



Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo



Highlights

Knowledge-driven stochastic modeling of geological geometry features conditioned on drillholes and outcrop contacts

Xiaolong Wei*, Zhen Yin, Wilson Bonner, Jef Caers

Computers and Geosciences xxx (xxxx) xxx

- We develop a method to incorporate geological knowledge, represented by a 2D geological diagram, into 3D stochastic modeling.
- The constructed 3D magmatic intrusion models exhibit consistency with geological knowledge, drillhole data, and outcrop contacts.
- We quantify the uncertainty of geometric characteristics for intrusion models.

Graphical abstract and Research highlights will be displayed in online search result lists, the online contents list and the online article, but **will not appear in the article PDF file or print** unless it is mentioned in the journal specific style requirement. They are displayed in the proof pdf for review purpose only.



Research paper

Knowledge-driven stochastic modeling of geological geometry features conditioned on drillholes and outcrop contacts

Xiaolong Wei^{a,*}, Zhen Yin^a, Wilson Bonner^b, Jef Caers^a

^a Department of Earth and Planetary Sciences, Stanford University, CA, USA

^b Kobold Metals, CA, USA

ARTICLE INFO

Keywords:

Knowledge-driven modeling
Geometry characteristics
Monte Carlo sampling
Uncertainty quantification

ABSTRACT

Geometry features in geosciences are crucial for understanding both near-surface and deeper Earth's interior structures. Geometry features have traditionally been delineated through diverse data sources such as surface mapping, geophysics, and core samples. However, the quantitative integration of geological knowledge and insights with the geoscientific data remains insufficiently addressed in the model construction. We have formulated a novel framework that employs stochastic level set simulation to model subsurface geometric features. The uniqueness of our framework involves geological knowledge, represented by two-dimensional (2D) geological diagrams, in the numerical modeling using Procrustes analysis. We account for the geological diagrams' intrinsic variability and uncertainty through transformations such as rotation, scaling, and translation. By using the designed loss functions, the resulting models align with the established geological knowledge and are also conditioned on the lithology from drillhole and surface observations (i.e., outcrop contacts). We apply the methodology to a field study of a Cu-Ni-PGE (copper-nickel-platinum group element) prospect hosted in a mafic intrusion, the Crystal Lake Gabbro (CLG), in northwest Ontario. Our study focuses on a single segment of the y-shaped CLG. The numerical outcomes of this application demonstrate that the incorporation of expert knowledge, drillholes, and surface data yields models with reliable geological geometry features, particularly the distribution of bottom boundary for intrusion models which is highly associated with economic mineralization.

1. Introduction

Geometrical features are significant in geosciences. For example, the fault geometry provides insights into the distribution and propagation of stresses caused by earthquakes (Lindsey and Fialko, 2013; Scognamiglio et al., 2018; Shimizu et al., 2020, 2021; Wei et al., 2023; Huang et al., 2023b). The geometry of magmatic intrusions plays a vital role in understanding geodynamic settings and strategically planning future exploration activities (Mungall, 2007a,b; Ding et al., 2012; Naldrett, 2013; Mota-e-Silva et al., 2013; Lightfoot and Evans-Lamswood, 2015). Understanding the geometry of sedimentary layers is essential in studies regarding basins (Flemings and Jordan, 1989; Miall, 2013; Civico et al., 2017) and lakes (Creighton et al., 2018; Yan et al., 2022). The diverse examples highlight the significance of geological geometry features in both fundamental research and practical applications.

Geometric features and shapes can be constructed explicitly as 3D models. The explicit way requires precise spatial locations of geological structures, such as fault systems (Frank et al., 2007; Cherpeau et al., 2012; Athens and Caers, 2022). Geoscientists can delineate the shape

for targets of interest manually in a commercial software, such as Leapfrog, using available drillhole data. This hands-on strategy allows for the modeling of geological structures informed by expert knowledge. 3D explicit modeling is computationally expensive, because it involves a detailed construction of each coordinate to directly define shapes of complex geological features. This approach, while straightforward, offers limited flexibility in constructing corners, boundaries, and capturing small structural variations. Perhaps more importantly, addressing uncertainty becomes challenging with this approach. To address these challenges, the implicit modeling method was developed.

The implicit method, which employs 3D scalar fields, has become increasingly important and is now the most favored approach for constructing geological models with complex geometric characteristics. The potential-field based interpolation (Lajaunie et al., 1997), as a typical example of implicit modeling, has been widely used in software or open-source packages, such as LoopStructural (Grose et al., 2021a,b), Gempy (de la Varga et al., 2019), and GeoModeller (Calcagno et al., 2008). Users can establish several points by leveraging available data,

* Corresponding author.

E-mail address: xwei2@stanford.edu (X. Wei).

for example, drillholes and field observations. To incorporate expert knowledge regarding subsurface characteristics, it is also possible to define a few points manually. All these points control the structure of interfaces in the modeling framework, serving as the foundation for interpolating the subsurface features using the potential-field method. The interpolation, based on several control points, typically results in a single optimal model without considering the uncertainty. A few studies quantified the uncertainty associated with 3D geological models by slightly varying the spatial location of control points using predefined probability distributions (Lindsay et al., 2012; Linde et al., 2015; Pakyuz-Charrier et al., 2018; Stamm et al., 2019; Athens and Caers, 2022; Huang et al., 2023a). We note that the quantified uncertainties reflect only a limited range of structural variability.

To enhance the quantification of uncertainty associated with geometric characteristics in implicit modeling, the previous work introduced a novel perturbation approach designed to construct geological interfaces that reflect the geometry characteristics (Fouedjio et al., 2021). The resulting models can also closely align with drillhole measurements. However, this method has limitations on adequately reflecting comprehensive geological knowledge. In Fouedjio et al. (2021), an issue with the resulting porphyry copper deposits is the rounded isosurfaces. This rounding effect is a direct consequence of employing a Gaussian assumption in the modeling process, where the level set of a Gaussian field is an ellipse or circle. Such geometric simplifications stand in contrast to complex geometries of copper deposits. For example, the porphyry copper deposits often exhibit strikes with the funnel shape (Piquer et al., 2021). A recent work introduces the knowledge-driven trend surface analysis to incorporate the geological knowledge into the modeling process (Wang et al., 2023). In this work, the essential geological knowledge, such as geological structures of a valley, was represented as a series of lines after the skeletonization. This skeletal representation then serves as a reference in the modeling process (Wang et al., 2023). A critical step of skeletonization involves selecting a threshold that dictates the degree of skeletonization. This predefined value significantly influences the skeletal representation of models. The challenge lies in determining a threshold that appropriately reflects geological structures, avoiding both over-reflecting and under-reflecting the geological knowledge.

Several methods have also been employed to integrate geological knowledge with multiple data sources for constructing geologically realistic models and quantifying uncertainties. Multiple-point statistics (MPS) can be used to generate subsurface models from diverse data (Caers and Zhang, 2004; Strebelle and Levy, 2008; Mariethoz and Caers, 2014; Jørgensen et al., 2015; Hoyer et al., 2017) and multiple 2D geological cross-sections (Chen et al., 2018; Hou et al., 2021). The manually interpreted 3D hydrostratigraphic models serve as inputs for the probabilistic modeling (Madsen et al., 2022). Moreover, deep learning techniques have gained prominence as intelligent and effective tools for achieving these objectives. For example, generative adversarial networks (GANs) are capable of generating complex sequences of geological models (Song et al., 2021; Jordão et al., 2022) and Bayesian neural networks provide a reliable framework for uncertainty quantification (Jordão et al., 2023).

In this paper, we introduce a novel approach for integrating geological knowledge into 3D stochastic modeling. When consulting textbooks, or consulting with geologists, geometric information is often conveyed using 2D diagrams, sometimes even sketches. These diagrams are important knowledge, but remain conceptual. By conceptual, we mean that the scales and orientations of these diagrams cannot be directly exported easily to 3D models. The aim of our work is to integrate these 2D conceptual diagrams to constrain 3D geological models for exhibiting realism. Our approach allows for flexible and automatic adjustments to the diagram, including rotation, scaling, and translation, acknowledging the inherent conceptualization associated with geological diagrams. This flexibility enables the application of a 2D diagram, which outlines specific geometric characteristics, to varied

areas with different orientations, sizes, and spatial locations. In our paper, we begin with a comprehensive overview of the methodologies used in our research. We also conduct a synthetic study to demonstrate the effectiveness and efficacy of our new approach. We apply our framework to a field data application, focusing on characterizing the geometrical features of magmatic intrusions conditioned on drillhole data and outcrop contacts at the Crystal Lake Gabbro in northwest Ontario. Lastly, we discuss and conclude our main findings.

2. Methodology

The proposed methodology integrates a geological diagram and drillhole data into the implicit stochastic modeling to generate subsurface intrusion models. It begins with geometric similarity analysis, aligning 2D cross-sections of 3D models with a reference 2D geological diagram, using Procrustes analysis. The lithologic information from drillholes and outcrop contacts can be incorporated by categorizing lithologies as binary indicators (intrusion or non-intrusion). We employ the implicit function to represent complex geological shapes, which allows for perturbations and uncertainty quantification using Monte Carlo sampling. The loss functions in Monte Carlo sampling balance the contributions of different data types.

2.1. A 2D geological diagram constraint

2.1.1. Geometric similarity analysis

The essence of modeling based on a 2D geological diagram lies in generating a multiple set of 3D models that exhibit geometric similarity with the geological diagram. To achieve this, it is important to quantify the similarity between two shapes, since our approach will be to match/compare a 2D geological diagram with the 2D cross-sections of a 3D geological model. We proposed the use of ordinary Procrustes analysis (OPA) to quantify the geometric similarity between two shapes (Gower, 1975; Goodall, 1991). By leveraging this distance, the 2D cross-section of a 3D model is superimposed onto a 2D geological diagram, after rotation, scaling, and translation, to achieve the best alignment with the diagram. This method minimizes the discrepancies between reference and comparison shapes. The loss function is formalized as follows:

$$L_p(\mathbf{X}_1, \mathbf{X}_2) = \|\beta \mathbf{X}_1 \mathbf{R} - \mathbf{X}_2 - \mathbf{C}\|_F^2 \quad (1)$$

where $\|\cdot\|_F$ is the Frobenius norm. \mathbf{X}_1 and \mathbf{X}_2 represent the coordinates of comparison shape and reference shape, respectively, each of which is a $k \times 2$ matrix constructed from k points in 2 dimensions. β is a positive scaling factor to zoom-in/out the comparison shape. \mathbf{R} is a rotation matrix with the dimension of 2×2 . \mathbf{C} indicates a translation vector, which can move the position of the comparison shape. The initial step in computing the ordinary Procrustes analysis is centering both shapes, which aligns their centroids. We note the aligned comparison and reference shapes as \mathbf{X}_1^c and \mathbf{X}_2^c . Then, the minimized solution of Eq. (1) is given by $\hat{\mathbf{C}} = 0$, $\hat{\mathbf{R}} = \hat{\mathbf{U}}\hat{\mathbf{V}}^T$, where \mathbf{U} and \mathbf{V} are left and right matrices from the singular value decomposition (SVD), and the optimal scaling factor $\hat{\beta} = \frac{\text{trace}(\mathbf{X}_2^c \mathbf{X}_1^c \hat{\mathbf{R}})}{\text{trace}(\mathbf{X}_1^c \mathbf{X}_1^c)}$. The derivative details can be found in Appendix A.

2.1.2. Evaluation metrics

We employed area and circumference as two metrics, independent of the Procrustes analysis, for assessing the performance of geometric similarity analysis. The valid comparative shapes would show the consistency in area and circumference with the reference shape, despite the varied spatial locations. In our work, the reference shape is a geological diagram, typically characterized by irregular and complex geometry. The geological diagram is in a binary format in our study, where 1 and 0 represent the intrusion and non-intrusion, respectively. For a binary model, we adopted the L_0 -norm as an approximation for the

computation of its area and circumference, because L_0 -norm is the sum of the number of non-zero elements within a model. The approximated area for a binary model (i.e., \mathbf{m}) can be written as:

$$E_1(\mathbf{m}) = \|\mathbf{m}\|_0 \quad (2)$$

We calculated gradients in the x and y directions to trace the boundary of a binary model. The approximated circumference is thus:

$$E_2(\mathbf{m}) = \left\| \sqrt{\left(\frac{\partial \mathbf{m}}{\partial x} \right)^2 + \left(\frac{\partial \mathbf{m}}{\partial y} \right)^2} \right\|_0 \quad (3)$$

2.2. Categorical drillholes and outcrop contacts

Our framework can incorporate lithologic information from drillholes and outcrop contacts. With a particular focus on modeling magmatic intrusions, the drillhole and outcrop data categorize lithology into two indicators: intrusion and non-intrusion, encoded as 1 and 0, respectively. The interface between intrusion and non-intrusion is termed the contact point, assigned a value of 0.5. The loss function on the contact point is expressed as (Fouedjio et al., 2021; Wang et al., 2023):

$$L_c(\mathbf{m}) = \frac{1}{n_c} \left\| 0 - \phi^n(\mathbf{m}) \right\|_2^2 + \left(\frac{1}{n_c} \sum_i^{n_c} (0 - \phi^n(\mathbf{m})) \right)^2 \quad (4)$$

where \mathbf{m} is the intrusion model in a binary format. n_c is the number of contact points for all available drillhole data. Function $\phi^n(\cdot)$ represents the implicit function at n th iteration (will introduce later). In our work, the theoretical value of contact points in the implicit function is 0. The right-hand side consists of two components. The first one indicates the mean squared residues (MSR) at contact points, which assess the variance for the prediction. The second component is the squared mean residual, reflecting the bias for the prediction.

The loss function applied to the contact point does not ensure the continuity of the intrusive body along the drillhole trajectory. It necessitates the need of an additional loss function that encompasses the entirety of the binary drillhole data. The foundational concept is to simply count the mismatch of lithologic indicators (i.e., 0 and 1). To render this loss function convex and implement in the negative-logarithmic space, it is formalized as follows (Wang et al., 2023):

$$L_b(\mathbf{m}) = \sum_i^{n_{b,1}} \frac{1}{\log(2)} \log(1 + \exp(-\phi^n(\mathbf{m}))) + \sum_i^{n_{b,2}} \frac{1}{\log(2)} \log(1 + \exp(\phi^n(\mathbf{m}))) \quad (5)$$

where $n_{b,1}$ and $n_{b,2}$, respectively, represent the number of intrusion and non-intrusion data points.

2.3. Implicit stochastic modeling

2.3.1. Implicit modeling and perturbation

In implicit modeling, geometrical configurations are delineated by means of implicit functions rather than explicit geometric entities, such as points, lines, or polygons. An implicit function in two dimensions is typically formalized as $\phi(x, y)$, where x and y are the coordinates of a point. $\phi(x, y) = 0$ represents the point located on boundary of the shape. The point inside and outside the shape satisfies $\phi(x, y) < 0$ and $\phi(x, y) > 0$, respectively. The advantage of implicit modeling lies in its ability to represent complex shapes without heavy representation for all vertices on an irregular shape. Another advantage is that the boundary delineated from an implicit function can be perturbed using the level set method (Osher and Fedkiw, 2001). The level set equation governs the evolution of the interface over time (t), which takes the typically discretized form as:

$$\phi^{n+1} = \phi^n - (v^* |\nabla \phi^n|) \Delta t \quad (6)$$

where v^* represents a normal velocity field in the normal direction. n is the number of perturbations. Δt indicates the step size in the perturbation. In our work, we implemented a special case of implicit functions, known as the signed distance function (Osher et al., 2004), where ϕ represents the distance to the boundary. The gradient of signed distance function equals to one ($\nabla \phi = 1$), which greatly simplifies the computation.

2.3.2. Metropolis–Hastings algorithm

The Metropolis–Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) is an essential component in the Markov chain Monte Carlo (MCMC) sampling. It provides an efficient way to approximate the a posterior probability via randomly sampling a large amount of models. The randomly sampled models will be accepted or rejected following the criteria:

$$\alpha = \frac{P[\phi(\mathbf{m}^i)]}{P[\phi(\mathbf{m}^j)]} \frac{L[\phi(\mathbf{m}^i)]}{L[\phi(\mathbf{m}^j)]} \frac{T[\phi(\mathbf{m}^j)|\phi(\mathbf{m}^i)]}{T[\phi(\mathbf{m}^i)|\phi(\mathbf{m}^j)]} \quad (7)$$

where, ϕ^i and ϕ^j are candidate and current implicit models, respectively. $P(\cdot)$ is the *a prior* probability density, $L(\cdot)$ is the likelihood function and $T(\cdot)$ is a proposal probability density. In our study, the proposal distribution is proven to be symmetric (Wang et al., 2023), manifesting as a Gaussian field with a zero mean and a standard deviation of one. The variables of Gaussian field are subject to uniform prior distributions. The symmetry of both the proposal and prior distributions allows for mutual cancellation. The acceptance ratio is thus determined solely by the likelihood.

Instead of using an exponential format, researchers commonly conduct MCMC simulations in the negative-logarithmic space for improved numerical stability and efficiency. The acceptance ratio, thus, is expressed as $\alpha = L_{nlog}[\phi(\mathbf{m}^j)] - L_{nlog}[\phi(\mathbf{m}^i)]$. In our work, the likelihood consists of four terms, namely, 2D geological diagram, contact points and binary data from drillholes, and geophysical data, which is expressed as follows:

$$L_{nlog}(\phi) = w_p L_p(\mathbf{X}_1, \mathbf{X}_2) + w_c L_c(\mathbf{m}) + w_b L_b(\mathbf{m}) \quad (8)$$

where w represents the weighting parameter to control the contribution of multiple terms. We ran all simulations on a MacBook Pro with 16 GB of memory and a 12-core processor. We will report the computational time for each case individually.

3. Synthetic study: the use of a 2D geological diagram as a constraint

3.1. Understanding a geological diagram constraint

We aim to show the effects of a geological diagram constraint on the stochastic model construction. We impose the geological diagram constraint only, excluding the use of additional data and/or constraints.

Fig. 1 presents a series of comparative results across three cases using different geological diagram constraints. **Fig. 1a** shows a 2D geological diagram characterized by a waterdrop shape. We employed a circle, located at the center of the model space, as the initial model (**Fig. 1b**). We conducted 10 sampling chains, each comprising 20 000 steps. A single chain took approximately 14 min of computational time on the laptop. To enhance model independence, we retained only the last 20% models for each chain and decimated with a factor of two to increase the independence. The resulting mean model (**Fig. 1c**) aligns closely with the geological diagram (**Fig. 1a**). The lower standard deviation suggests a strong consensus for the core area among the sampled models (**Fig. 1d**). The transition from lower (blue) to higher (red) standard deviation visually represents the variability inherent in delineating the boundaries of the models. Panels A-a to A-l in **Fig. 2** displays multiple realizations of the randomly sampled models. Notably, these models exhibit slightly varying rotation angles, such

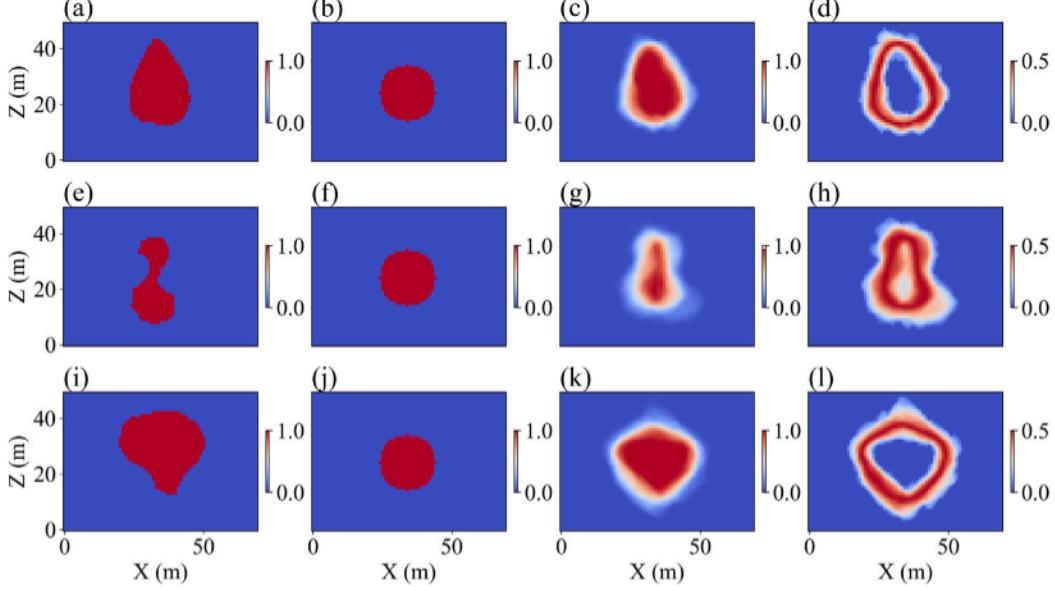


Fig. 1. The comparison across three cases using different geological diagram constraints. Each case is organized in a separate row. The first column displays three distinct geological diagrams. The second column presents the initial model, while the third and final columns show mean and standard deviation models, respectively, derived from accepted models. The color scale represents the lithologic information, with 1 and 0 respectively indicating intrusion and non-intrusion zones.

as rightward (Fig. 2A-a), vertically (Fig. 2A-b), and leftward orientations (Fig. 2A-c). These angles are governed by the rotation matrix R , as specified in Eq. (1). These models also show the variability of geometric characteristics, reflecting the inherent spatial uncertainty in realizations constrained by a geological diagram.

The second and last rows of Fig. 1 introduce the other two cases using distinct geological diagrams. The resulting realizations are shown in Fig. 2B and C, respectively. This comparative study illustrates the effectiveness of a geological diagram constraint. We also employed circumference and area as two metrics for the evaluation of geometric similarity among the true geological diagram and model realizations. These metrics, which are not considered during the sampling process, can serve as independent and unbiased validation for our analysis. We calculated the true circumference and area for the three geological diagrams represented in Fig. 1a, e, and i. These calculations were also extended to all accepted models from the sampling process, the mean of which are shown in Fig. 1c, g, and k. The calculated circumference and area were then plotted on a scatter plot (Fig. 4a). Fig. 4a illustrates that the geometry characteristics of randomly sampled realizations conform to the true geological diagrams, as evidenced by the true values of circumference and area falling into the range of realizations.

To investigate the influence of the initial model, we performed an additional synthetic study using the same diagram (Fig. 3a, e, and i) but varying initial models (Fig. 3b, f, and j). We implemented 10 sampling chains, each of which includes 20 000 steps. We kept only the last 20% models to increase the independence of models. Fig. 3 demonstrates that the initial model affects the location of model realizations during the sampling process. The accepted models tend to center around the initial model's position. However, it is important to recognize that the shape of the models remains uninfluenced by the initial model, adhering to a provided geological diagram constraint. Fig. 4b confirms the geometric similarity among the true geological diagram and multiple realizations based on different initial models. To reduce the bias of any single initial model on the final ensemble realizations, we randomly sample initial models across multiple sampling chains. We aim to achieve a more representative and unbiased sampling that is constrained by a geological diagram, rather than the starting models.

3.2. A geological diagram conditioned to drillhole data

We integrate drillhole data with a geological diagram constraint to refine our model construction. The drillhole data indicates the intrusion distribution along the drilling trajectory. We anticipate that the randomly sampled intrusion models will align with the actual subsurface intrusion indicated by the drillhole observations while exhibiting geometric characteristics that are consistent with a geological diagram. We conducted three experiments to validate our framework.

The first experiment involves three different geological diagrams, each subject to the same drillhole data (Fig. 5). We implemented a single sampling chain for each diagram constraint with a randomly generated initial model. The mean model in Fig. 5d shows a similar shape to the diagram constraint in Fig. 5b, while this mean model presents a leftward rotation. The mean model is also consistent with the drillhole data, as indicated by the overlap between the red intrusion areas and the black drilling trajectory in Fig. 5d. The lower standard deviation area (Fig. 5e) is consistent with the intrusion indicated by the drillhole data. This suggests that the variability in model realizations is lower in the region where the drillhole data provides a direct constraint. In Fig. 5, the panels, from f to j and k to o, present the modeling process using the other two geological diagrams. These results illustrate the efficacy of the modeling approach, wherein the randomly sampled models are conditioned on both the geological diagram constraint and drillhole data.

In the second experiment, we explored the influence of varying drillhole data given a same geological diagram constraint. This experiment is designed to demonstrate the adaptability of a geological diagram constraint to various information from drillholes. The first drillhole reveals a substantial intrusion (Fig. 6a). This drillhole penetrates the left segment of the mean intrusion model (Fig. 6d). The second drillhole presents a dipping trajectory, as shown in Fig. 6f. The resulting mean model (Fig. 6i) aligns with the dipping drillhole at the top of the intrusive body. The third drillhole scenario shows a thin intrusion feature (Fig. 6k). This drillhole is positioned at the left margin of the mean model in Fig. 6n. Regardless of the drillhole data, ranging from thick to thin intrusions, and from vertical to dipping trajectory, the

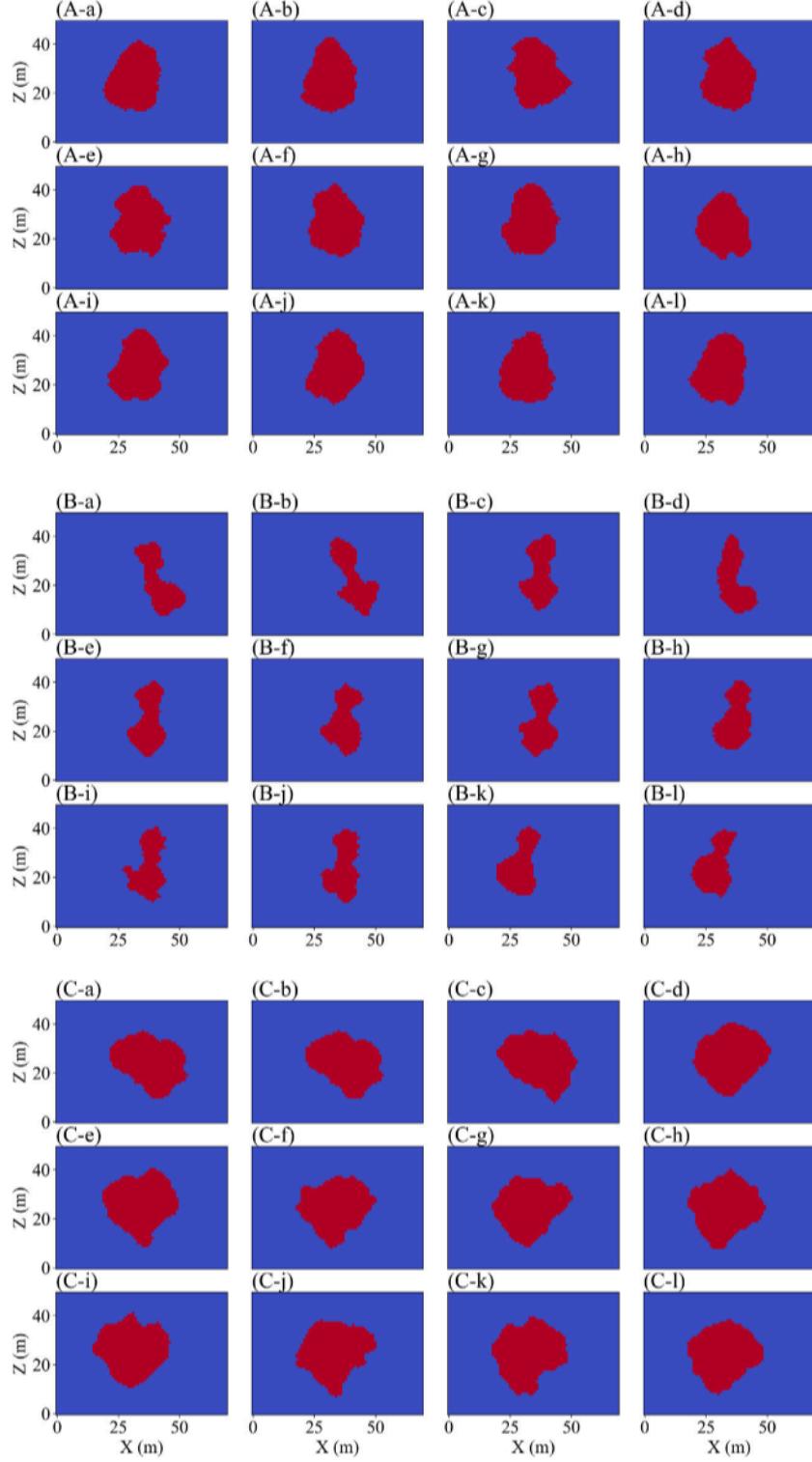


Fig. 2. Model realizations based on different geological diagram constraints. Panels A-a to A-l associate with the geological diagram shown in Fig. 1a. Panels B-a to B-l correspond to the diagram in Fig. 1e. Lastly, panels C-a to C-l are resulted from the geological diagram in Fig. 1i.

geological diagram constraint demonstrates a capacity to accommodate these differences constrained by a geological diagram.

Given a drillhole and a geological diagram, it is legitimate to expect the intrusion models to have variabilities in the location and orientation. The drillhole acts as an anchor point, providing a reference for the presence of the intrusion. The randomly sampled models can exhibit a range of spatial orientations around this anchor point, akin to a pendulum that can swing in various directions, from left to right,

around the drillhole. To verify our expectations, we conducted the third experiment. We implemented five sampling chains with random initial models using the same drillhole data and a geological diagram. Fig. 7 shows the results of nine sampling chains as examples. We observe that the drillhole data are situated variably at the left (Fig. 7d), right (Fig. 7i), and center (Fig. 7n) of intrusion models with different orientations, which demonstrate the variability in the location and orientation during the sampling process. Fig. 8 shows the mean and

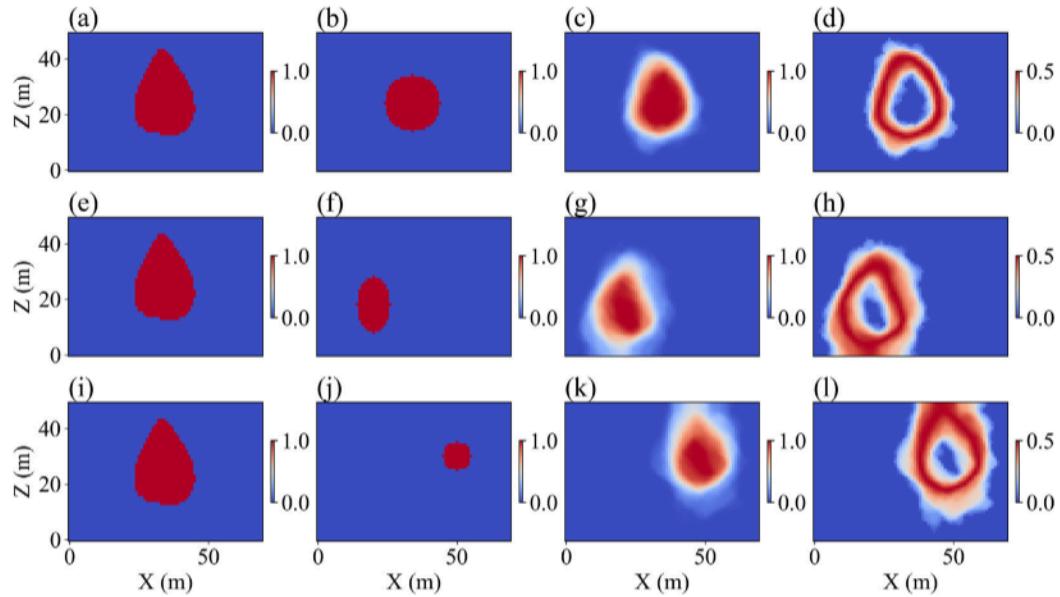


Fig. 3. The comparative outcomes from three cases, which employ the same geological diagram constraint but different initial models. Each case is organized in a separate row. The initial column displays the geological diagram. The second column presents different initial models. The third and final columns show mean and standard deviation models, respectively.

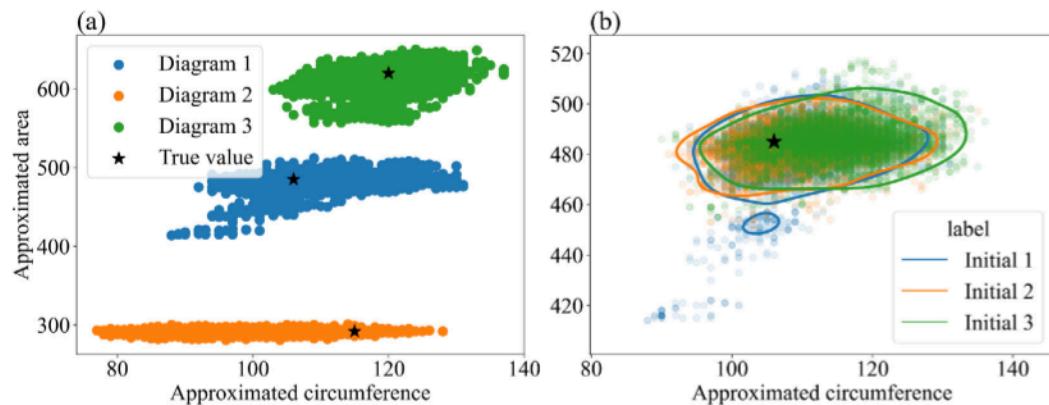


Fig. 4. The performance evaluation of the geological diagram constraint. (a) links to Fig. 1, demonstrating the efficacy of different geological constraints. (b), corresponding to Fig. 3, illustrates that variations in initial models do not influence the geometric similarity.

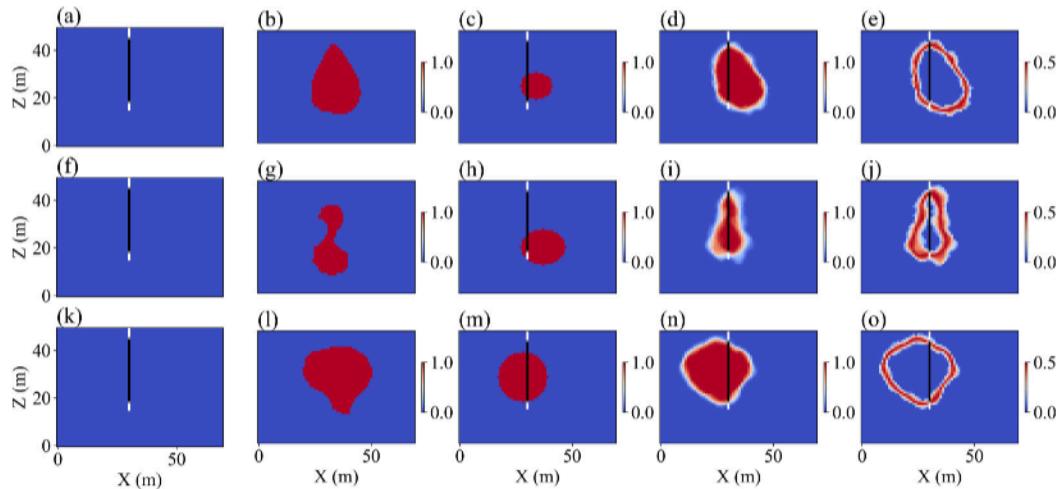


Fig. 5. The comparison of three different geological diagram constraints conditioned on the same drillhole data. The first column shows the drillhole data where the black and white colors indicate the intrusion and non-intrusion, respectively. The second column shows three different diagram constraints. The third column presents the randomly sampled initial models. The last two columns are mean and standard deviation models.

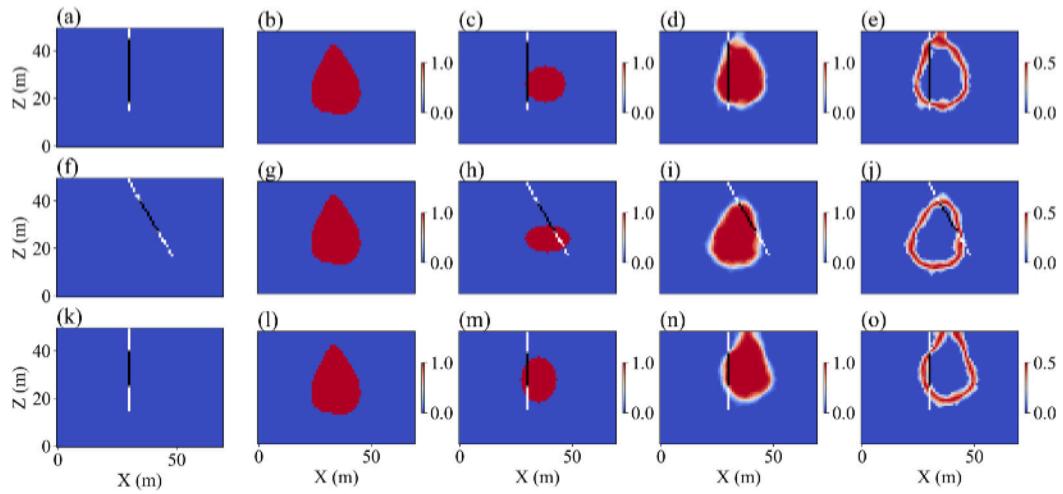


Fig. 6. The comparison of a geological diagram constraint conditioned on three different drillholes. The first column presents varied drillhole data. (a) depicts a vertical drillhole characterized by a thick intrusion zone along its drilling trajectory. (f) shows a drillhole with a dipping orientation. (k) is a vertical drillhole that reveals a thin intrusion zone. The second column shows a diagram constraint with a water-drop shape. The third column presents the randomly sampled initial models. The last two columns are mean and standard deviation models.

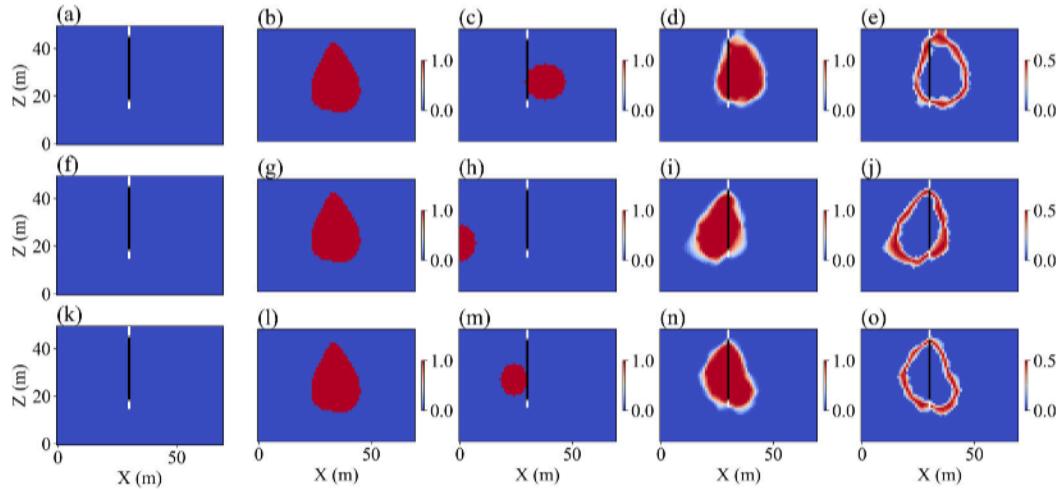


Fig. 7. Multiple sampling chains using the same drillhole and geological diagram constraint. The columns, from left to right, represent drillholes, diagrams, initials, mean and standard deviation models.

standard deviation models among the five sampling chains using the same drillhole and geological diagram. The mean (Fig. 8c) and standard deviation (Fig. 8d) models clearly exhibit randomly sampled models swinging in various directions around the drillhole.

The results of these three experiments illustrate the effectiveness and robust capability of our modeling framework, successfully integrating and adapting to both drillhole data and geological constraints. Note that, compared to the previous case, incorporating drillhole data increases the computational expense. A single chain costed around 21 min.

3.3. The effects of weighting parameters

The likelihood function, Eq. (8), includes multiple terms corresponding to different data sets, each associated with a specific weighting parameter. The selection of these weighting parameters is important during MCMC sampling, because they influence both the contribution of various data sets and computational efficiency. Choosing appropriate weighting parameters often involves a trial-and-error process. In this work, we propose and follow two key guidelines to determine the weighting parameters: first, we aim for equal contribution from all data

sets during the sampling process, and second, we target an acceptance rate of approximately 20% (Earl and Deem, 2005).

We tested various weighting parameters using the drillhole and geological diagram data shown in Fig. 5a and b, respectively. We performed MCMC sampling with a fixed weighting parameter of 1 for the drillhole, and varied the weighting parameter for the geological diagram. The drillhole-to-diagram weighting ratio reflects the relative contribution of these two data sets. As the ratio increases, the acceptance rate correspondingly rises, as demonstrated from Fig. 9a to e. Fig. 9b, c, and d show that when the ratio falls between the range of $\frac{1}{25}$ and $\frac{1}{5}$, the loss value curves of drillhole and diagram are closely aligned. In contrast, when the geological diagram dominates the sampling process, its loss value exceeds that of the drillhole, resulting in a lower acceptance rate (Fig. 9a). Conversely, when the loss value of the drillhole is much larger than the value of the diagram, it leads to a higher acceptance rate (Fig. 9e). Both excessively lower and higher acceptance rates are undesirable. To achieve the target acceptance rate of around 20%, the appropriate weighting parameters should fall within the range shown in Fig. 9b and c. Therefore, in our synthetic studies, we set a weighting parameter of 1 for the drillhole data and 20 for the geological diagram.

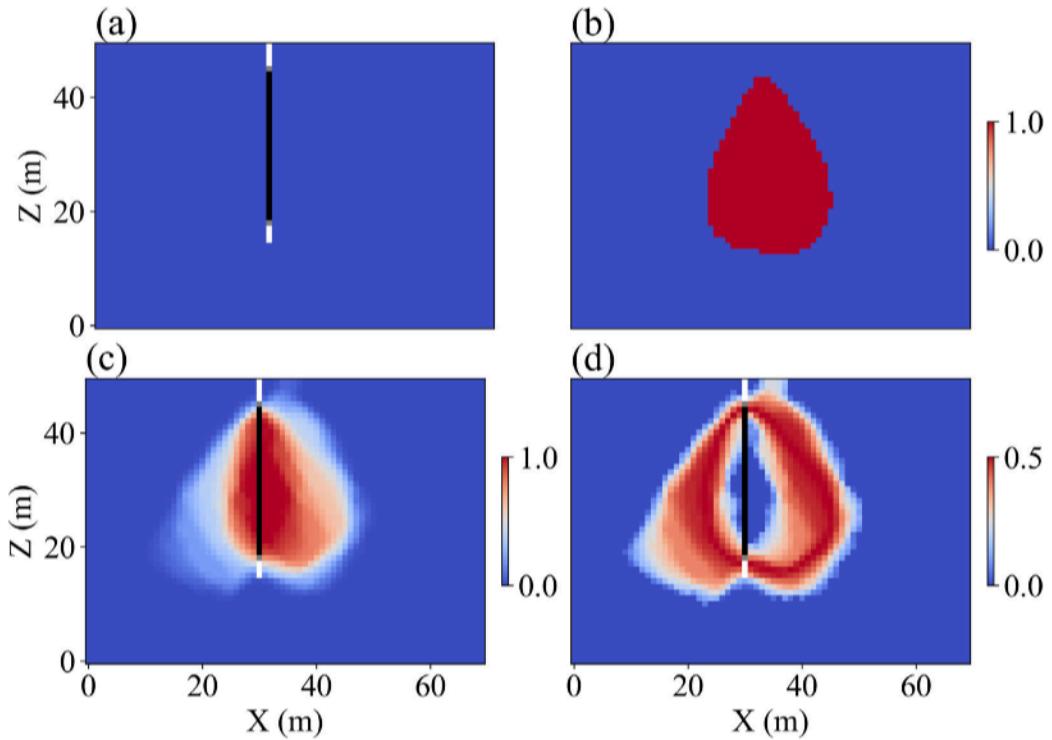


Fig. 8. Using the same drillhole (a) and geological diagram constraint (b), we computed the mean (c) and standard deviation (d) models from five sampling chains. Note that Fig. 7 shows the results from three out of the five sampling chains.

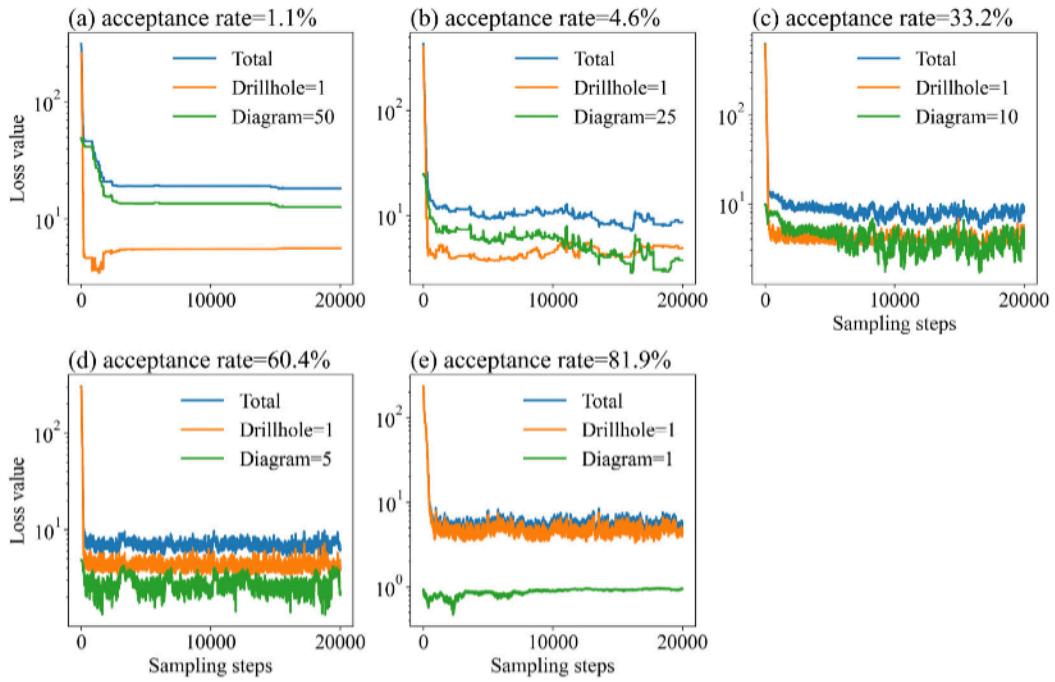


Fig. 9. The loss values of MCMC sampling chains using different weighting parameters given drillhole data (i.e., Fig. 5a) and a geological diagram (i.e., Fig. 5b).

4. Field application: a study of the crystal lake gabbro in north-west Ontario

4.1. Geological setting and data

The Lake Superior region in both the United States of America (USA) and Canada has experienced a complex geological history culminating in the Mesoproterozoic (1115–1086 Ma), intracontinental

Midcontinent Rift (MCR) and associated large igneous province (Canon, 1992). Two significant phases of magmatism within the lifespan of the rift (Miller and Nicholson, 2013) produced Ni-Cu-PGE-bearing mafic and ultramafic intrusions including the Eagle Ni deposit in Michigan, the Tamarack Ni deposit in Minnesota, and the many Cu-Ni-PGE deposits of the Duluth Complex in Minnesota (Ding et al., 2010; Taranovic et al., 2015; Severson et al., 2002). Another mineralized mafic intrusion emplaced during the MCR is the Crystal Lake Gabbro (CLG)

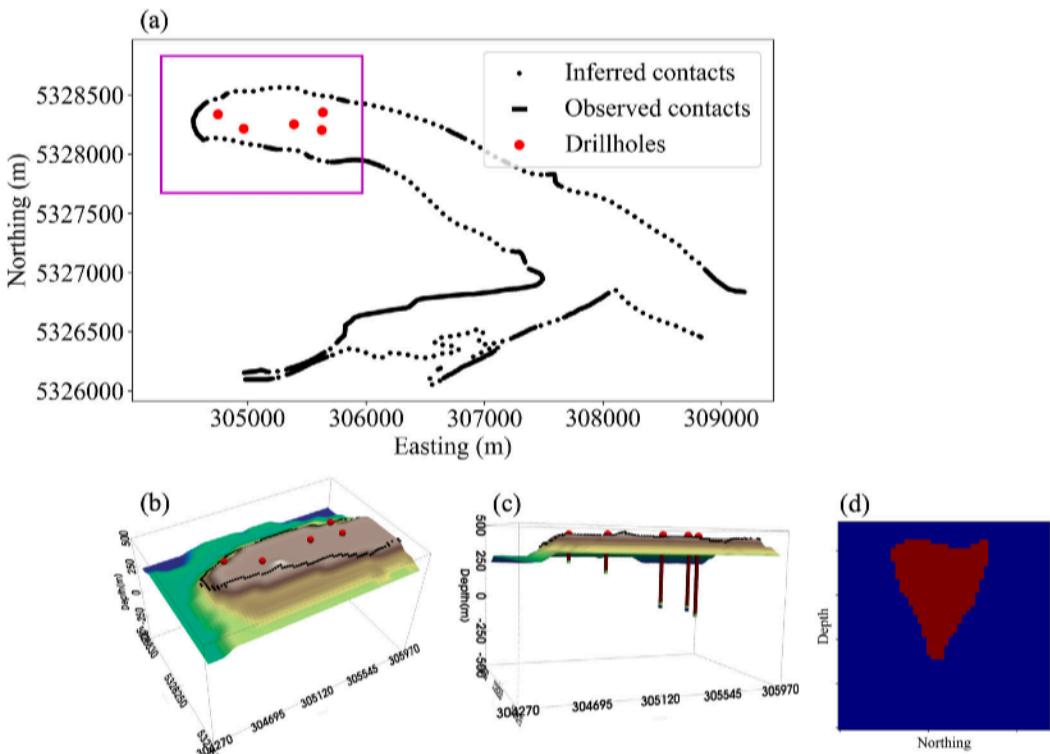


Fig. 10. (a) shows the CLG intrusion. The solid black lines mark outcrop contacts directly observed in the field from Geul (1970). The black dots are inferred contacts which were blind in the field and inferred with geophysics and topography. The magenta box outlines the study area, where red dots indicate the locations of five drillholes used in the modeling. (b) presents a 3D visualization of the study area where the background color indicates the topography, red balls and black dots mark the locations of drillholes and outcrops, respectively. (c) offers a 3D visualization from a different viewing angle, enhancing the visibility of drillholes along the depth. (d) is a manually crafted 2D geological diagram for the CLG intrusion based on mineral system models for Eagle-Kalatongke type Ni-Cu-PGE deposits (Ding et al., 2012; Lightfoot and Evans-Lamswood, 2015).

along the western shore of Lake Superior and approximately 7 km north of the USA-Canada border. The CLG is a y-shaped, Ni-Cu-PGE-bearing mafic intrusion (Fig. 10a) emplaced into the Paleoproterozoic Rove Formation, which was the principal source of sulfur for the Cu-Ni mineralization (Ripley, 2014; Benkó et al., 2015; O'Brien, 2018). It is up to 750 m wide and informally subdivided into a 5 km “north arm” and a 2.75 km “south arm” (Geul, 1970; Smith and Sutcliffe, 1987; Pardee and Townships, 0000). Our study focuses on the north arm of the CLG. The north arm consists of four distinct zones: Basal, Lower, Middle, and Upper. The Basal Zone comprises a 7 m zone of aphanitic to fine-grained gabbro, containing assimilated xenoliths of the wall rocks. The Lower Zone, up to 50 m in thickness, consists of coarse-grained gabbro with pegmatitic patches, blocks, and disseminated sulfides. The Middle Zone, 30 m thick, exhibits distinct layering features of anorthosite, olivine leucogabbro, chromite-rich anorthosite, and melanocratic olivine gabbro. The Upper Zone, 80 m thick, is characterized by coarse-grained olivine gabbro overlain by medium-grained troctolite, with the disappearance of chromite-rich layers (O'Brien, 2018). All significant mineralization is hosted in the Lower Zone.

This study focuses on a segment of the north arm of the CLG, indicated by the magenta box in Fig. 10a. This area is formed by a single pulse magma (O'Brien, 2018). Our primary objective is to delineate the geometry of intrusions and to assess the associated uncertainty on their spatial distributions. In the survey area, the west section benefits from a robust constraint owing to numerous drillholes, and the east portion is comparatively less constrained. We expect the constructed models, however, to inform the geometric configuration of intrusions in the less constrained area to aid future exploration efforts and to test the validity of deposit models.

Our study relies on categorical lithology information from drillholes and outcrop contacts as well as a 2D geological diagram of intrusions. We employed five drillholes and a set of outcrop contacts in the

modeling (Fig. 10). The outcrop contacts indicate a boundary that separates one lithology from another. In our work, the outcrop contacts were classified as *observed* or *inferred*. The observed contacts represent the exposed rocks that were directly observed in the field (the solid black lines in Fig. 10a). If visible outcrops are absent, experts can infer the location of contacts using geophysics, topography, and information from drilling. These deduced boundaries are referred to as inferred contacts indicated by the black dots in Fig. 10a. In the 3D modeling, inferred contacts were treated with the same level of constraint as observed contacts for two primary reasons. First, the y-shape of CLG is well-documented in the literature (Geul, 1970; O'Brien, 2018; Smith et al., 2020), based on geological, geophysical, and geochemical evidence, which supports the reliability of inferred contacts. Second, using more outcrop contacts can enhance the constraint quality, leading to more reliable and continuous intrusion features near the surface. Barnes et al. (2016) summarized several typical geometric characteristics of mafic and ultramafic intrusions that host magmatic Ni-Cu sulfide mineralization, such as an elongate sill known as Noril'sk type (Naldrett, 2013), a tubular chonolith based on Nebo-Babel type (Seat et al., 2007; Mota-e-Silva et al., 2013), a blade shaped dyke referred to as Expo-Savannah type (Mungall, 2007a,b), and a funnel shape known as Eagle-Kalatongke type (Zhou et al., 2002; Song et al., 2011; Ding et al., 2012; Lightfoot and Evans-Lamswood, 2015). The drillhole data in the west of the survey area suggest a higher possibility of conforming to a funnel shape for the CLG intrusion. We thus manually crafted a 2D geological diagram illustrating the funnel shape (Fig. 10d), which will serve as a vital geological diagram constraint in our modeling endeavors.

4.2. 3D stochastic modeling of intrusions

4.2.1. Case 1: modeling from drillholes and outcrop contacts

We used categorical drillholes and outcrop contacts to construct intrusion models. We performed nine sampling chains, each comprising

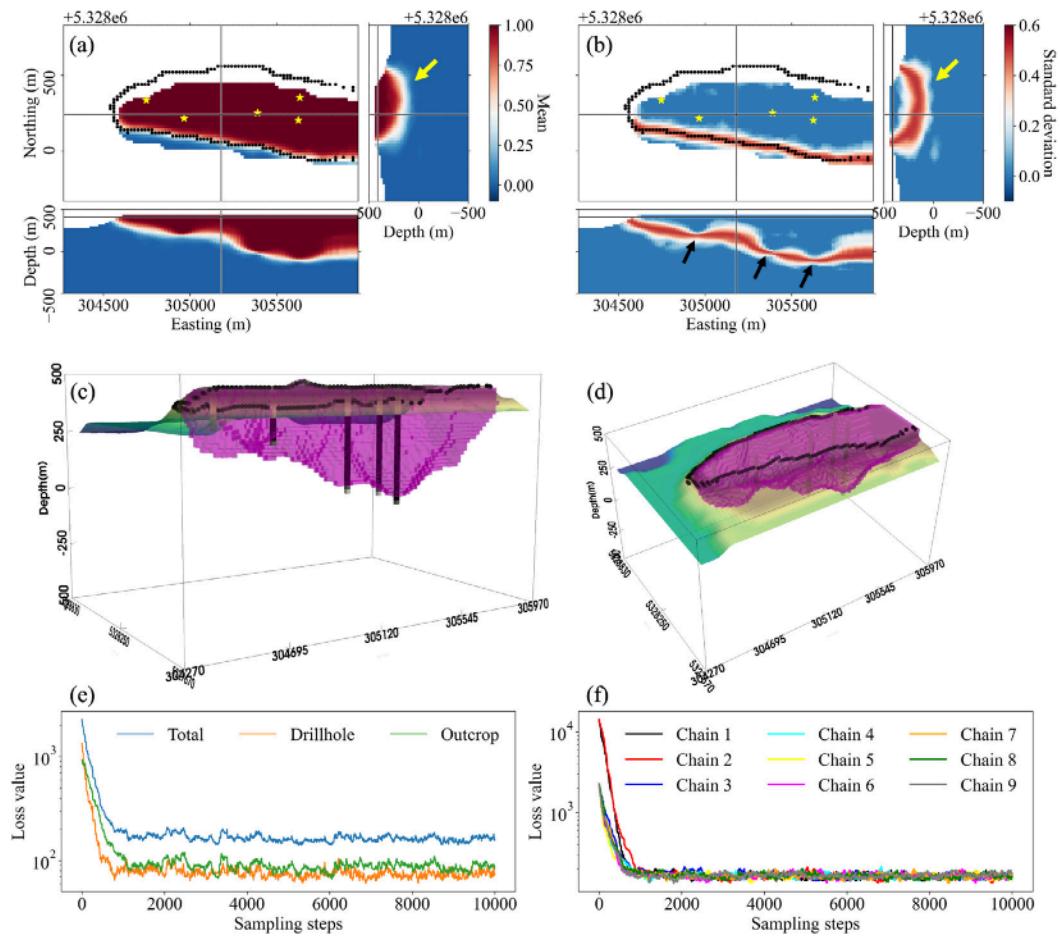


Fig. 11. The mean (a) and standard deviation (b) models across the nine sampling chains. The yellow stars mark the locations of drillholes. The back dots indicate the locations of outcrop contacts. (c) and (d) show the 3D visualization of the last sampled model from the third sampling chain in different viewing angles, where the constructed model matches with both drillholes and outcrop contacts. (e) shows the loss values of drillholes and outcrop contacts separately as well as the total loss value for the one sampling chain. (f) represents the total loss value for nine sampling chains.

10,000 steps initiated from a randomly generated model. A single chain required approximately 7 h to run on the laptop. For each chain, we kept the last 20% of models for the subsequent analysis. We calculated the mean and standard deviation models across all three chains.

Fig. 11a presents the mean model. The mean model's depth slice aligns with the location of contacts indicated by black dots in Fig. 11a. The west-east cross-section reveals the general trend of intrusion models, where a decrease in thickness suggests magma flow direction from east to west. The cross-section along the south-north direction, as highlighted by the yellow arrow in Fig. 11a, shows the rounded bottom feature of intrusion models. This rounded geometry contradicts the expected funnel or anti-triangle shape typical of Cu-Ni-PGE intrusions, indicating a deviation from our geological understanding. The standard deviation model (Fig. 11b) captures the uncertainty associated with intrusion models. The higher standard deviation values closely match the spatial distribution of contacts, indicating variability in the fitting of outcrop contacts. The near-zero standard deviation of the intrusion body within outcrop locations reflects a high confidence level in the existence of intrusion. The depth slice highlights reduced standard deviation at drillhole sites (marked by black arrows), demonstrating the efficacy of drillhole constraints. The depth slice along the southwest direction exhibits higher standard deviation aligning with the rounded bottom of the intrusion models (yellow arrow in Fig. 11b). We visualized the last sampled model with topography from the third chain in a 3D view, enabling an enhanced assessment of the goodness-of-fitting for both drillholes and outcrop contacts. Fig. 11c and d show these 3D models from different viewing angles. Fig. 11e shows the loss values

of drillholes and outcrop contacts separately as well as the total loss value which is the sum of losses for drillholes and outcrop contacts in the one sampling chain. Fig. 11e reflects a promising convergence of the sampling chain, as evidenced by the “hairy-caterpillar” features. Fig. 11f shows the total loss value across all three chains.

We also examined the model cross-sections for each drillhole. Fig. 12a–e display the mean models that closely align with the drillholes. The standard deviation models (Fig. 12f–j) emphasize the confidence in the main body intersected by drillholes and reflect the high variability at the bottom of drillholes. Both mean and standard deviation models exhibit a rounded-flat bottom rather than the geologically anticipated funnel or anti-triangle geometry. This discrepancy highlights the essential need to integrate geological knowledge into the modeling process.

4.2.2. Case 2: knowledge-driven modeling from drillholes and outcrop contacts

We integrated geological knowledge into constructing CLG intrusion models. The geometry of intrusions resembles a funnel or anti-triangle shape. According to the literature, we manually crafted a 2D geological diagram to represent the geometric characteristics of intrusions (Fig. 10d). We employed this 2D geological diagram as a constraint in the modeling, alongside the data from drillholes and contacts. Our objective was to generate intrusion models that not only align with the data from drillholes and outcrop contacts but also adhere to the geological knowledge reflected by a 2D geological diagram. We conducted nine sampling chains using the same initial models as the previous section.

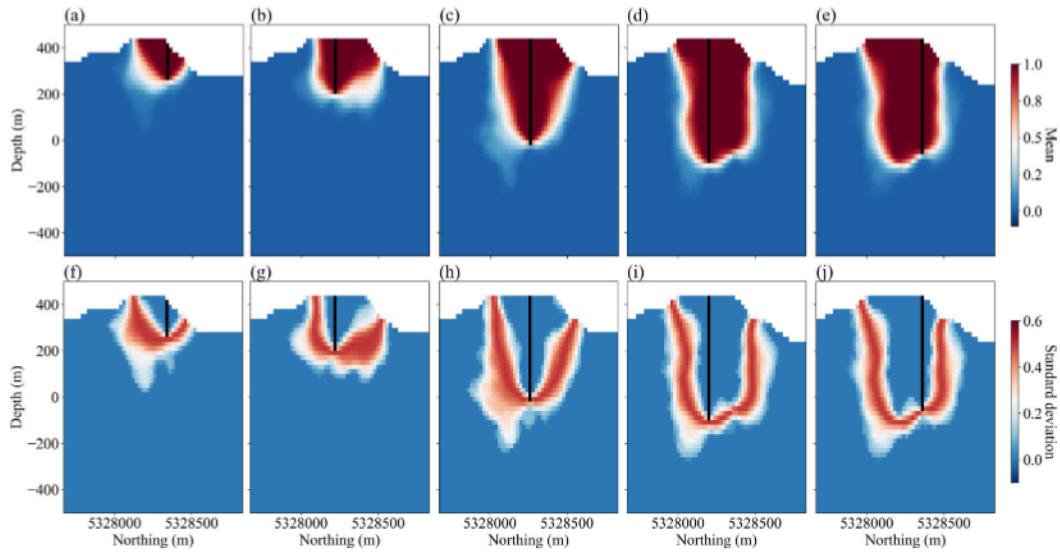


Fig. 12. The model cross-sections for each drillhole along the south–north direction. In each row, the sub-panels are arranged according to the sequence of drillholes from west to east shown in Fig. 11a. (a)–(e) represent the mean model. (f)–(j) show the standard deviation model.

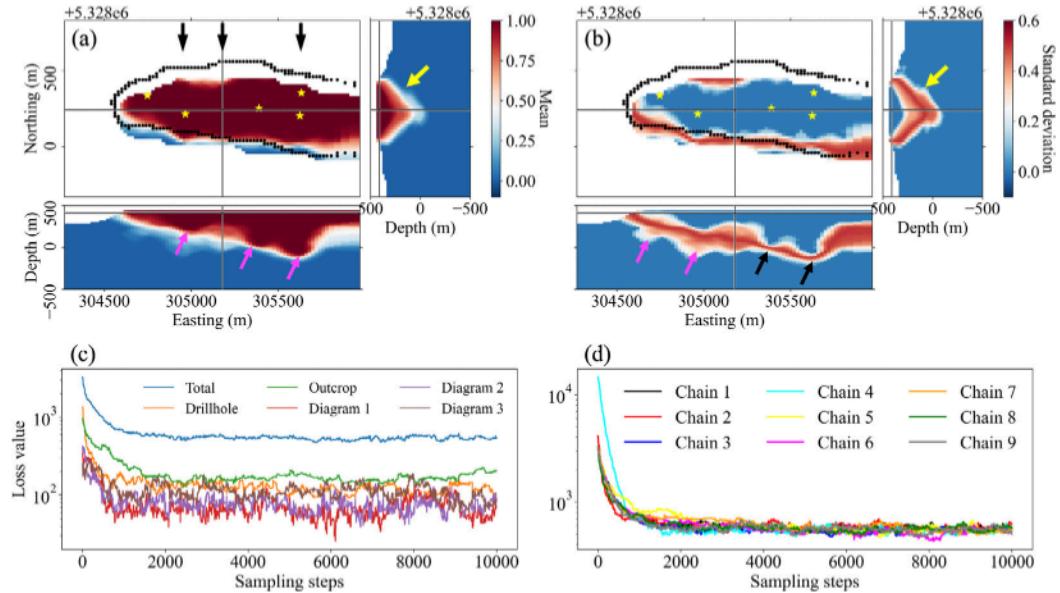


Fig. 13. (a) The mean model across the nine sampling chains constrained by the geological diagrams. The black arrows indicate the spatial positions where the geological diagram constraints are imposed. (b) displays the knowledge-driven standard deviation model. (c) shows the loss values of drillholes, outcrop contacts, and geological diagrams separately as well as the total loss value for the one sampling chain. (d) represents the total loss value for nine sampling chains.

We imposed the 2D geological diagram constraint on three slices indicated by the black arrows in Fig. 13a. The south–north cross-section of the mean model (Fig. 13a) reveals the anti-triangle shape of intrusion models, highlighted by a yellow arrow, which is deemed geologically plausible when contrasted with the rounded-flat shape presented in Fig. 11a. The profile extending in a west–east direction exhibits a similar trend observed in Fig. 13a. The use of a geological diagram constraint was not anticipated to significantly alter the general trend of intrusions, given the same drillhole and outcrop contact data in the modeling process. Nevertheless, observable improvements are noted. For instance, the west–east cross-section reveals several protrusions, denoted by magenta arrows in Fig. 13a. These protrusions indicate the magma flow depressed into sedimentary layers, which are geologically plausible.

The standard deviation model shows an anti-triangle shape marked by the yellow arrow in Fig. 13b. The black arrows point to the areas with a lower variability, reflecting the constraints from drillholes. In

Fig. 13b, we also observe the intrusive bodies characterized by an intermediate level of standard deviation indicated by the magenta arrows. These intrusive bodies are likely to be artifacts. The caterpillar features in Fig. 13c and d verify the convergence of the sampling process. Fig. 14 shows the model cross-sections for each drillhole. The mean models (Fig. 14a–e) align with the drillholes and the standard deviation models (Fig. 14f–j) emphasize the confidence of the main body intersected by drillholes. Both the mean and standard deviation models display an anti-triangle shape, which is consistent with our geological knowledge.

Fig. 15 presents a comparison of multiple realizations from models with and without geological diagram constraints. Panels a–l display 2D cross-sections extracted from models that do not incorporate geological diagram constraints. These models exhibit inconsistencies with established geological knowledge regarding magmatic intrusions, such as the anti-triangle shape documented in the literature (Lightfoot and Evans-Lamswood, 2015). The depicted random geometric features in

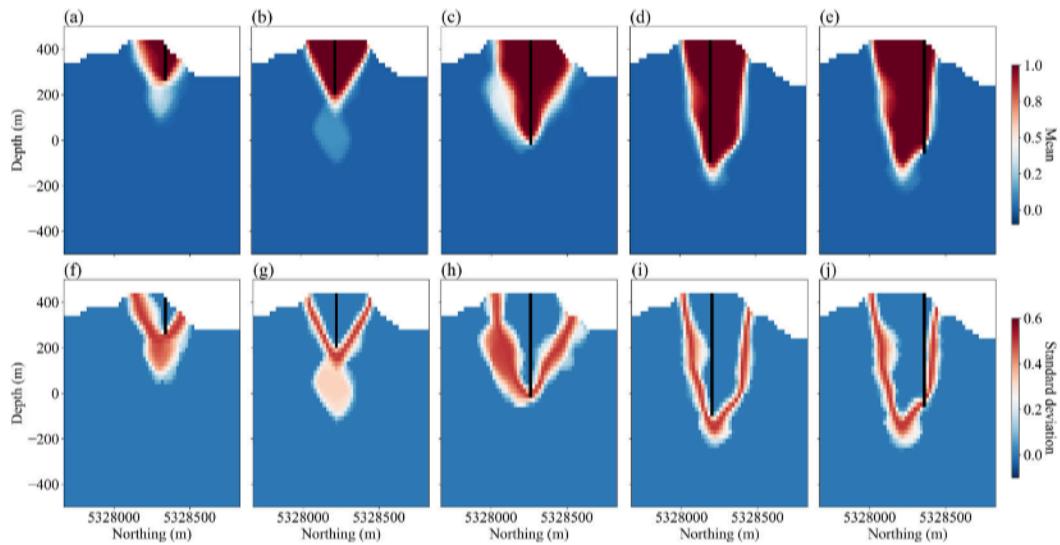


Fig. 14. The model cross-sections for each drillhole along the south-north direction. (a)–(e) represent the mean model constrained by the geological diagrams. (f)–(j) show the standard deviation model.

Fig. 15a–l are influenced by the stochastic sampling process. **Fig. 15m–x** feature 2D cross-sections that are constrained by the geological diagram. The general anti-triangle shape is evident in each panel, illustrating the success of our framework. The variations among the panels (**Fig. 15m–x**) highlight the inherent uncertainty of intrusion models.

5. Discussions

The developed framework is highly versatile and can be applied to a wide range of research problems, including the modeling of karst conduits, water aquifers, and volcanic structures. Our implementation is also compatible with different inversion algorithms. The resulting models can serve as excellent prior for the Bayesian inversion. We encourage researchers to explore combining a geological diagram constraint with inversion algorithms in future work. We parameterize the model space into cells (i.e., prisms), where the cell size directly influences the resolution of the models. A practical approach to determining cell size is to base it on the spacing of the available drillholes. We recommend that the cell size should be smaller than the minimum drillhole interval to capture adequate structural variations between drillholes. The parameterization strategy should also depend on the available computational resources, ensuring computational feasibility while maintaining model resolution.

Addressing several challenges could further enhance our framework and implementation. First, the current model perturbation, based on the normal direction of a Gaussian field, struggles with highly curved geometries (e.g., **Fig. 1e** and i). Two alternative approaches could improve the recovery of high curvature features. The one is to perturb the model shape based on curvature and incorporating curvature misfit into the objective function. Curvature is a more effective metric of capturing highly curved variations. Another one is employing multiple-level sets to construct the complex geometry. [Yang et al. \(2019\)](#) performed multiple level sets to construct the 3D geological surfaces for several lithologies. Also, the 3D stochastic level set poses computational challenges related to perturbation and forward modeling. Future improvements in computational efficiency include wavelet domain sampling, unstructured mesh parameterization, and advanced parallel techniques. For example, [Cai et al. \(2021, 2022\)](#) implemented an advanced parallel direct solver on unstructured mesh parameterization to greatly accelerate the large-scale computations involved in electromagnetic simulations. Further advancements can harness NVIDIA Kepler GPUs and hybrid

MPI-CUDA programming ([Tu et al., 2024](#)). We believe that this combination of hardware and software, as demonstrated by [Tu et al. \(2024\)](#), represents a promising approach to enhance computational efficiency in the stochastic level set modeling. Moreover, properly adjusting weights for different data sources is crucial. Trial-and-error experiments were performed to achieve convergence in our work. The development of an advanced, automatic, and adaptive algorithm will be interesting and significant. This algorithm should be capable of dynamically adapting user-specified parameters based on the data and geological setting, such as hierarchical Bayesian inversion ([Wang et al., 2022](#)).

Note that a geological diagram, used to constrain models, represents a hypothesis, and may not always be accurate. It is thus essential to test the diagram when integrating it into the modeling process. In regions such as the Crystal Lake area, where extensive drillhole data provide a solid understanding of the subsurface, the geometric characteristics of intrusions can be delineated with a high degree of confidence. This depth of understanding supports the use of a 2D geological diagram as a constraint for modeling. However, in less explored areas where drillhole data are sparse, the use of a 2D geological diagram without rigorous falsification may bias the model construction. Incorporating geophysical data, such as gravity measurements, enables the falsification of hypotheses of geological diagrams in new exploration areas. By assigning density contrasts to the randomly sampled intrusion models and conducting forward modeling computations, we can generate the predicted geophysical data. A direct comparison between the predictions and the observation is beneficial for falsifying geological diagrams. If the predicted geophysical responses from the models fail to align with the observed data, this discrepancy suggests that the geological diagram may not accurately represent the subsurface structure. Conversely, a good match between predictions and the observation validates the geological diagram.

6. Conclusions

Our innovative framework bridges the gap between geological knowledge and geoscientific data in stochastic modeling. It involves two-dimensional (2D) geological diagrams representing the geological knowledge and accommodates diagrams' variability and uncertainty using Procrustes analysis. The resulting models can align with geological insights and lithologic information from drillholes and outcrop contacts. Applied to the Crystal Lake Gabbro (CLG) in northwest Ontario, the results demonstrate the effectiveness of our framework in modeling geological geometry features.

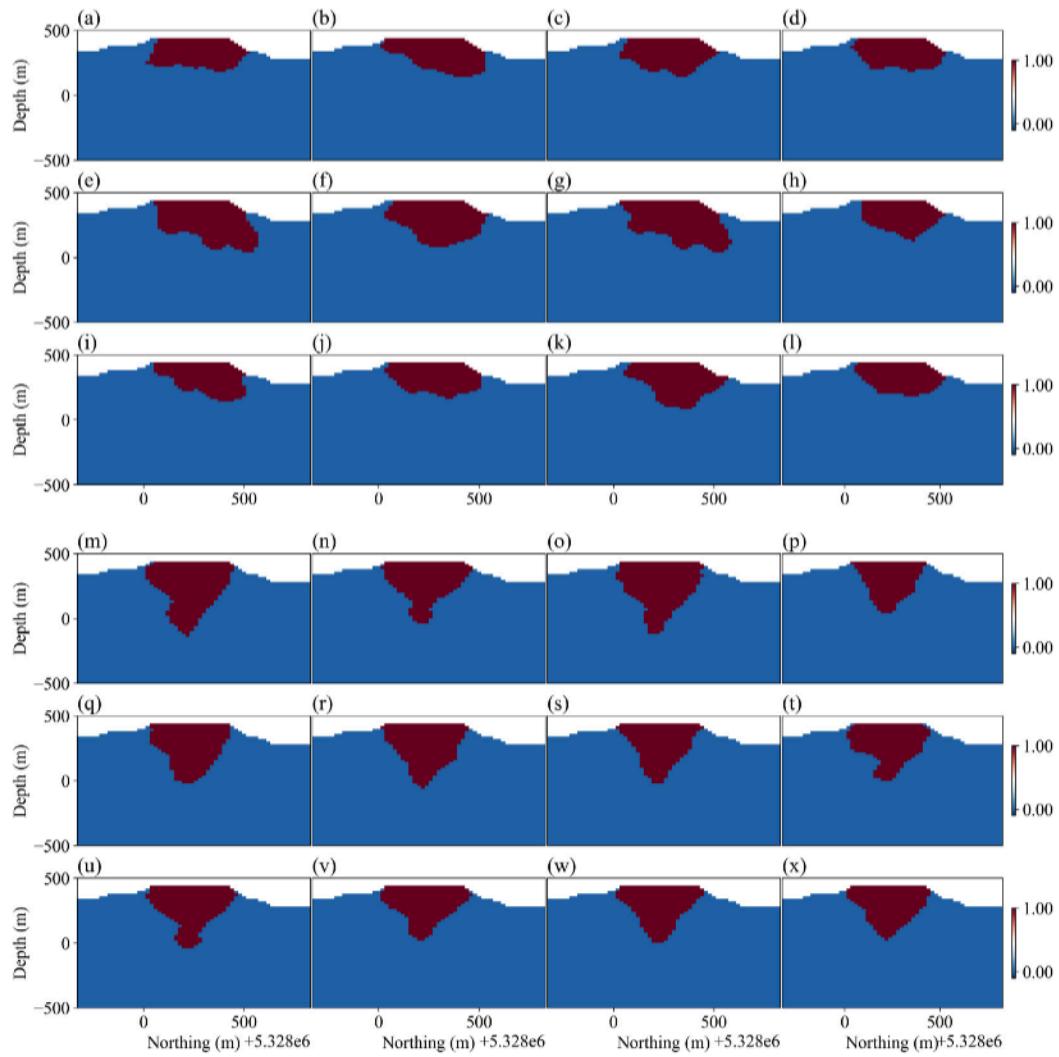


Fig. 15. 2D cross-sections extracted at the east position of 305,180 m from 3D models. Panels (a)–(l) represent models without the geological diagram constraint corresponding to *Case 1: modeling from drillholes and outcrop contacts*. Panels (m)–(x) depict models that incorporate the geological diagram constraint, which associates with *Case 2: knowledge-driven modeling from drillholes and outcrop contacts*.

CRediT authorship contribution statement

Xiaolong Wei: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Zhen Yin:** Writing – review & editing, Validation, Project administration. **Wilson Bonner:** Writing – review & editing, Validation, Resources, Data curation. **Jef Caers:** Writing – review & editing, Supervision, Resources, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Xiaolong Wei reports financial support was provided by Stanford University. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Kobold Metals and Rio Tinto for funding and supporting this research.

Appendix A. Derivative of ordinary procrustes analysis (OPA)

A.1. Objective function

The objective of ordinary Procrustes analysis (OPA) is to find the best-fit transformation (i.e., translation, rotation, and scaling) to map one configuration of points (\mathbf{X}_1) onto another configuration (\mathbf{X}_2). The objective function aims to minimize the Frobenius norm of the difference between \mathbf{X}_1 and \mathbf{X}_2 :

$$L_p(\mathbf{X}_1, \mathbf{X}_2) = \|\beta\mathbf{X}_1\mathbf{R} - \mathbf{X}_2 - \mathbf{C}\|_F^2 \quad (9)$$

A.2. Translation: \mathbf{C}

The initial step in computing the ordinary Procrustes analysis is centering both shapes, which aligns their centroids. The centering guarantees that no further translation is required in subsequent calculations (i.e., $\mathbf{C} = 0$). To align the centroids, we subtract the means of \mathbf{X}_1 and \mathbf{X}_2 from each point. We perform the subsequent analysis with centered matrices, expressed as follows:

$$\mathbf{X}_1^c = \mathbf{X}_1 - \mu_1, \mathbf{X}_2^c = \mathbf{X}_2 - \mu_2 \quad (10)$$

A.3. Rotation: \mathbf{R}

The goal is to find a rotation matrix \mathbf{R} that rotates \mathbf{X}_1^c to align as closely as possible with \mathbf{X}_2^c . Mathematically, we want to minimize the Frobenius norm of the difference between $\mathbf{X}_1^c \mathbf{R}$ and \mathbf{X}_2^c . The essential in solving this optimization problem is the cross-covariance matrix \mathbf{C} between \mathbf{X}_1^c and \mathbf{X}_2^c :

$$\mathbf{C} = \mathbf{X}_1^c \mathbf{X}_2^c \quad (11)$$

A rotation matrix \mathbf{R} should be orthogonal, rotating a set of points without distorting distances and angles. The SVD provides two orthogonal matrices, $\mathbf{C} = \mathbf{U}\Sigma\mathbf{V}^T$. The matrices \mathbf{U} and \mathbf{V} can capture the directions that point sets need to be rotated. The optimal rotation matrix $\hat{\mathbf{R}}$ is:

$$\hat{\mathbf{R}} = \mathbf{V}\mathbf{U}^T \quad (12)$$

where \mathbf{U} rotates the directions of \mathbf{X}_1^c and \mathbf{V} rotates the directions of \mathbf{X}_2^c .

A.4. Scaling: β

To find the optimal scaling factor $\hat{\beta}$, we set the derivative of objective function, Eq. (9), equal to zero:

$$\|\beta\mathbf{X}_1^c \mathbf{R} - \mathbf{X}_2^c\|_F^2 = \beta^2 \text{trace}(\mathbf{R}^T \mathbf{X}_1^c \mathbf{X}_1^c \mathbf{R}) - 2\beta \text{trace}(\mathbf{X}_2^c \mathbf{X}_1^c \mathbf{R}) + \text{trace}(\mathbf{X}_2^c \mathbf{X}_2^c) \quad (13)$$

$$\frac{\partial \|\beta\mathbf{X}_1^c \mathbf{R} - \mathbf{X}_2^c\|_F^2}{\partial \beta} = 2\beta \text{trace}(\mathbf{R}^T \mathbf{X}_1^c \mathbf{X}_1^c \mathbf{R}) - 2\text{trace}(\mathbf{X}_2^c \mathbf{X}_1^c \mathbf{R}) = 0 \quad (14)$$

where \mathbf{R} is an orthogonal matrix and $\mathbf{R}^T \mathbf{R} = \mathbf{I}$. Thus, the optimal scaling factor $\hat{\beta}$ is:

$$\hat{\beta} = \frac{\text{trace}(\mathbf{X}_2^c \mathbf{X}_1^c \hat{\mathbf{R}})}{\text{trace}(\mathbf{X}_1^c \mathbf{X}_1^c)} \quad (15)$$

where $\text{trace}(\mathbf{X}_2^c \mathbf{X}_1^c \hat{\mathbf{R}})$ measures the goodness-of-alignment between \mathbf{X}_1^c and \mathbf{X}_2^c . $\text{trace}(\mathbf{X}_1^c \mathbf{X}_1^c)$ is a measure of the size for \mathbf{X}_1^c .

Appendix B. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.cageo.2024.105779>.

Data availability

Data will be made available on request.

References

- Athens, N., Caers, J., 2022. Stochastic inversion of gravity data accounting for structural uncertainty. *Math. Geosci.* 54 (2), 413–436.
- Barnes, S.J., Cruden, A.R., Arndt, N., Saumur, B.M., 2016. The mineral system approach applied to magmatic Ni–Cu–PGE sulphide deposits. *Ore Geol. Rev.* 76, 296–316.
- Benkó, Z., Mogessie, A., Molnár, F., Severson, M.J., Hauck, S.A., Raič, S., 2015. Partial melting processes and Cu-Ni-PGE mineralization in the footwall of the South Keweenawan intrusion at the Spruce Road deposit, Duluth Complex, Minnesota. *Econ. Geol.* 110 (5), 1269–1293.
- Caers, J., Zhang, T., 2004. Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models.
- Cai, H., Liu, M., Zhou, J., Li, J., Hu, X., 2022. Effective 3D-transient electromagnetic inversion using finite-element method with a parallel direct solver. *Geophysics* 87 (6), E377–E392.
- Cai, H., Long, Z., Lin, W., Li, J., Lin, P., Hu, X., 2021. 3D multinary inversion of controlled-source electromagnetic data based on the finite-element method with unstructured mesh. *Geophysics* 86 (1), E77–E92.
- Caltagno, P., Chilès, J.-P., Courrioux, G., Guillen, A., 2008. Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules. *Phys. Earth Planet. Inter.* 171 (1–4), 147–157.
- Cannon, W.F., 1992. The Midcontinent rift in the Lake Superior region with emphasis on its geodynamic evolution. *Tectonophysics* 213, 41–48.
- Chen, Q., Mariethoz, G., Liu, G., Comunian, A., Ma, X., 2018. Locality-based 3-D multiple-point statistics reconstruction using 2-D geological cross sections. *Hydrol. Earth Syst. Sci.* 22 (12), 6547–6566.
- Cherpeau, N., Caumon, G., Caers, J., Lévy, B., 2012. Method for stochastic inverse modeling of fault geometry and connectivity using flow data. *Math. Geosci.* 44, 147–168.
- Civico, R., Sapia, V., Di Giulio, G., Villani, F., Pucci, S., Baccheschi, P., Amoroso, S., Cantore, L., Di Nuccio, D., Hailemikael, S., et al., 2017. Geometry and evolution of a fault-controlled Quaternary basin by means of TDEM and single-station ambient vibration surveys: The example of the 2009 L’Aquila earthquake area, central Italy. *J. Geophys. Res.: Solid Earth* 122 (3), 2236–2259.
- Creighton, A.L., Parsekian, A.D., Angelopoulos, M., Jones, B.M., Bondurant, A., Engram, M., Lenz, J., Overduin, P.P., Grosse, G., Babcock, E., et al., 2018. Transient electromagnetic surveys for the determination of talik depth and geometry beneath thermokarst lakes. *J. Geophys. Res.: Solid Earth* 123 (11), 9310–9323.
- de la Varga, M., Schaaf, A., Wellmann, F., 2019. GemPy 1.0: open-source stochastic geological modeling and inversion. *Geosci. Model Dev.* 12 (1), 1–32.
- Ding, X., Li, C., Ripley, E.M., Rossell, D., Kamo, S., 2010. The Eagle and East Eagle sulfide ore-bearing mafic-ultramafic intrusions in the Midcontinent Rift System, upper Michigan: Geochronology and petrologic evolution. *Geochem. Geophys. Geosyst.* 11 (3), <http://dx.doi.org/10.1029/2009GC002546>.
- Ding, X., Ripley, E.M., Li, C., 2012. PGE geochemistry of the Eagle Ni–Cu–(PGE) deposit, Upper Michigan: constraints on ore genesis in a dynamic magma conduit. *Miner. Deposita* 47, 89–104.
- Earl, D.J., Deem, M.W., 2005. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.* 7 (23), 3910–3916.
- Flemings, P.B., Jordan, T.E., 1989. A synthetic stratigraphic model of foreland basin development. *J. Geophys. Res.: Solid Earth* 94 (B4), 3851–3866.
- Fouedjio, F., Scheidt, C., Yang, L., Achtziger-Zupančič, P., Caers, J., 2021. A geostatistical implicit modeling framework for uncertainty quantification of 3D geo-domain boundaries: Application to lithological domains from a porphyry copper deposit. *Comput. Geosci.* 157, 104931.
- Frank, R., Tertois, A.-L., Mallet, J.-L., 2007. 3D-reconstruction of complex geological interfaces from irregularly distributed and noisy point data. *Comput. Geosci.* 33 (7), 932–943.
- Geul, J., 1970. Geology of Devon and Pardee Townships and the Stuart Location, District of Thunder Bay. Ontario Department of Mines.
- Goodall, C., 1991. Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 53 (2), 285–321.
- Gower, J.C., 1975. Generalized procrustes analysis. *Psychometrika* 40, 33–51.
- Grose, L., Aillères, L., Laurent, G., Caumon, G., Jessell, M., Armit, R., 2021a. Modelling of faults in LoopStructural 1.0. *Geosci. Model Dev.* 14 (10), 6197–6213.
- Grose, L., Aillères, L., Laurent, G., Jessell, M., 2021b. LoopStructural 1.0: time-aware geological modelling. *Geosci. Model Dev.* 14 (6), 3915–3937.
- Hastings, W.K., 1970. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. Oxford University Press.
- Hou, W., Liu, H., Zheng, T., Shen, W., Xiao, F., 2021. Hierarchical MPS-based three-dimensional geological structure reconstruction with two-dimensional image (s). *J. Earth Sci.* 32 (2), 455–467.
- Hoyer, A.-S., Vignoli, G., Hansen, T.M., Vu, L.T., Keefer, D.A., Jørgensen, F., 2017. Multiple-point statistical simulation for hydrogeological models: 3-D training image development and conditioning strategies. *Hydrol. Earth Syst. Sci.* 21 (12), 6069–6089.
- Huang, J., Deng, H., Chen, J., Li, N., Wang, J., Liu, Z., Mao, X., 2023a. Assessing geometrical uncertainties in geological interface models using Markov chain Monte Carlo sampling via abstract graph. *Tectonophysics* 864, 230032.
- Huang, J., Deng, H., Mao, X., Chen, G., Yu, S., Liu, Z., 2023b. 3D modeling of detachment faults in the Jiaodong gold province, eastern China: A Bayesian inference perspective and its exploration implications. *Ore Geol. Rev.* 154, 105307.
- Jordão, H., Azevedo, L., Sousa, A.J., Soares, A., 2022. Generative adversarial network applied to ore type modeling in complex geological environments. *Math. Geosci.* 54 (7), 1165–1182.
- Jordão, H., Sousa, A.J., Soares, A., 2023. Using Bayesian neural networks for uncertainty assessment of ore type boundaries in complex geological models. *Nat. Resour. Res.* 32 (6), 2495–2514.
- Jørgensen, F., Hoyer, A.-S., Sanderson, P.B., He, X., Foged, N., 2015. Combining 3D geological modelling techniques to address variations in geology, data type and density—An example from Southern Denmark. *Comput. Geosci.* 81, 53–63.
- Lajaunie, C., Courrioux, G., Manuel, L., 1997. Foliation fields and 3D cartography in geology: principles of a method based on potential interpolation. *Math. Geol.* 29, 571–584.
- Lightfoot, P.C., Evans-Lamswood, D., 2015. Structural controls on the primary distribution of mafic-ultramafic intrusions containing Ni–Cu–Co–(PGE) sulfide mineralization in the roots of large igneous provinces. *Ore Geol. Rev.* 64, 354–386.
- Linde, N., Renard, P., Mukerji, T., Caers, J., 2015. Geological realism in hydrogeological and geophysical inverse modeling: A review. *Adv. Water Resour.* 86, 86–101.
- Lindsay, M.D., Aillères, L., Jessell, M.W., de Kemp, E.A., Betts, P.G., 2012. Locating and quantifying geological uncertainty in three-dimensional models: Analysis of the Gippsland Basin, southeastern Australia. *Tectonophysics* 546, 10–27.

- Lindsey, E., Fialko, Y., 2013. Geodetic slip rates in the southern San Andreas Fault system: Effects of elastic heterogeneity and fault geometry. *J. Geophys. Res.: Solid Earth* 118 (2), 689–697.
- Madsen, R.B., Høyar, A.-S., Andersen, L.T., Møller, I., Hansen, T.M., 2022. Geology-driven modeling: A new probabilistic approach for incorporating uncertain geological interpretations in 3D geological modeling. *Eng. Geol.* 309, 106833.
- Mariethoz, G., Caers, J., 2014. Multiple-Point Geostatistics: Stochastic Modeling with Training Images. John Wiley & Sons.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21 (6), 1087–1092.
- Miall, A.D., 2013. Principles of Sedimentary Basin Analysis. Springer Science & Business Media.
- Miller, J., Nicholson, S., 2013. In: Miller, J. (Ed.), Precambrian Research Center Guidebook. pp. 1–50.
- Mota-e-Silva, J., Filho, C.F.F., Giustina, M.E.S.D., 2013. The Limoeiro deposit: Ni-Cu-PGE sulfide mineralization hosted within an ultramafic tubular magma conduit in the Borborema Province, northeastern Brazil. *Econ. Geol.* 108 (7), 1753–1771.
- Mungall, J.E., 2007a. Crustal contamination of picritic magmas during transport through dikes: the Expo Intrusive Suite, Cape Smith fold belt, New Quebec. *J. Petrol.* 48 (5), 1021–1039.
- Mungall, J.E., 2007b. Crystallization of magmatic sulfides: An empirical model and application to Sudbury ores. *Geochim. Cosmochim. Acta* 71 (11), 2809–2819.
- Naldrett, A.J., 2013. Magmatic Sulfide Deposits: Geology, Geochemistry and Exploration. Springer Science & Business Media.
- O'Brien, S., 2018. Petrology of the Crystal Lake Gabbro and the Mount Mollie Dyke, Midcontinent Rift, Northwest Ontario (Unpublished doctoral dissertation).
- Osher, S., Fedkiw, R.P., 2001. Level set methods: an overview and some recent results. *J. Comput. Phys.* 169 (2), 463–502.
- Osher, S., Fedkiw, R., Piechor, K., 2004. Level set methods and dynamic implicit surfaces. *Appl. Mech. Rev.* 57 (3), B15.
- Pakyuz-Charrier, E., Lindsay, M., Ogarko, V., Giraud, J., Jessell, M., 2018. Monte Carlo simulation for uncertainty estimation on structural data in implicit 3-D geological modeling, a guide for disturbance distribution selection and parameterization. *Solid Earth* 9 (2), 385.
- Pardee, D., Townships, C., Rio Tinto Exploration Canada Inc. Assessment Report On the Cu-Ni-PGE Crystal Lake Project.
- Piquer, J., Sanchez-Alfaro, P., Pérez-Flores, P., 2021. A new model for the optimal structural context for giant porphyry copper deposit formation. *Geology* 49 (5), 597–601.
- Ripley, E.M., 2014. Ni-Cu-PGE mineralization in the Partridge River, South Kawishiwi, and Eagle intrusions: a review of contrasting styles of sulfide-rich occurrences in the Midcontinent Rift System. *Econ. Geol.* 109 (2), 309–324.
- Scognamiglio, L., Tinti, E., Casarotti, E., Pucci, S., Villani, F., Cocco, M., Magnoni, F., Michelini, A., Dreger, D., 2018. Complex fault geometry and rupture dynamics of the MW 6.5, 30 October 2016, Central Italy earthquake. *J. Geophys. Res.: Solid Earth* 123 (4), 2943–2964.
- Seat, Z., Beresford, S.W., Grgruric, B.A., Waugh, R.S., Hronsky, J.M., Gee, M.M., Groves, D.I., Mathison, C.I., 2007. Architecture and emplacement of the Nebo-Babel gabbro-hosted magmatic Ni-Cu-PGE sulphide deposit, West Musgrave, Western Australia. *Miner. Deposita* 42, 551–581.
- Severson, M.J., Miller, J.D., Peterson, D.M., Green, J.C., Hauck, S.A., 2002. Report of Investigations - Minnesota Geological Survey, 58. Minnesota Geological Survey, pp. 164–200.
- Shimizu, K., Yagi, Y., Okuwaki, R., Fukahata, Y., 2020. Development of an inversion method to extract information on fault geometry from teleseismic data. *Geophys. J. Int.* 220 (2), 1055–1065.
- Shimizu, K., Yagi, Y., Okuwaki, R., Fukahata, Y., 2021. Construction of fault geometry by finite-fault inversion of teleseismic data. *Geophys. J. Int.* 224 (2), 1003–1014.
- Smith, J., Bleeker, W., Hamilton, M., Petts, D., Kamo, S., Rossell, D., 2020. Timing and controls on Ni-Cu-PGE mineralization within the Crystal Lake Intrusion, 1.1 Ga Midcontinent Rift. Target. *Geosci. Initiative* 5, 37–63.
- Smith, A., Sutcliffe, R., 1987. Keweenawan intrusive rocks of the Thunder Bay area. In: Summary of Field Work. pp. 248–255.
- Song, S., Mukerji, T., Hou, J., 2021. GANSim: Conditional facies simulation using an improved progressive growing of generative adversarial networks (GANs). *Math. Geosci.* 1–32.
- Song, X.-Y., Xie, W., Deng, Y.-F., Crawford, A.J., Zheng, W.-Q., Zhou, G.-F., Deng, G., Cheng, S.-L., Li, J., 2011. Slab break-off and the formation of Permian mafic-ultramafic intrusions in southern margin of Central Asian Orogenic Belt, Xinjiang, NW China. *Lithos* 127 (1–2), 128–143.
- Stamm, F.A., de la Varga, M., Wellmann, F., 2019. Actors, actions and uncertainties: Optimizing decision making based on 3-D structural geological models. *Solid Earth* 10, 2015–2043.
- Strebelle, S., Levy, M., 2008. Using multiple-point statistics to build geologically realistic reservoir models: the MPS/FDM workflow. *Geol. Soc. Lond. Spec. Publ.* 309 (1), 67–74.
- Taranovic, V., Ripley, E.M., Li, C., Rossell, D., 2015. Petrogenesis of the Ni-Cu-PGE sulfide-bearing Tamarack Intrusive Complex, Midcontinent Rift System, Minnesota. *Lithos* 212–215, 16–31. <http://dx.doi.org/10.1016/j.lithos.2014.10.012>.
- Tu, X., Bowles-Martinez, E.J., Schultz, A., 2024. Massively parallel modeling of electromagnetic field in conductive media: An MPI-CUDA implementation on Multi-GPU computers. *Comput. Geosci.* 105710.
- Wang, L., Kitandis, P.K., Caers, J., 2022. Hierarchical Bayesian inversion of global variables and large-scale spatial fields. *Water Resour. Res.* 58 (5), e2021WR031610.
- Wang, L., Peeters, L., MacKie, E.J., Yin, Z., Caers, J., 2023. Unraveling the uncertainty of geological interfaces through data-knowledge-driven trend surface analysis. *Comput. Geosci.* 178, 105419.
- Wei, G., Chen, K., Meng, H., 2023. Bayesian inversion of finite-fault earthquake slip model using geodetic data, solving for non-planar fault geometry, variable slip, and data weighting. *J. Geophys. Res.: Solid Earth* 128 (2), e2022JB025225.
- Yan, S., Blankenship, D.D., Greenbaum, J.S., Young, D.A., Li, L., Rutishauser, A., Guo, J., Roberts, J.L., van Ommen, T.D., Siegert, M.J., et al., 2022. A newly discovered subglacial lake in East Antarctica likely hosts a valuable sedimentary record of ice and climate change. *Geology* 50 (8), 949–953.
- Yang, L., Hyde, D., Grujic, O., Scheidt, C., Caers, J., 2019. Assessing and visualizing uncertainty of 3D geological surfaces using level sets with stochastic motion. *Comput. Geosci.* 122, 54–67.
- Zhou, M.-F., Yang, Z., Song, X., Lesher, C., Keays, R., 2002. Magmatic Ni-Cu-(PGE) sulfide deposits in China. In: The Geology, Geochemistry, Mineralogy, Mineral Beneficiation of the Platinum-Group Elements. Canadian Institute of Mining, Metallurgy and Petroleum.