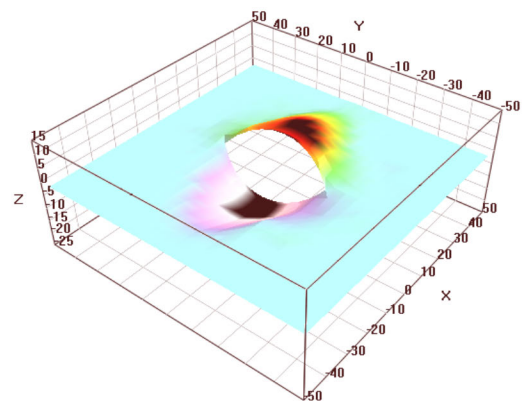


FIGURE 10. Constructing fault on the surface interpolated by D-OK method.

local frame was implemented based on the work presented in reference [32] and detailed in section III. B.

Fig. 10 depicts the fault modeling result using an additional fault modeling method after the D-OK interpolation. Compare to the result illustrated in Fig. 9(b), the surface in Fig. 10 appears some dark areas within the fault displacement zone. Because the deformation of the surface in the fault modeling

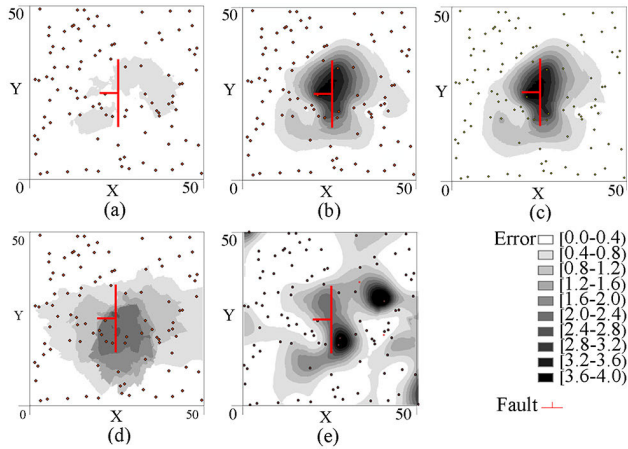


FIGURE 11. Distribution of errors produced by ten cross-validation results of (a) the proposed TIN-OK, (b) D-OK, (c) Q-OK, (d) HRBF, and (e) Multi-CSRBF methods.

process caused a misfit of sample points. Although it can be corrected by implementing optimization methods mentioned in Section I, the increasing time consumption is another concern. The same conclusion can be derived from the Q-OK, HRBF, and multi-CSRBF methods, since they also have no built-in methods to handle fault structures.

2) COMPARISON OF INTERPOLATION ACCURACY

A K-fold cross-validation (with K equal to ten here) was subsequently implemented to compare the accuracy of these interpolators. In each cross-validation, sample points were divided into K subsets, of which nine were interpolated, and the remaining one was used to acquire errors to validate. This process ends after all subsets have served as a validation set. The interpolation error at point (x_j, y_j) is defined as an unsigned value of difference between the estimated value z'_j and the z_j .

This K-fold cross-validation operated ten times for each method in our experiment. The average errors from ten cross-validations were then interpolated and are illustrated in Fig. 11. The darker the color on the map, the greater the error in cross-validation.

In all four maps, the error increases when the sample point approaches the center of fault. The results from D-OK and Q-OK interpolation have larger dark areas (bigger errors), and the errors reach a maximum value of 4.55 m for the D-OK method (Fig. 11b) and 4.53 m for the Q-OK method (Fig. 11c). However, errors in most areas in the results from the proposed TIN-OK interpolation were below 0.40, with a maximum value of 1.52 (Fig. 11a). HRBF has better accuracy than D-OK and Q-OK, but it exceeded others in its error area, the same as a maximum value of 9.17 m (Fig. 11d). The multi-CSRBF approach also produced large errors in the center and two of four corners (Fig. 11e), maximally 7.19 m, which may be caused by a sharp change in the normal vectors near the fault. Thus, the proposed TIN-OK showed the highest interpolation accuracy compared to other approaches.

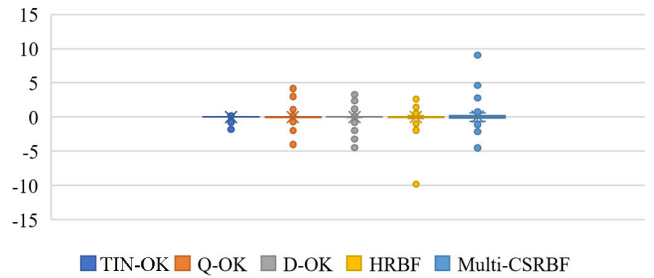


FIGURE 12. Boxplot of errors from one cross-validation result for previously tested methods.

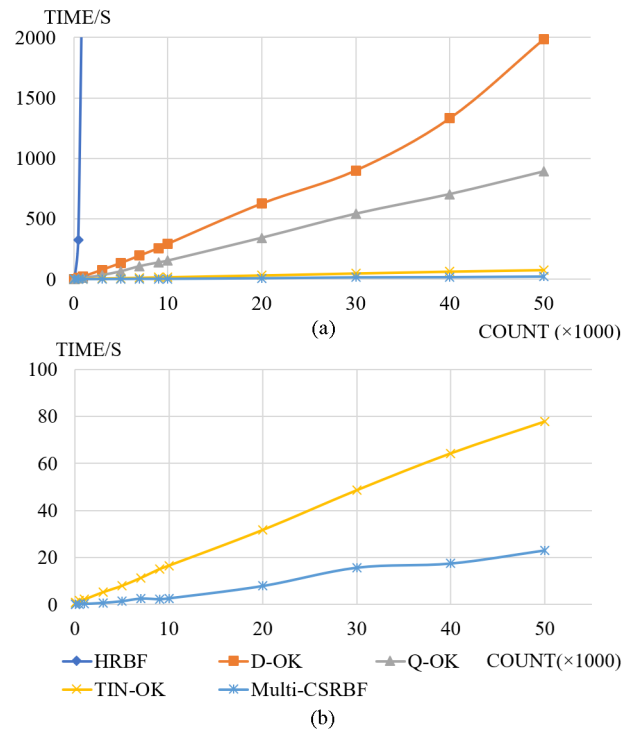


FIGURE 13. Time consumption changes as point size increases in testing methods.

One of ten K-fold cross-validation results is illustrated in the boxplot in Fig. 12. The result shows the same trend illustrated in Fig. 11, i.e. the introducing of TIN improves the accuracy of OK method when interpolating faulted surface.

3) COMPARISON OF INTERPOLATION TIME

The time efficiency of the four methods was compared using simulated datasets with different sizes (50, 100, 500, 1000, 5000, 10000, 20000, 30000, 40000, and 50000) on a laptop with an Intel^R CoreTM i5-8250U processor clocked at 1.60 GHz and 8 GB of random-access memory. Fig. 13a shows the time consumption trend of the four methods when the size of the simulated dataset increases. Limiting the count of neighbors, selected by the different neighborhood methods, that participate in each estimation is expected to linearly increase time in all kriging methods. The slope of the line

produced by the TIN-OK method is smaller than the other two OK methods with a reduction in the time used of 96.1% (D-OK) and 91.3% (Q-OK) when the count of simulated points reaches fifty thousand. In terms of the other methods, the HRBF method uses all points in each estimation, which means that the computational cost required to solve the system of RBF equations increases cubically and the laptop used in this experiment cannot implement interpolation with more than one thousand points; the multi-CSRBF method, in the opposite situation, runs even more efficiently than all of the TIN-OK method benefits from the Octree indexing and multi-scale interpolations (Fig. 13b).

Table 3 summarizes the time complexity of the four comparison methods. The whole interpolation process consists of three parts in our time consumption analysis: constructing a sample point index for the neighborhood search, interpolating, and visualizing. In the first part, the TIN-OK and multi-CSRBF methods have a worst-case time complexity of $O(n \log n)$ for operating the CDT and Octree construction algorithms, respectively, for the whole dataset whereas D-OK, Q-OK, and HRBF cost no time in this stage.

In the interpolating stage, for the three OK methods (TIN-OK, D-OK, and Q-OK), each estimation in interpolation is calculated after indexing k neighbors and making the prediction. TIN-OK has an $O(n)$ time complexity to locate a triangle, as we mentioned in Section 3.3.2.2; the D-OK and Q-OK methods sort sample points based on distance (in each sector for the Q-OK method) and create an index dynamically to find the k nearest neighbors in each estimation, which produces a worst-case time complexity of $O(n)$ and $O(n \log n)$, respectively. All of these methods index neighbors with $O(1)$ time complexity, except the HRBF. Then, to solve the linear equation constructed in the OK methods and get an estimate from selected neighbors, the $O(k^3)$ time is taken; the HRBF reaches $O((n + 3n)^3)$ complexity to calculate the isovalue surface in this process; whereas the multi-CSRBF, using a sparse matrix benefit from the compactly supported radial basis neighborhood search, can approach $O(n^2)$ time complexity in this stage.

The HRBF and multi-CSRBF also requires a visualization algorithm to explicitly represent the isovalue surface, such as the marching cubes algorithm with a worst-case time complexity of $O(q)$, where q depends on the spatial granularity of the visualized surface.

In summary, in terms of time efficiency, although the TIN-OK method costs extra time in constructing the TIN-based search network, this method accelerates the neighbor selection process and leads to an improvement in overall time efficiency. Furthermore, fault displacement and optimization methods are needed after geological surfaces are interpolated from the D-OK, Q-OK, HRBF, and multi-CSRBF methods in a geological surface modeling workflow, which makes the modeling workflow even more complicated and may cause unpredictable time expenditure. As a comparison, our proposed method has no such steps because fault displacements have been considered in the interpolation step.

TABLE 3. Time consumption of four methods in experiment.

	Construct index	Visualize
TIN-OK	$O(n \log n)$	-
D-OK	-	-
Q-OK	-	-
HRBF	-	$O(q)$
Multi-CSRBF	$O(n \log n)$	$O(q)$
Interpolate		
	Search neighbors	Solve linear equation
TIN-OK	Locating: $O(n)$, and indexing: $O(1)$	$O(k^3)$
D-OK	Get distance: $O(n)$, sort: $O(n \log n)$, and indexing: $O(1)$	$O(k^3)$
Q-OK	Get distance: $O(n)$, sort: $O(n \log n)$, and indexing: $O(1)$	$O(k^3)$
HRBF	-	$O((n+3n)^3)$
Multi-CSRBF	Indexing: $O(\log n)$	$O(n^2)$

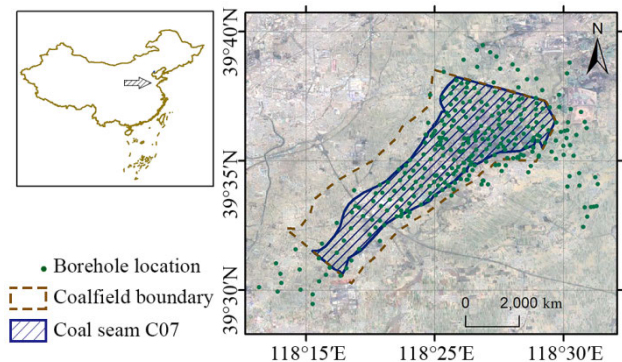


FIGURE 14. The geographical location and boundary of Qianjiaying coal mine, and borehole distribution.

B. CASE STUDY

The Qianjiaying coal mine is located approximately 15 kilometers southeast of Tangshan, China, with an area of 88 square kilometers (Fig. 14). The strata in this mine have a smaller dip angle (generally $10^\circ \sim 15^\circ$) and simpler construction. The coal-bearing strata of this field belong to the Upper Carboniferous and the Lower Permian, and the basement strata are the limestone of the Middle Ordovician Majiagou Formation. The total thickness of the coal-bearing strata is about 500 meters, containing more than ten layers of coal with a total thickness of 19.79 meters and a coal content of 3.96%.

Five horizontal coal seams (C05, C07, C08, C09, and C12a) and 26 faults are expected to be modeled in this area. Available data sources are listed in Table 4, including boreholes drilling, underground sketch mapping, and fault observation.

TABLE 4. Parameters of the study area.

Data source	Data Amount
Borehole	259
Underground borehole	57
Geological sketch map (extracted points)	985
Fault observation	26

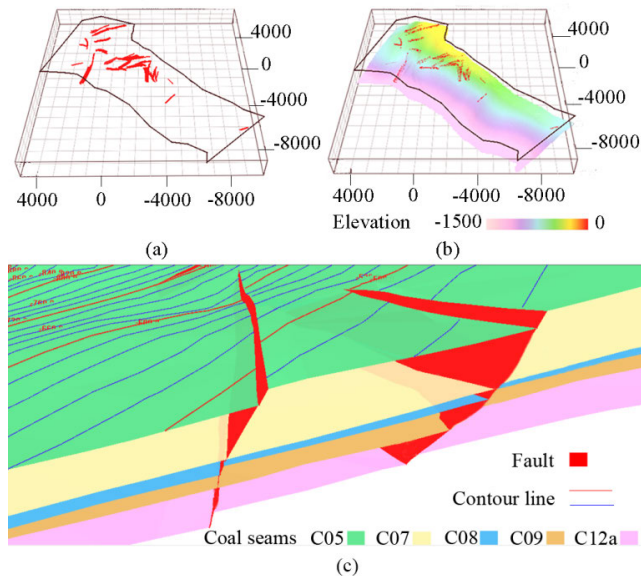


FIGURE 15. Modeling of a faulted area in the Qianjiaying coal mine. (a) Fault surface models. (b) Faulted coal seam models. (c) Detailed models of fault and seams.

Based on the proposed workflow, fault surfaces were modeled (Fig. 15a), and their horizontal cutoff lines were computed for each intersecting fault and coal seam. Then, the sample points extracted from drills and underground surveys and fault horizontal cutoff lines were interpolated using the proposed method. Finally, all faulted coal seams models were reconstructed (Fig. 15b). Fig. 15c illustrated detailed models of two faults intersecting coal seams.

V. CONCLUSION

In this paper, a TIN-constrained OK method is proposed to build the faulted surfaces in geological context. To the best of our knowledge, this is a pioneering attempt to apply TIN to kriging as a neighborhood search constraint. Compared to previous methods, this method makes three important methodological contributions:

(1) A search network based on TIN is developed, and neighborhood search is conducted along this network, which ensures that each estimation uses sample points from a stable and continuous area.

(2) The proposed method performs better in both accuracy and efficiency than the OK method using distance-based neighborhood search approaches and the HRBF method in the geological modeling of the faulted surface, particularly regarding near faults.

(3) The fault displacement and the optimization steps after interpolation are simplified in our method, which may also improve the time efficiency of modeling.

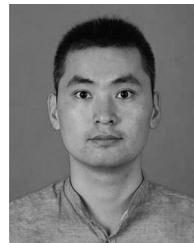
We provide a simplified modeling workflow based on the proposed TIN-constrained OK method without introducing the complex kinematics and mechanics of faults. The accuracy of modeling improves as the available data increases, and time efficiency exceeds two OK methods using the distance-based neighborhood method, and two implicit approaches (HRBF and multi-CSRBF).

However, because the fault displacement direction is not always parallel to the slip direction at the fault center, the method of computing cutoff lines does not always apply to all fault types [31]. In the future, other fault modeling methods will be introduced to address this problem [19]. Moreover, as an explicit modeling method, the proposed method present elevation value (z) in terms of x and y coordinates, which means it cannot be applied to present the overfolding or erosion of a surface such as an overturned fold or salt surfaces, etc. A possible solution is to project such structure to a medium plane aligned with the slope direction of such surface [31], or by using a central cylindrical projection.

REFERENCES

- [1] P. Thore, A. Shtuka, M. Lecour, T. A. Ettajer, and R. Cognot, "Structural uncertainties: Determination, management, and applications," *Geophysics*, vol. 67, no. 3, pp. 840–852, May 2002.
- [2] G. Caumon, P. Collon-Drouaillet, C. Le Carlier de Veslud, S. Viseur, and J. Sausse, "Surface-based 3D modeling of geological structures," *Math. Geosci.*, vol. 41, no. 8, pp. 927–945, Nov. 2009.
- [3] M. J. Hillier, E. M. Schetselaar, E. A. de Kemp, and G. Perron, "Three-dimensional modelling of geological surfaces using generalized interpolation with radial basis functions," *Math. Geosci.*, vol. 46, no. 8, pp. 931–953, Nov. 2014.
- [4] S. J. Jolley, H. Dijk, J. H. Lamens, Q. J. Fisher, T. Manzocchi, H. Eikmans, and Y. Huang, "Faulting and fault sealing in production simulation models: Brent province, northern north sea," *Petroleum Geosci.*, vol. 13, no. 4, pp. 321–340, Nov. 2007.
- [5] T. Manzocchi, C. Childs, and J. J. Walsh, "Faults and fault properties in hydrocarbon flow models," *Frontiers Geofluids*, vol. 10, pp. 94–113, Dec. 2010.
- [6] M. H. Herbert, C. B. Jones, and D. S. Tudhope, "Three-dimensional reconstruction of geoscientific objects from serial sections," *Vis. Comput.*, vol. 11, no. 7, pp. 343–359, Jul. 1995.
- [7] J. Ming, M. Pan, H. Qu, and Z. Ge, "GIS: A 3D geological multi-body modeling system from netty cross-sections with topology," *Comput. Geosci.*, vol. 36, no. 6, pp. 756–767, Jun. 2010.
- [8] E. Isaaks, R. Srivastava, *An Introduction to Applied Geostatistics*. New York, NY, USA: Oxford Univ. Press, 1989.
- [9] J. B. Boisvert and C. V. Deutsch, "Programs for kriging and sequential Gaussian simulation with locally varying anisotropy using non-Euclidean distances," *Comput. Geosci.*, vol. 37, no. 4, pp. 495–510, Apr. 2011.
- [10] X. Li, P. Li, and H. Zhu, "Coal seam surface modeling and updating with multi-source data integration using Bayesian geostatistics," *Eng. Geol.*, vol. 164, pp. 208–221, Sep. 2013.
- [11] I. Hassen, H. Gibson, F. Hamzaoui-Azaza, F. Negro, K. Rachid, and R. Bouhlila, "3D geological modeling of the kasserine aquifer system, central tunisia: New insights into aquifer-geometry and interconnections for a better assessment of groundwater resources," *J. Hydrol.*, vol. 539, pp. 223–236, Aug. 2016.
- [12] J. L. Mallet, "Discrete Smooth Interpolation in geometric modelling," *ACM Trans. Graph.*, vol. 8, no. 2, pp. 121–144, 1989.
- [13] P. Collon, W. Steckiewicz-Laurent, J. Pellerin, G. Laurent, G. Caumon, G. Reichart, and L. Vaute, "3D geomodelling combining implicit surfaces and Voronoi-based remeshing: A case study in the lorraine coal basin (France)," *Comput. Geosci.*, vol. 77, pp. 29–43, Apr. 2015.

- [14] G. Turk and J. F. O'Brien, "Shape transformation using variational implicit functions," in *Proc. ACM SIGGRAPH Courses SIGGRAPH*, 2005, pp. 335–342.
- [15] S. A. Vollgger, A. R. Cruden, L. Ailleres, and E. J. Cowan, "Regional dome evolution and its control on ore-grade distribution: Insights from 3D implicit modelling of the navachab gold deposit, Namibia," *Ore Geol. Rev.*, vol. 69, pp. 268–284, Sep. 2015.
- [16] Y.-H. Tseng and S. Wang, "Semiautomated building extraction based on CSG model-image fitting," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 2, pp. 171–180, Feb. 2003.
- [17] Y. Ohtake, A. Belyaev, and H. P. Seidel, "A multi-scale approach to 3D scattered data interpolation with compactly supported basis functions," in *Proc. Shape Modeling Int.*, May 2003, pp. 153–161.
- [18] Q. Jia, D. Che, and W. Li, "Effective coal seam surface modeling with an improved anisotropy-based, multiscale interpolation method," *Comput. Geosci.*, vol. 124, pp. 72–84, Mar. 2019.
- [19] Q. Wu, H. Xu, X. Zou, and H. Lei, "A 3D modeling approach to complex faults with multi-source data," *Comput. Geosci.*, vol. 77, pp. 126–137, Apr. 2015.
- [20] G. Laurent, G. Caumon, A. Bouziat, and M. Jessell, "A parametric method to model 3D displacements around faults with volumetric vector fields," *Tectonophysics*, vol. 590, pp. 83–93, Apr. 2013.
- [21] G. Caumon, G. Gray, C. Antoine, and M.-O. Titeux, "Three-dimensional implicit stratigraphic model building from remote sensing data on tetrahedral meshes: Theory and application to a regional model of la popa basin, NE Mexico," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1613–1621, Mar. 2013.
- [22] P. Calcagno, J. P. Chilès, G. Courrioux, and A. Guillen, "Geological modelling from field data and geological knowledge—Part I. Modelling method coupling 3D potential-field interpolation and geological rules," *Phys. Earth Planet. Interiors*, vol. 171, nos. 1–4, pp. 147–157, 2008.
- [23] G. Machado, A. Alali, B. Hutchinson, O. Olorunsola, and K. J. Marfurt, "Display and enhancement of volumetric fault images," *Interpretation*, vol. 4, no. 1, pp. SB51–SB61, Feb. 2016.
- [24] J. Qi, G. Machado, and K. Marfurt, "A workflow to skeletonize faults and stratigraphic features," *Geophysics*, vol. 82, no. 4, pp. O57–O70, Jul. 2017.
- [25] X. Wu and Z. Zhu, "Methods to enhance seismic faults and construct fault surfaces," *Comput. Geosci.*, vol. 107, pp. 37–48, Oct. 2017.
- [26] J. Mallet, *Geomodeling*. London, U.K.: Oxford Univ. Press, 2002.
- [27] G. Caumon, G. Laurent, J. Pellerin, N. Cherpeau, F. Lallier, R. Merland, and F. Bonneau, "Closing the gap: Advances in applied geomodeling for hydrocarbon reservoirs," *Can. Soc. Petroleum Geologists*, vol. 20, pp. 43–52, Dec. 2013.
- [28] J. A. M. Barnett, J. Mortimer, J. H. Rippon, J. J. Walsh, and J. Watterson, "Displacement geometry in the volume containing a single normal fault," *AAPG Bull.*, vol. 71, no. 8, pp. 925–937, 1987.
- [29] B. Freeman, G. Yielding, and M. Badley, "Fault correlation during seismic interpretation," *First Break*, vol. 8, no. 1229, Mar. 1990.
- [30] F. Georgsen, P. Røe, A. R. Syversveen, and O. Lia, "Fault displacement modelling using 3D vector fields," *Comput. Geosci.*, vol. 16, no. 2, pp. 247–259, Mar. 2012.
- [31] N. Cherpeau, G. Caumon, and B. Lévy, "Simulations stochastiques de réseaux de failles en modélisation structurale 3D," *Comp. Rendus Geosci.*, vol. 342, no. 9, pp. 687–694, 2010.
- [32] G. Godefroy, G. Caumon, M. Ford, G. Laurent, and C. A.-L. Jackson, "A parametric fault displacement model to introduce kinematic control into modeling faults from sparse data," *Interpretation*, vol. 6, no. 2, pp. B1–B13, May 2018.
- [33] R. W. Allmendinger, "Inverse and forward numerical modeling of trishear fault-propagation folds," *Tectonics*, vol. 17, no. 4, pp. 640–656, Aug. 1998.
- [34] N. Cardozo and J. P. Brandenburg, "Kinematic modeling of folding above listric propagating thrusts," *J. Struct. Geol.*, vol. 60, pp. 1–12, Mar. 2014.
- [35] L. P. Chew, "Constrained delaunay triangulations," in *Proc. 3rd Annu. Symp. Comput. Geom. SCG*, 1987, pp. 215–222, 1987.
- [36] L. Tacher, I. Pomian-Szrednicki, and A. Parriaux, "Geological uncertainties associated with 3-D subsurface models," *Comput. Geosci.*, vol. 32, no. 2, pp. 212–221, Mar. 2006.
- [37] P. Thierry, A.-M. Prunier-Leparmier, C. Lembezat, E. Vanoudheusden, and J.-F. Vernoux, "3D geological modelling at urban scale and mapping of ground movement susceptibility from gypsum dissolution: The paris example (France)," *Eng. Geol.*, vol. 105, nos. 1–2, pp. 51–64, Apr. 2009.
- [38] N. Cressie, *Statistics for Spatial Data*. New York, NY, USA: Wiley, 1993.
- [39] J. Chilès and P. Delfiner, *Geostatistics: Modeling Spatial Uncertainty*. New York, NY, USA: Wiley, 1999.
- [40] L. Li, T. Lossner, C. Yorke, and R. Piltner, "Fast inverse distance weighting-based spatiotemporal interpolation: A Web-based application of interpolating daily fine particulate matter PM_{2.5} in the contiguous U.S. using parallel programming and k-d tree," *Int. J. Environ. Res. Public Health*, vol. 11, no. 9, pp. 9101–9141, 2014.
- [41] W. Tong, L. Li, X. Zhou, and J. Franklin, "Efficient spatiotemporal interpolation with spark machine learning," *Earth Sci. Informat.*, vol. 12, no. 1, pp. 87–96, Mar. 2019.
- [42] Y. Zhang, X. Zheng, Z. Wang, G. Ai, and Q. Huang, "Implementation of a parallel GPU-based space-time kriging framework," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 5, p. 193, 2018.
- [43] G. Mei, N. Xu, and L. Xu, "Improving GPU-accelerated adaptive IDW interpolation algorithm using fast kNN search," *SpringerPlus*, vol. 5, no. 1, p. 1389, Dec. 2016.
- [44] X. Emery, "The kriging update equations and their application to the selection of neighboring data," *Comput. Geosci.*, vol. 13, no. 3, pp. 269–280, Sep. 2009.
- [45] W. K. HAMBLLIN, "Origin of 'reverse drag' on the Downthrown side of normal faults," Ph.D. dissertation, Dept. Geol., Brigham Young Univ., Prom, Utah, 1965, pp. 1145–1164.
- [46] W. Tobler, "A computer movie simulating urban growth in the detroit region," *J. Chem. Inf. Model.*, vol. 46, pp. 234–240, Jun. 1970.



QINGREN JIA born in Rizhao, China, in 1990. He received the B.S. degree in surveying engineering and the M.S. degree in geographic information engineering from Northeastern University, China, in 2013 and 2015, respectively, where he is currently pursuing the Ph.D. degree in digital mining.

He has also been a Visiting Student with the School of Geography and Urban Planning, Arizona State University, since 2018. His main research interests are three-dimensional geological modeling, spatial interpolation, and spatio-temporal data analysis.



WENWEN LI received the B.S. degree in computer science from Beijing Normal University, Beijing, China, in 2004, the M.S. degree in signal and information processing from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, in 2007, and the Ph.D. degree in earth system and geoinformation science from George Mason University, VA, USA, in 2010.

She is currently an Associate Professor with Arizona State University, AZ, USA. Her research interest is geographic information science with a focus on cyberinfrastructure, big data, semantic interoperability, spatial information retrieval, and distributed geospatial information processing.



DEFU CHE born in Yantai, China, in 1970. He received the Ph.D. degree in geographic information engineering from the China University of Mining and Technology, Beijing, in 2006.

From 1994 to 2001, he worked at Gold Mine, Shandong, China, as a Survey Engineer. From 2007 to 2010, he was an Associate Professor with the School of Resources and Civil Engineering, Northeastern University, China. Since 2011, he has been a Professor with Northeastern University.

He is also a member of the Professional Committee of the International Society for Mine surveying (ISM) and a member of the Mine Measurement Committee of the China Coal Society. His main research interests are spatial geometry algorithms in the geographic information systems, digital mine, digital cities, three-dimensional geoscience modeling, and software development.

...