# Reply to Referee #2's comments and suggestions

In is paper, the authors present a method to propagate data and model uncertainty in a Mineral Prospectivity Mapping workflow. The manuscript is rather well structured and pleasant to read. However, some strong reformulation in the introduction, additional justifications and discussion point are missing. It would make the manuscript stronger in terms of addressed scope and impact. More details are given below.

Re: Thanks for your interests. We have carefully considered all of your suggestions and revised the manuscript accordingly. We sincerely appreciate your thoughtful comments, which have helped improve the quality of the manuscript.

Introduction

Lines 41-42: spread the citations across the specific mentioned techniques?

Re: We have revised the manuscript by distributing the citations to the corresponding techniques.

Lines 62-63: in the context of MPM, or in a more general context (e.g. modelling). These two categories are more general, than the four introduced earlier line 55. I would thus reverse the order of presenting these ways of categorizing uncertainties, from a more general approach (2 categories) to the more specific MPM context (4 categories).

Re: Thanks for your valuable suggestion. We have revised the structure of the Introduction in this revision.

Lines 71-115: this part of the introduction is hard to follow especially if the reader is not familiar with deep learning model for MPM. It would be helpful if the authors could recall the main components considered in MPM (e.g. structural features, degree of mineralization, etc.), the common architecture of such workflows (step 1-inferring layers, step 2- extracting features /, step 3- combining information to produce a map) and which methods are usually involved at each step. Then explain how uncertainty is currently ignored or considered at each step and what methods are available for it. Then clearly summarize the knowledge-gap or current limitations and define what you specifically address in this paper. Given the relatively important number of self-citations, it is important to clearly put in

perspective the contribution and limitations of previous work by the authors to underline the novelty of this approach.

Re: Thanks for this insightful and constructive suggestion. We have reorganized this section of the Introduction for better expression. Specifically, we first introduce the concept of MPM and summarize the evolution of MPM methodologies, ranging from traditional statistical approaches to deep learning algorithms. We then describe the general workflow of deep learning-based MPM and emphasize that many existing deep approaches primarily focus on improving predictive performance while overlooking inherent sources of uncertainty, which may result in overconfident predictions. To address this issue, we review several classical definitions, sources, and classification schemes of uncertainty in MPM. Based on these concepts, we clarify that uncertainty in this study is categorized into data-related uncertainty and model-related uncertainty, and focus on the quantification of data uncertainty, model uncertainty, and their combined effect on total prediction uncertainty in MPM. Subsequently, we systematically review the origins, modeling approaches, recent advances, and limitations of data uncertainty and model uncertainty quantification in MPM. This discussion provides the motivation for developing the proposed uncertainty quantification framework.

2.1.1 Line 143: can you explain why multifractal singularity theory offers a rigorous mathematical foundation to simulate geological evidential layer uncertainty?

Re: Geological evidential layers are commonly constructed by discretizing geological features and transforming them into weighted evidential features using weighting function. This process introduces uncertainty because the weighting functions are typically determined based on geological knowledge and expert judgment rather than objective observations. To simulate the uncertainty associated with this weighting functions, multifractal singularity theory offers a mathematical foundation. However, multifractal singularity theory itself does not directly simulate uncertainty. Instead, it provides the basis for constructing weighting functions by describing the power-law relationship between mineralization intensity and the distance to geological features. The uncertainty arises because one of the key parameters in this weighting function, namely the maximum controlling distance ($d_m$), cannot be determined uniquely from observations. It is typically estimated empirically according to metallogenic models and geological expert judgment. Different reasonable choices of $d_m$ lead to different weighting functions and, consequently, different geological evidential layers. To characterize this parameter

uncertainty, we treat $d_m$ as a stochastic variable and randomly sample it within a predefined interval to generate multiple realizations of the geological evidential layers. Therefore, multifractal singularity theory provides the physically and mathematically justified weighting model, while the stochastic sampling of $d_m$ propagates the uncertainty associated with expert-defined parameters into the evidential layers.

The revised sentence is:

"The fractal dimension has proven effective for quantifying the spatial association between mineral deposits and geological features:

$\rho = c\varepsilon^{-D}$, (1)

where $D$ denotes the fractal dimension and $c$ is a constant. The fractal dimension has proven effective for quantifying the spatial association between mineral deposits and geological features, thereby facilitating the construction of weighting functions for various geological evidential features with the help of GIS. Generally, for a given location $(i, j)$ near a geological feature, areas closer to the feature tend to contain more significant mineralization information and should therefore be assigned higher weights. Conversely, areas farther away are less relevant, or even irrelevant to mineralization, should be assigned lower weights. Accordingly, the weighting function for an evidential feature can be expressed as a piecewise formula:

$$\omega_{ij} = \begin{cases} cd^{-D} & d \le d_m \\ 0 & d > d_m \end{cases}, (2)$$

where $d$ denotes the distance from a given location $(i, j)$ to the geological feature, $d_m$ represents the maximum effective controlling distances of that feature. The weight $\omega$ quantifies the relative contribution of the geological feature to mineralization, with a larger $\omega$ indicating stronger ore-controlling effect.

In practice, the weight $\omega$ can be estimated by fitting a linear relationship to the log–log plot of $\rho$ versus $d$ using the least-squares method:

$\log(\rho) = -\omega \log(d) + c$. (3)

In the formula 2, the controlling distance $d_m$, also known as the spatial extent of a feature's influence, is typically set empirically based on metallogenic models, which vary according to expert interpretation. This subjectivity inevitably introduces data uncertainty into MPM. To account for this uncertainty, the parameter $d_m$ is treated as a stochastic variable and randomly sampled within a predefined interval to

generate multiple realizations of geological evidential features. Consequently, the weighting function for geological features is expressed as:

$$\omega = \frac{1}{1-d_m{}^{-D}}(x^{-D} - d_m), d_m \in [0,1]. \quad (4)$$

This formulation accounts for both the relative influence of geological features on mineralization and the uncertainty in the modeling process of evidential feature. The resulting weight values range continuously from 1 and 0. A weight of 1 indicates a strong statistical correlation between the geological feature and known mineral deposits, whereas decreasing values represent weaker spatial association. When the location $x$ exceeds the maximum controlling distances $d_m$, the weight reduces to 0, implying that the geological feature in those regions exerts no measurable impact on mineralization. After obtaining the weight $\omega$, the geological evidence layer $g_{(\omega)}$ can be constructed as:

$$g_{(\omega)} = g \cdot \omega, \quad (5)$$

where $g$ is the original gridded geological feature and $g_{(\omega)}$ is the weighted geological evidence layer."

2.1.2 What training images are available for geochemical and geophysical evidential layers? In the example given in Figure 2, the TI seems to be a realization from a multi-Gaussian Random Field. What is the advantage of using the Direct Sampling algorithm in such a case? Would Sequential Gaussian Simulations not be sufficient?

Re: In this study, the training image (TI) is a raster map obtained from geochemical sampling points. In other words, the observed geochemical (or geophysical) measurements themselves are used as both the conditioning data and the TI. The primary reason for adopting DS rather than Sequential Gaussian Simulation (SGS) is that DS does not require Gaussian transformation or variogram model. Although SGS is an effective geostatistical simulation technique, it primarily reproduces two-point spatial statistics represented by the variogram and may smooth localized anomalies, thereby fails to reproduce more complex spatial configurations. In contrast, DS is a multiple-point simulation method that reproduces complex spatial patterns directly from the TI through neighborhood matching. Consequently, it can better preserve local spatial structures, nonlinear spatial relationships, and high-value anomalies that are important for MPM. Our objective is not only to interpolate unsampled locations but also to generate multiple realizations that preserve the observed spatial patterns while quantifying uncertainty. Therefore, DS was selected because it better satisfies these requirements without imposing restrictive distributional assumptions.

2.2 Using only a distance criterion to randomly sample negative samples seems a bit limited. Isn't there a risk of creating a bias? Would it be possible to consider other factors? Which ones? A map showing the locations of negative samples (or a sampling density map) would be meaningful to this study to compare later with the results.

Re: Thank you for this valuable comment. In general, negative samples are challenging to define because the absence of known mineralization does not necessarily indicate the absence of undiscovered deposits. Consequently, negative samples are commonly constructed from randomly selected background points under geological constraints, known occurrences of other mineral deposit types, or drill holes with economically insignificant mineralization (Nykänen et al., 2015; Lindsay et al., 2022; Montsion et al., 2024). In this study, we adopted a distance-based random sampling strategy in which negative samples were selected from areas sufficiently distant from known mineral deposits. Carranza et al. (2008) has demonstrated that randomly sampling points from regions located far from known deposits is an available strategy for generating negative samples in MPM uncertainty quantification. Nonetheless, we acknowledge that relying solely on a distance criterion may introduce sampling bias. We have added this discussion to the revised manuscript as a limitation of the current study.

References:

Carranza, E., Hale, M., and Faassen, C.: Selection of coherent deposit-type locations and their application in data-driven mineral prospectivity mapping, Ore Geology Reviews, 33(3–4), 536–558, https://doi.org/10.1016/j.oregeorev.2007.07.001, 2008.

Nykänen, V., Lahti, I., Niiranen, T., and Korhonen, K.: Receiver operating characteristics (ROC) as validation tool for prospectivity models—A magmatic Ni–Cu case study from the Central Lapland Greenstone Belt, Northern Finland, Ore Geology Reviews, 71, 853–860, https://doi.org/10.1016/j.oregeorev.2014.09.007, 2015.

Lindsay, M. D., Piechocka, A. M., Jessell, M. W., Scalzo, R., Giraud, J., Pirot, G., and Cripps, E.: Assessing the impact of conceptual mineral systems uncertainty on prospectivity predictions, Geoscience Frontiers, 13(6), 101435, https://doi.org/10.1016/j.gsf.2022.101435, 2022.

Montsion, R. M., Perrouty, S., Lindsay, M. D., Jessell, M. W., and Sherlock, R.: Development and application of feature engineered geological layers for ranking magmatic, volcanogenic, and

orogenic system components in Archean greenstone belts, Geoscience Frontiers, 15(2), 101759, https://doi.org/10.1016/j.gsf.2023.101759, 2024.

Figure 7: it is not trivial to see the differences between the realizations from one row to the other. A fifth row displaying the standard deviation over the generated ensemble or weighted evidence layer might highlight the variability. A colour bar with the value range is missing for the evidence layers.

Re: Thank you for this valuable suggestion. Figure 7 has been revised to improve its clarity and interpretability. We have added a fifth row showing the standard deviation of generated geological evidential layers, which more clearly illustrates the spatial variability among the realizations. In addition, we have added color bars to the evidential layer maps.

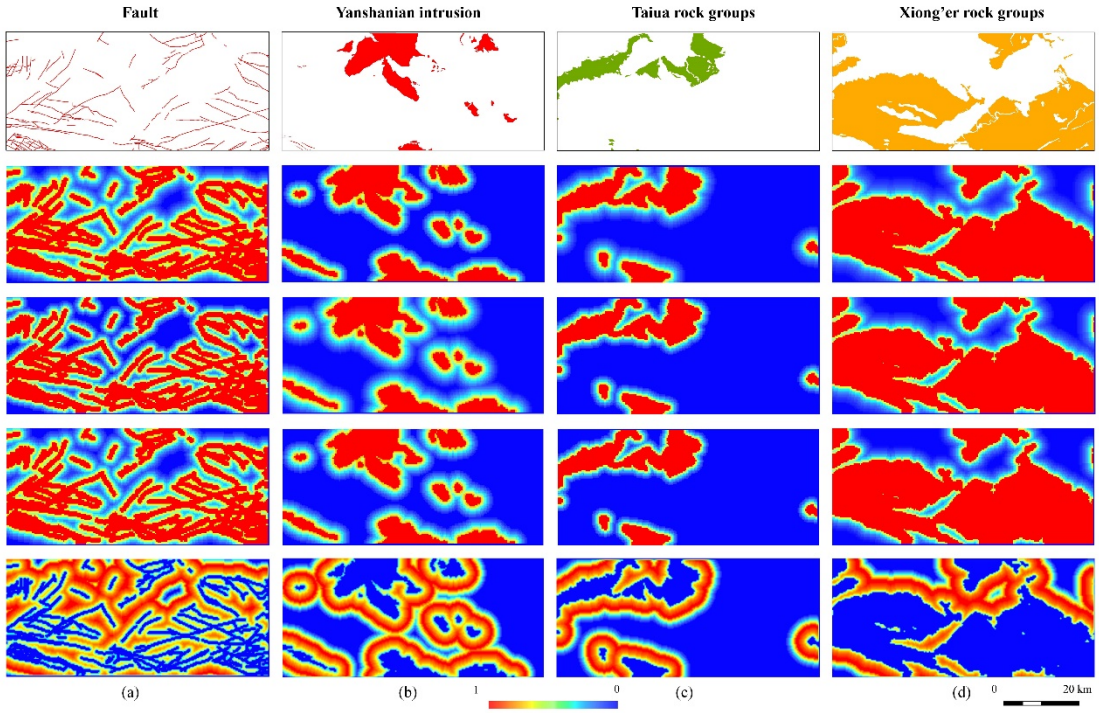


Figure 7: Weighted evidence layers: (a) faults, (b) Yanshanian igneous intrusions, (c) Taihua rock groups strata, (d) Xiong' er rock groups strata. The first row is the original geological features, the second to fourth rows are the weighted evidence layers. The fifth row is the standard deviation of the generated evidence layer, which clearly illustrates the spatial variability among the realizations.

4.2 The Training Images need to be presented and clearly justified (what support the interpretation of such property field structures).

Re: Following previous applications of DS to geoscientific data, the observed geochemical and geophysical datasets are transformed into raster and used as the training images. To justify the suitability of these training images, we compared the simulated and observed evidential layers using cumulative frequency distributions and experimental variograms. Using Au as an example, the cumulative frequency distribution curve (Figure 9a) clearly demonstrate that the DS reproduces the statistical distribution pattern of geochemical features in the study area. The close agreement between the variogram curves of the observed and simulated realizations (Figure 9b) further indicates that the DS approach successfully reproduces the major spatial structures of the original geochemical layer. These results indicate that the selected training images adequately capture the spatial patterns required for stochastic simulation.

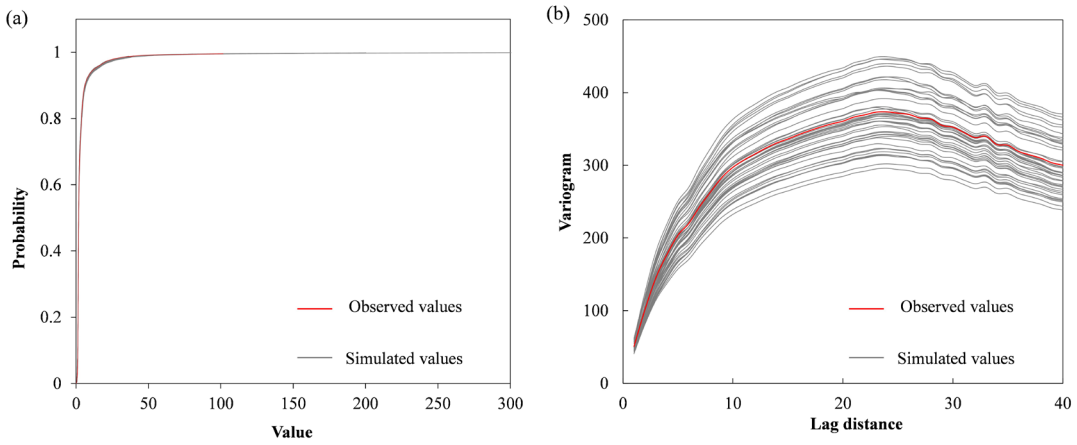


Figure 9: Cumulative frequency distribution curves (a) and experimental variogram curves (b) derived from 100 simulated Au geochemical layers.

Figure 8: though the differences are more visible than in Figure 7, a fifth row displaying the standard deviation of the simulated evidence layers would be great.

Re: Thank you for this valuable suggestion. We have added a fifth row showing the standard deviation of generated geochemical and geophysical evidential layers.

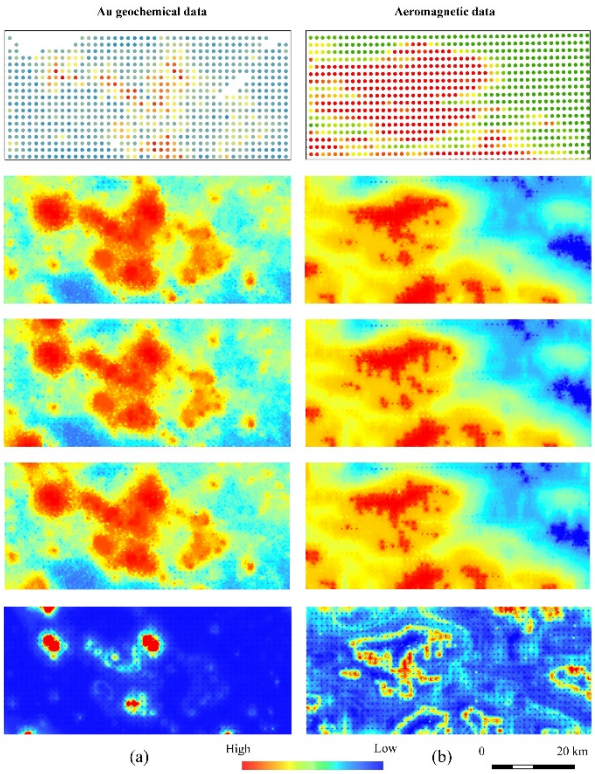


Figure 8: Simulated evidence layers: (a) Au geochemical layer, (b) geophysical aeromagnetic layer. The first row is the original geological features, the second to fourth rows are the simulated evidence layers. The fifth row is the standard deviation of the simulated evidence layer, which illustrates the spatial variability among the realizations.

5.1: how were the key hyper-parameters determined? How do you justify the values employed?

Re: The key hyperparameters are adjusted through grid search strategy. The optimal values are selected based on the best predictive performance and convergence curves of training loss and accuracy across epochs. We have added a description of the hyperparameter tuning in the revised manuscript.

Section 6

Line 405: 'Results' instead of 'Result'

Re: Revised.

Lines 452-453 "Overall, areas exhibiting high mean and low entropy values represent the most reliable exploration targets": Could you provide such an exploration map e.g. by multiplying the mean map with (1-entropy map)?

Re: Thanks for this valuable suggestion. We have added an exploration reliability map by combining the normalized mean prospectivity map with the entropy map.

To facilitate exploration target prioritization, an exploration reliability map (Figure 18) is delineated by integrating both the mean prediction probability with predictive uncertainty:

$R(p) = \bar{p}^* - H(p)^*$,

where $\bar{p}^*$ and $H(p)^*$ denote the normalized mean predictive prospectivity and normalized predictive entropy, respectively. This normalization places both variables on the same scale, allowing prediction potential and prediction reliability to be considered simultaneously. Areas with high values represent locations correspond to both high prospectivity and low uncertainty, and are therefore regarded as the most reliable exploration targets. Compared with the mean prospectivity map alone, this map suppresses regions with high prediction uncertainty while preserving areas with consistently high prospectivity, thereby providing a more reliable basis for exploration decision-making.

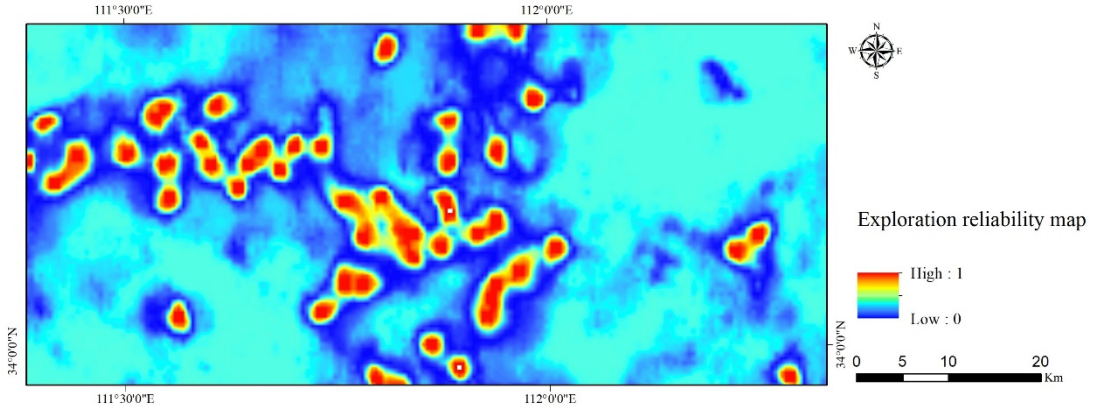


Figure 18: The exploration reliability map.

A single figure grouping figures 12, 13 and 14 might facilitate the comparison of the results.

Re: Thank you for the helpful suggestion. We carefully considered combining Figures 12-14 into a single composite figure. However, we decided to retain them as separate figures for two reasons. First, the three figures present the results of data uncertainty quantification, model uncertainty quantification, and prediction uncertainty quantification, respectively, which correspond to different sections of the Results and are discussed independently. Maintaining separate figures allows each uncertainty component to be presented and interpreted more clearly within its corresponding context. Second, each figure contains multiple high-resolution spatial maps. Combining them into a single figure would substantially increase the number of subfigures, reduce the display size of each panel, and compromise the readability of important spatial patterns and uncertainty distributions. Nevertheless, following the

reviewer's suggestion, we have strengthened the connections among Figures 12-14 in the revised manuscript.

6.3 While 100 evidence layers realizations are used to quantify data uncertainty, how many models are generated using the Bayesian CNN for model uncertainty quantification? 100? from a single (original) set of evidence layers?

Re: Yes. This process is equivalent to sampling 100 subnetworks from a trained Bayesian CNN for model uncertainty quantification. Specifically, during inference, dropout layers were kept active, and 100 stochastic forward passes were performed on the same set of original evidence layers. Each forward pass samples a different thinned subnetwork by randomly deactivating neurons, which can be interpreted as sampling from the posterior distribution over the network parameters. The resulting ensemble of 100 predictions approximates the posterior predictive distribution, from which the predictive mean, variance, and entropy were calculated to quantify model uncertainty.

6.4 Where do the 10,000 realizations come from? 100 stochastic sets of evidence layers x 100 Bayesian CNN models?

Re: The 10,000 predictions do not arise from training 100 independent Bayesian CNN models with 100 stochastic sets of evidence layers. Instead, they are obtained by combining the stochastic sampling of both data and model uncertainties. Specifically, 100 stochastic realizations of the evidence layers are first generated to represent data uncertainty. For each realization, the same trained Bayesian CNN is evaluated using 100 Monte Carlo Dropout stochastic forward passes, resulting a total of $100 \times 100 = 10{,}000$ predictions.

7.1, the uncertainty decomposition should also be compared and discussed with figures 12 and 13. What are the similarities or differences? Why is it valuable to perform this decomposition rather than relying on data uncertainty and model uncertainty as generated in Figures 12 and 13?

Re: We have expanded the discussion to compare the uncertainty decomposition results (Figure 15) with the uncertainty maps shown in Figures 12 and 13.

Uncertainty decomposition analysis enables the partitioning of total predictive variance into components attributable to data variability and model stochasticity, thereby providing a clearer understanding of uncertainty sources and supporting more informed exploration decisions. Figures 12

and 13 characterize the spatial distributions of data uncertainty and model uncertainty, respectively. In contrast, uncertainty decomposition analysis does not represent uncertainty magnitude itself but decomposes the total prediction uncertainty into the relative contributions of data uncertainty and model uncertainty. This contribution map additionally reveals which source dominates the overall uncertainty at each location which cannot be obtained by Figures 12 and 13 independently. In the uncertainty decomposition maps, areas dominated by data uncertainty suggest that improving the quality or representativeness of the input evidence layers would most effectively reduce uncertainty. Conversely, areas dominated by model uncertainty indicate that improvements in model architecture, training strategy, or parameter estimation may be more beneficial.

7.2 For this specific study case, prospective areas seem to be in the vicinity of existing mines. Is it related to the location of the negative samples? What is learned specifically for this area from the uncertainty and prospectivity maps from previous MPM.

Re: The predicted high-prospectivity areas are clustered around known Au deposits. However, this pattern is not primarily related to the selection of negative samples. Negative samples were selected from non-mineralized areas while avoiding the vicinity of known deposits to reduce the likelihood of including undiscovered mineralization.

In the study area, previous mineral exploration has already discovered some Au deposits. The proposed uncertainty framework provides additional information by quantifying the reliability of these predictions. Specifically, areas exhibiting both high prospectivity and low uncertainty represent the most reliable exploration targets and can therefore be prioritized for exploration. In contrast, high-prospectivity areas with elevated uncertainty should be prioritized for additional geological exploration investment. Furthermore, the uncertainty decomposition indicates that prediction uncertainty is dominated by data uncertainty rather than model uncertainty, suggesting that improving the quality and representativeness of the evidence layers would be more effective than further optimizing the prediction model. Therefore, the uncertainty analysis not only improves the interpretability of the prospectivity map but also provides practical guidance for exploration risk assessment and future data acquisition.