RESEARCH ARTICLE

# Uncertainty assessment of a 3D geological model by integrating data errors, spatial variations and cognition bias

Dong Liang[1] · WeiHua Hua[1] · Xiuguo Liu[1] · Yabo Zhao[1] · Zhipeng Liu[1]

## Abstract

A 3D geological structural model is an approximation of an actual geological phenomenon. Various uncertainty factors in modeling reduce the accuracy of the model; hence, it is necessary to assess the uncertainty of the model. To ensure the credibility of an uncertainty assessment, the comprehensive impacts of multi-source uncertainties should be considered. We propose a method to assess the comprehensive uncertainty of a 3D geological model affected by data errors, spatial variations and cognition bias. Based on Bayesian inference, the proposed method utilizes the established model and geostatistics algorithm to construct a likelihood function of modeler's empirical knowledge. The uncertainties of data error and spatial variation are integrated into the probability distribution of geological interface with Bayesian Maximum Entropy (BME) method and updated with the likelihood function. According to the contact relationships of the strata, the comprehensive uncertainty of the geological structural model is calculated using the probability distribution of each geological interface. Using this approach, we analyze the comprehensive uncertainty of a 3D geological model of the Huangtupo slope in Badong, Hubei, China. The change in the uncertainty of the model during the integration process and the structure of the spatial distribution of the uncertainty in the geological model are visualized. The application shows the ability of this approach to assess the comprehensive uncertainty of 3D geological models.

**Keywords** 3D geological model · Multi-source uncertainties · Uncertainty assessment · Integration · Bayesian

## Introduction

A 3D geological structural model can describe complicated geological phenomena in an intuitive way. In general, the accuracy of the geological structural model is affected by many factors, such as the geological complexity, the measurement errors, the sparsity of samples, the subjectivity of the modeler, the limitations of the modeling method and the software, and so on. These factors result in inevitable uncertainties in the geological models. A 3D geological model can be only an approximate description of objective geological phenomena. To further revise and improve a model, it is necessary to determine the uncertainty factors that can affect the model.

To date, many efforts have been made to analyze and quantify the uncertainties of geological models (Chilès et al. 2004; Caumon and Journel 2005; Bistacchi et al. 2008; Suzuki et al. 2008; Jessell et al. 2010; Wellmann and Regenauer-Lieb 2012; Thiele et al. 2016; Schneeberger et al. 2017; Edwards et al. 2017). Uncertainties in 3D geological models can be classified into three types, according to their different sources (Mann 1993; Bárdossy and Fodor 2004; Wu et al. 2005; Wellmann et al. 2010; Zhu and Zhuang 2010; Bond et al. 2011; Lindsay et al. 2012; Bond 2015). (1) Data error (Fig. 1a), e.g., measurement error in observations and accuracy loss in data processing. (2) Spatial variation (Fig. 1b). This type of uncertainty is derived from the random process defined in the modeling algorithm to describe geological phenomena. Similar conditions may produce different results. (3) Cognitive uncertainty (Fig. 1c). Cognitive uncertainty is mainly caused by a geologist's cognitive bias or incomplete knowledge, e.g., empirical knowledge bias in the delimitation of ore bodies.

Many studies of uncertainty analysis have focused only on uncertainties from specific sources, or have ignored the interactions between different sources. This makes it hard to precisely estimate the comprehensive impact from multi-source

✉ Xiuguo Liu
liuxiuguo@cug.edu.cn

[1] School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, People's Republic of China
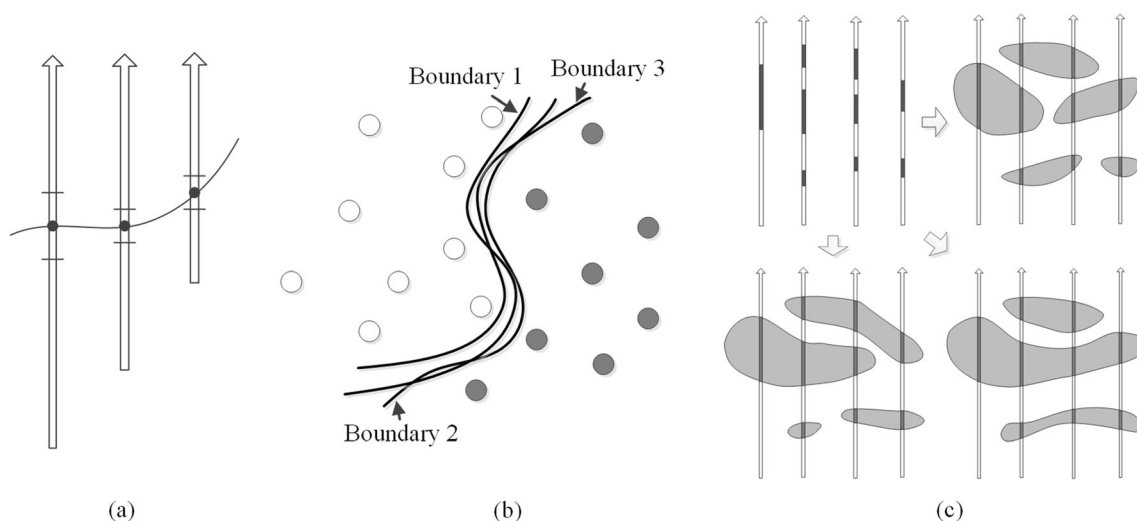
**Fig. 1** Multi-source uncertainties in geological models: (**a**) Data error in contact measurements; (**b**) different boundary realizations via a stochastic simulation with the same data; (**c**) Diverse ore bodies based on the same observations

uncertainties on model accuracy. To determine a more accurate assessment of this impact, some researchers have considered multiple uncertainty sources and have integrated the uncertainties from different sources (Riddick et al. 2005; Lelliott et al. 2009; Lark et al. 2013; Pakyuz-Charrier et al. 2018). Numerous geostatistical modelling methods, such as Bayesian Maximum Entropy (Christakos 1990), Bayesian sequential Gaussian simulation (Doyen and Den Boer 1996), kriging with measurement errors (Fazekas and Kukush 2005), and so on, can process data with measurement errors. These methods can also be used to evaluate the comprehensive uncertainty of data error and the spatial variation in their posterior distribution (e.g., Li et al. 2013), various realizations (e.g., Thore et al. 2002), or estimation error (Kang et al. 2017). In addition to the methods mentioned above, geostatistical methods can also be combined with stochastic simulation to assess the comprehensive uncertainty (Wellmann et al. 2010; Wellmann and Regenauer-Lieb 2012; Røe et al. 2014; Schweizer et al. 2017; Hou et al. 2017; Soares et al. 2017; Pakyuz-Charrier et al. 2019). Although cognitive uncertainty is difficult to evaluate (Bárdossy and Fodor 2001), some researchers have presented uncertainty assessment methods that consider a geologist's cognition (Wellmann and Regenauer-Lieb 2012; Wellmann et al. 2017; de la Varga and Wellmann 2016; Demyanov et al. 2019). To assess the impact of spatial variation and cognitive bias, Tacher et al. (2006) assumed that a geological model is the best guess based on modeler's cognition, and used a Gaussian random field around the expected model to evaluate the comprehensive uncertainty. This is a very useful method to effectively mine information from the established model. To satisfy the hypotheses of the respective methods used, many uncertainty assessment studies have more or less ignored the effects of data error and the modeler's subjective cognition. This simplification of uncertainty factors

may affect the outcome of the assessment. Therefore, it is necessary to conduct an overall assessment of multi-source uncertainties.

In this paper, an uncertainty assessment of 3D geological models that uses integrated data error, spatial variation and cognition is proposed. This method is applied to the uncertainty assessment of an established 3D geological model of the Huangtupo Slope in Badong, Yichang. In the following sections, we will analyze the sources of the uncertainties of the model and introduce a geological interface random function to formulize multi-source uncertainties. In a Bayesian framework, we will integrate the various data errors, the spatial variation and the geologist's cognition into the posterior probability distribution of the geological interface. Then, the probability distributions of the geological interfaces are converted into a series of conditional probabilities on the basis of the contact relationships of strata. Lastly, to assess the uncertainty at any specific location in the geological model, the probability field for individual stratum and the comprehensive uncertainty for the entire model are calculated.

## Materials and method

### Data and model description

The Huangtupo slope is located on the south limb of the Guandongkou syncline in the Three Gorges reservoir area, on the south bank of the Yangtze River in Badong, Hubei, China. The facility sliding strata at this location pose a landslide threat to the residents and towns on the Huangtupo slope. Huangtupo Landslide is a complex landslide mass formed after multiple slumps (including Riverside slump mass 1#, Riverside slump mass 2#, Garden Spot Landslide and

Substation Landslide) covering an area of 1.35×km$^2$ with a volume of 69 million m$^3$. The strata on the Huangtupo slope mainly belong to the Badong Formation, Triassic system (T$_2$b): the slump mass of Huangtupo is developed in slip stratum of the Middle Triassic Badong Formation Section 2 (T$_2$b$^2$) and Section 3 (T$_2$b$^3$), mainly composed of mud rock, pelitic siltstone and muddy limestone (Yang and Chen 2001). The Huangtupo slope shows the alternations of soft and hard rocks:

**Fig. 2** Geological map and location of the Huangtupo slope
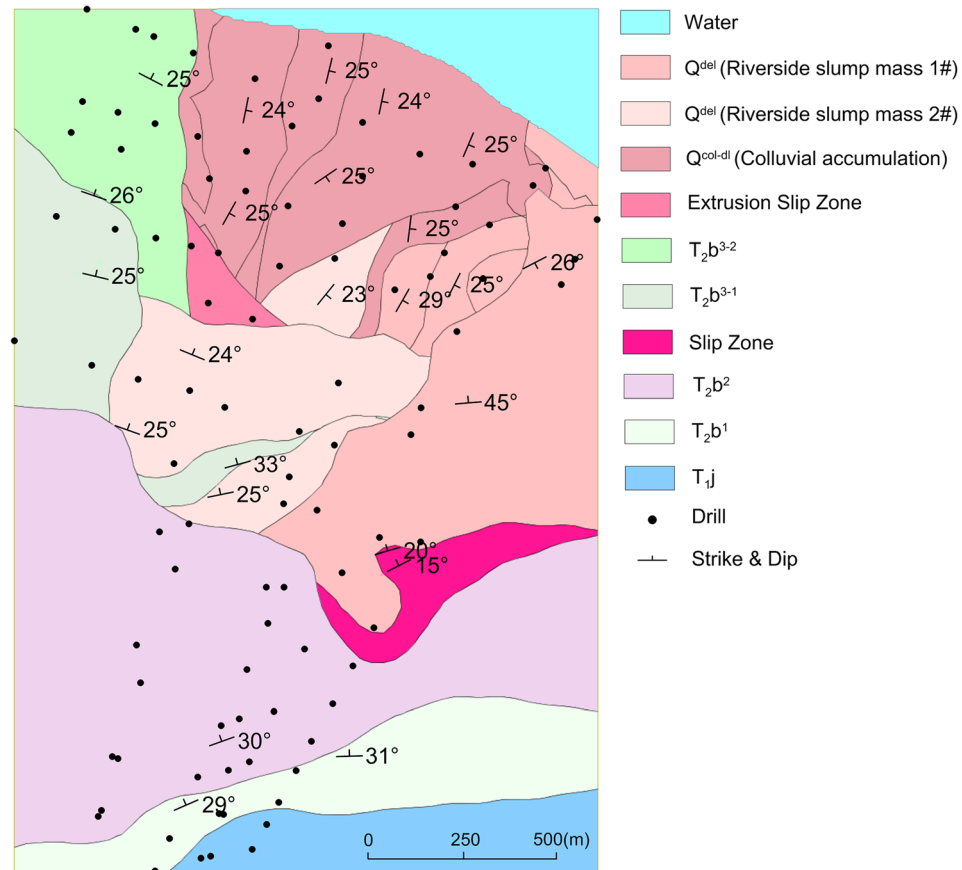
upper soft rock ($T_2b^2$), medium hard rock ($T_2b^{3-1}$) and lower soft rock ($T_2b^{3-2}$) (Hu et al. 2012). The surface of underlying bedrock in the landslide area is undulating. The rock structure of the slide body presents fragmented. The geological map and location of the Huangtupo slope are shown in Fig. 2.

Many detailed geological investigations have been conducted in this area. A 3D structural model of the Huangtupo slope was built in the previous geological modeling work to assist the geological hazard analysis (as shown in Fig. 3 and Fig. 4). However, many uncertainty factors may limit the accuracy of the subsequent analysis. Therefore, it is necessary to determine the impact of these uncertainties on the established

model. The model was built by modeling software with the artificial auxiliary. The modeler ignored data errors in the modeling process. The model mainly used the contact information for 22 strata from 95 boreholes, 6 sections, and a detailed geological map. Because the area has few observational data, some virtual boreholes were added based on the modeler's expertise. The size of the study area is 1530 × 2265 × 702 m. Because of the absence of initial modelers and relevant record, the details of the parameter settings used in the modeling procedure are unknown. We can only obtain observational data and an established model to use in the uncertainty assessment in this paper.
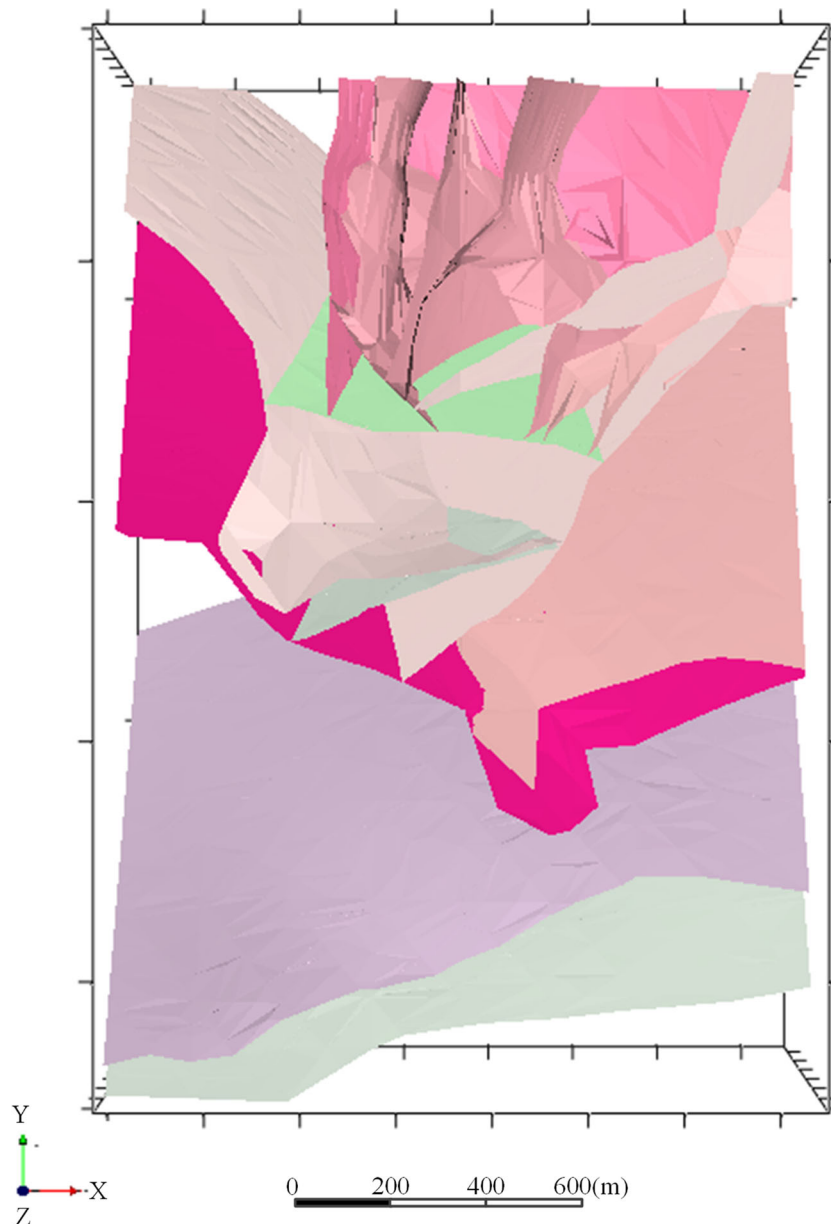


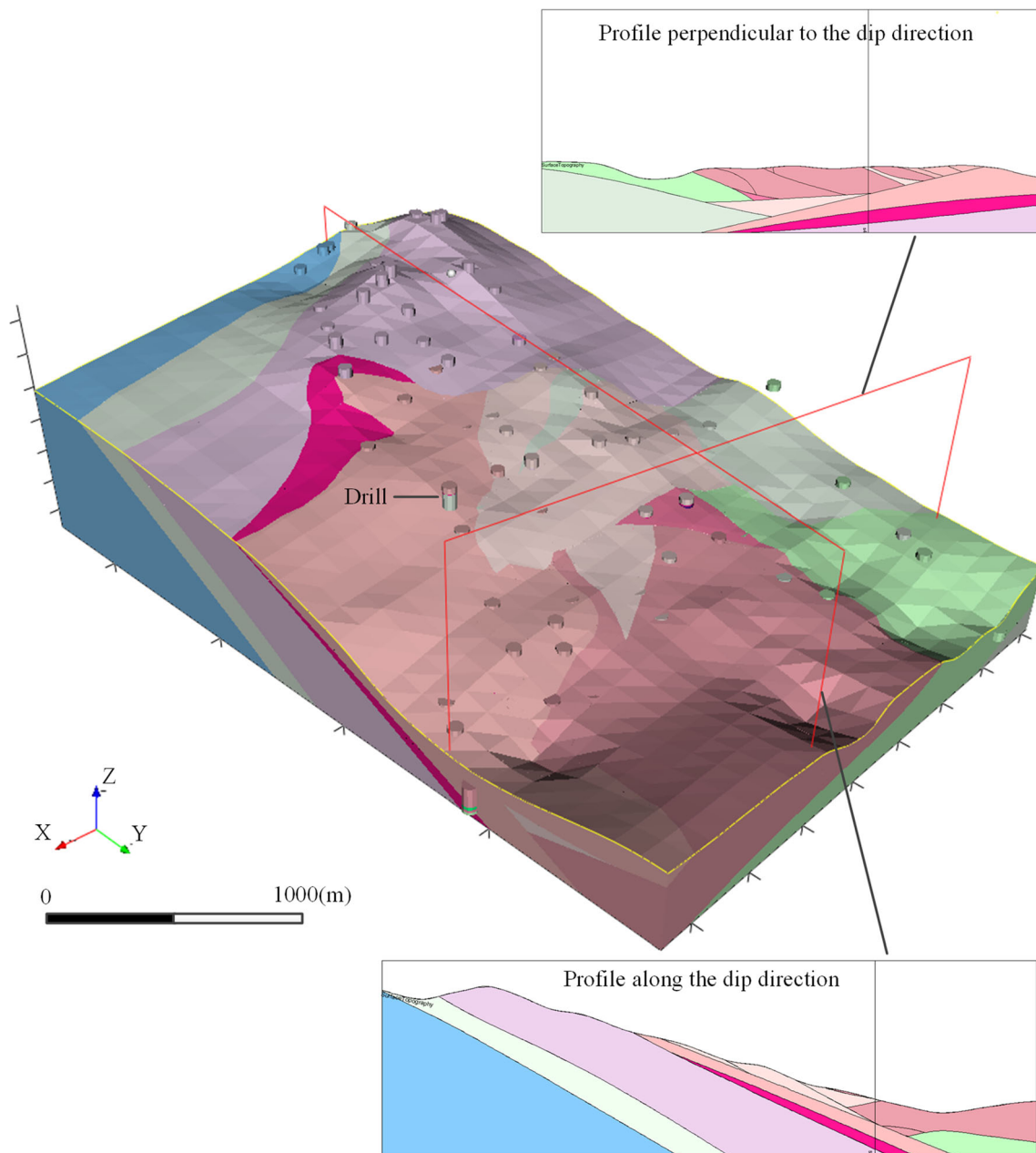**Fig. 3** Geological subsurface model of the Huangtupo slope

**Fig. 4** Two geological profiles of the Huangtupo slope model

## Source analysis of the uncertainty of the model

Before the uncertainty assessment is conducted, we need to ascertain the sources of uncertainty in our study. The 3D geological model is built in three stages: data acquisition, data processing, and structural modeling (Fig. 5). The observations from the measurements are processed into a dataset. Within the constraints of the available geological information and the dataset, the geological interfaces are modeled and assembled into a geological structural model. During data acquisition, measurement error is inevitable. The data processing procedure may cause accuracy loss. In

the procedure for structural modeling, the randomness of spatial variation and the cognitive uncertainty of the modeler are introduced. All of these uncertainties propagate and cumulate as the modeling progresses.

To quantify the impact of multi-source uncertainty factors on the geological model, a random function that describes the probable events affecting the geological structure should be defined. We defined a geological interface random function based on Abrahamsen and Omre's work (Abrahamsen and Omre 1994). First, for each interface, a coordinate system $(U, V, Z)$ is defined according to the occurrence of formation: U is parallel to the strike direction,
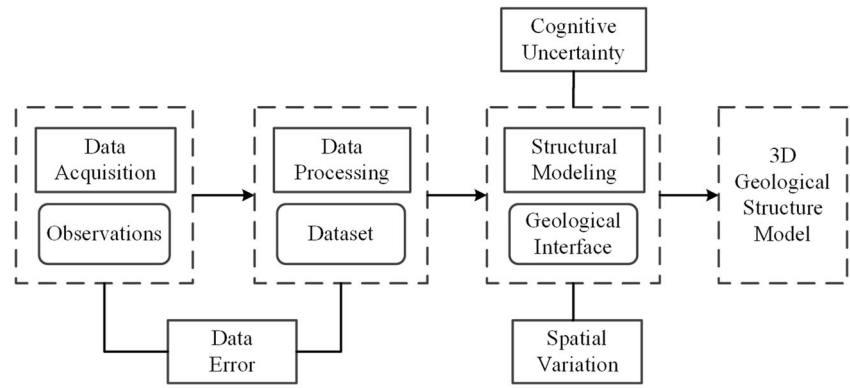
V is parallel to the dip direction and Z is perpendicular to the UV plane and oriented upwards. The UV plane is horizontal, and the Z-direction is vertical. The local coordinate system $(U, V, Z)$ and global coordinate system $(X, Y, Z)$ adopt the same origin. The coordinates in the $(U, V, Z)$ coordinate system can be obtained by the rotation transformation according to the azimuth of dip direction. Then, the random function $Z(u)$ is defined as the elevation value of the geological interface at the location $u = (u, v)$. The purpose of coordinate transformation is to consider the anisotropy of spatial variation. Since the trend of different interfaces are not the same, each interface should be handled individually. A diagram of the random function $Z(u)$ in the interface between strata $T_2b^2$ and $T_2b^1$ is shown in Fig. 6.

In geostatistics, the spatial variation in the geological structures presents two types of characteristics: a structural trend and random fluctuations. The structural trend is caused by the spatial continuity of geological structures. The randomness reflects the measurement error and the spatial variation. The geological model can be seen to be a mixture of deterministic and random information (Abrahamsen et al. 1991). Therefore, $Z(u)$ consists of two distinct parts: a local drift $m(u)$ and a local variation $\sigma(u) \cdot \varepsilon(u)$.

$$Z(u) = m(u) + \sigma(u) \cdot \varepsilon(u). \tag{1}$$

Where $m(u)$ is the most probable value that $Z(u)$ will have at the location $u$ and $\sigma(u) \cdot \varepsilon(u)$ expresses the random

**Fig. 6** Random function $Z(u)$ of the interface between the strata $T_2b^2$ and $T_2b^1$. The enlarged section shows the distribution of elevation value of the geological interface in the plane of $U = u_0$
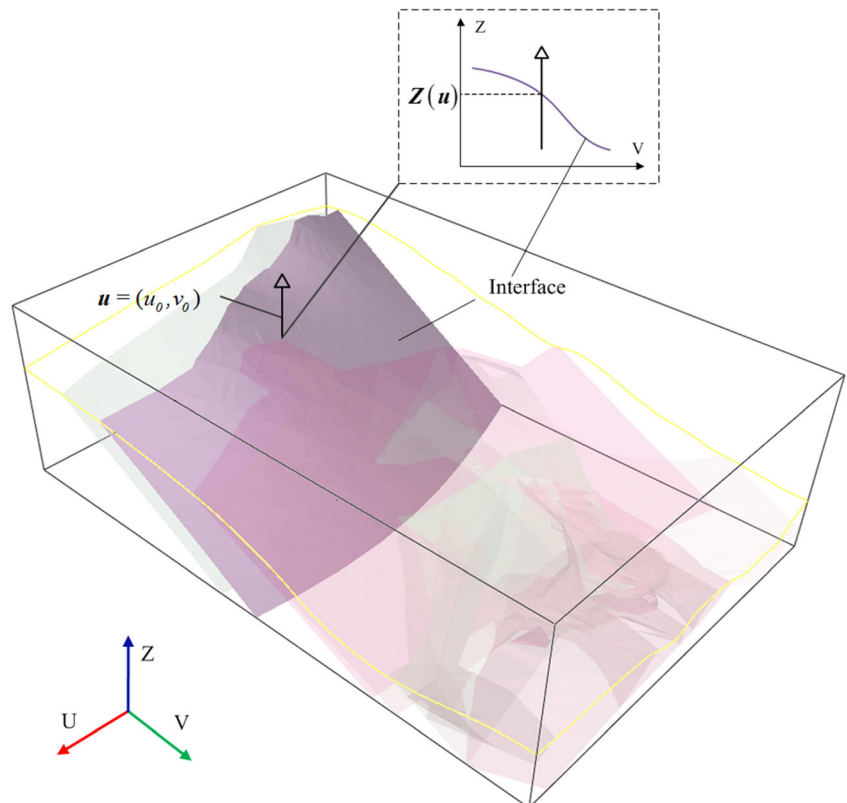
fluctuation around the best guess, $m(\boldsymbol{u})$. $m(\boldsymbol{u})$, is inferred by the geologist based on his knowledge of the geological phenomenon. $\sigma(\boldsymbol{u})$, is the standard deviation of $Z(\boldsymbol{u})$ at the location $\boldsymbol{u}$, which means the fluctuation amplitude of the random distribution. $\varepsilon(\boldsymbol{u})$, is a random function, which $Var\{\varepsilon(\boldsymbol{u})\}=1$. The distribution characteristic of $\varepsilon(\boldsymbol{u})$ depends on the spatial correlation structure in the local area surrounding the location $\boldsymbol{u}$. Because of the insufficient information, we usually simply assumed that $E\{\varepsilon(\boldsymbol{u})\}=0$ and the distribution of $\varepsilon(\boldsymbol{u})$ follows Gaussian distribution. However it should be noted that $\varepsilon(\boldsymbol{u})$ is not limited only to subject to Gaussian distribution(Gunning 2000). $m(\boldsymbol{u})$, $\sigma(\boldsymbol{u})$ and $\varepsilon(\boldsymbol{u})$ are spatially dependent, these terms may change with location $\boldsymbol{u}$.

Since $Z(\boldsymbol{u})$ is a random function describes the geologic structure, the three types of uncertainty we mentioned in the introduction also exist in the subterms of the random function $Z(\boldsymbol{u})$:

1. Data error. $\sigma(\boldsymbol{u})$ is the random amplitude of the summation of the spatial variation and the data error at the location $\boldsymbol{u}$. $\sigma(\boldsymbol{u})$ is partially composed of data errors. The rest of the composition of $\sigma(\boldsymbol{u})$ reflects the randomness of spatial variation.

2. Spatial variation. From the viewpoints of the probability and statistics, the observations of geological phenomena are the outcome of a regionalized random process (Tacher et al. 2006). In the unobserved area, the distribution of geological phenomena is characterized by randomness. In practical applications, we make assumptions about the study variables based on a certain application scenario that helps us select the appropriate modeling methods, such as intrinsic assumption, stationary assumption, Gaussian assumption, and so on. Various modeling methods have different theoretical models and assumptions to describe spatial variation; therefore, the descriptions of the randomness and the assumptions used in these theories may take a variety of forms of expression (Arnold et al. 2019). The random function $\varepsilon(\boldsymbol{u})$ reflects the pattern of spatial variation simulated by the chosen modeling method.

3. Cognitive uncertainty. Cognitive bias is a systematic (non-random) error in modeler's mind (Haselton et al. 2005; Ariely 2008). Due to lack of knowledge or scientific ignorance, the subjective bias in the cognition of the modeler will inevitably lead to the bias and limitation in the research methods adopted by the modeler and the consequent conclusions (Mann 1993). The modeling method, modeling software, and all of the components that used as modeling tools are inevitably influenced by the cognition of the modeler. Therefore, the uncertainty caused by cognitive bias is derived not only from the human-computer interaction in the modeling but also from the selection of the modeling method, software parameter setting, and so on. The subjective decisions made by the modeler result in the cognitive uncertainty of

geological model. Although the cognitive uncertainty is difficult to quantify, the established model is based on available observations and the modeler's own subjective understanding. Therefore, a geological model is the most likely instance of the geological structure in modeler's cognition. We assume $m(\boldsymbol{u})$ is chosen by the modeler as the best guess of the geological structure at the location $\boldsymbol{u}$, which the modeler developed after considering the comprehensive information, and is consistent with observations from the sampling locations. The established geological model is an assemblage of the best guess $m(\boldsymbol{u})$ for the entire region. The deviation between the model and the reality reflects the uncertainty caused by cognitive bias. It should be noted that in the cognitive uncertainty, the parts "We don't know what we don't know" and "We cannot know" (Caers 2011), i.e. the "unknown unknowns", are theoretically unknowable on the basis of the given information (Mann 1993; Wellmann et al. 2010). The potential effect of the "unknown unknowns" is outside the scope of this paper.

In error theory, $m(\boldsymbol{u})$ represents the predicted value with a systematic error, and $\sigma(\boldsymbol{u}) \cdot \varepsilon(\boldsymbol{u})$ represents a random error. $m(\boldsymbol{u})$ expresses the most likely result, with a modeler's cognitive bias. $\sigma(\boldsymbol{u})$ expresses the degree of random uncertainty. $\varepsilon(\boldsymbol{u})$ expresses the distribution pattern of random uncertainty. The cognitive bias (as a systematic error) may not reduce the precision of the geological model, but it may change the accuracy of the model. Multi-source uncertainty factors may be correlative. For example, modeling data can introduce measurement error into the randomness of spatial variation. To assess the comprehensive influence of these factors, multi-source uncertainties should be integrated following certain rules.

## Methodology of uncertainty assessment

We know that Bayesian theory is widely used in multi-source information integration (Dowd 2018). In some ways, the propagation and accumulation of uncertainty in a modeling procedure can be identified as a process of information integration (Caers 2011). Thus, we utilize a Bayesian approach to integrate multi-source uncertainties. In this section, uncertainties from data error, spatial variation and the modeler's cognition are considered and integrated.

Statistically, data error and the randomness of spatial variation are all included in random uncertainty. We integrate these two types of uncertainties into the PDF (Probability Density Function) of $Z(\boldsymbol{u})$ based on the Bayesian Maximum Entropy (BME) principle. Although the cognitive uncertainty has a different nature than the random uncertainty (Mann 1993; Ariely 2008), both of them can be quantified as probabilities in Bayesian theory. Hence we adopt a Bayesian Inference (BI) method to update the BME PDF. A posterior PDF is calculated

for each individual geological interface; thus, it is insufficient for analyzing the uncertainty of the entire model. The posterior PDF of each interface is transformed into a stratigraphic type probability field and integrated by the stratigraphic relationship. Ultimately, the comprehensive uncertainty will have a unified representation; therefore, there will be a holistic picture of the uncertainty in the model. The workflow of the uncertainties assessment is shown in Fig. 7. The details of the workflow will be explained in the following sections.

### Uncertainty integration of data error and spatial variation

In this paper, we adopt the spatial BME method (Christakos 1990), implemented using the BMELib code library to integrate data error and spatial variation into the posterior PDF of every interface. BMElib is a spatiotemporal geostatistics toolbox run in MATLAB (for more details, see reference Christakos et al. 2002). BMElib can be downloaded freely (https://mserre.sph.unc.edu/BMElib_web/). One of the advantages of using this



**Fig. 7** Workflow of the uncertainty assessment

method is that it can integrate data with various errors. Many geostatistical prediction methods are based on the minimum prediction error criterion and an assumption of a Gaussian distribution. However, the multi-source data with the measurement error may not conform to the assumption of a Gaussian distribution. In addition, there may be more than one type of error distribution existent in the data error, such as uniform distributions, Gaussian distributions, Bernoulli distributions, skewed distributions, and so on. Because of the complexity of multi-source uncertainties, many linear integration methods may not be applicable to data with such intricate errors. The BME method can process both hard data (e.g., accurate observations without errors) and soft data (e.g., data with uncertainties, such as interval data or data expressed by PDFs) (Bordwell 2002). BME does not require the data to conform to a Gaussian distribution. Another advantage is that BME is allowed to add many aspects of general knowledge (such as physical laws, expertise, and statistical information) as constraints (Christakos 2017). These non-numerical constraints are treated as prior information of the variables. Then, BME calculates the prior distribution of the maximum information entropy using prior information. The maximum entropy principle can ensure that information will be fully used in the estimation procedure to obtain the optimum result. BME not only generates a single prediction value but also calculates the PDFs of all potential values. Using the sample information (i.e., data from boreholes, sections, geological maps, etc.), the prior distribution will be updated to the BME probability distribution in accordance with the Bayesian inference. If the constraints are only the mean and the variogram of the variable, and the sample data conform to a Gaussian distribution without measurement error, then the solution of the BME will coincide with that given by a simple kriging prediction (Christakos 1990). In this condition, spatial BME and simple kriging are equivalent.

It's necessary to note that, as a method for spatial interpolation in 2D mapping, BME cannot be applied directly to some complex geological settings. To some scenarios, geological structures may have multiple $Z(\boldsymbol{u})$ (elevation value of the boundary) in an $\boldsymbol{u} = (u, v)$. To solve this problem, a strategy of boundary decomposition (Pomian-Srzednicki 2001) is taken into account. In this strategy, the boundaries of complex geological structures such as faults, lenticular bodies, overturned folds, etc. can be divided into different parts in order to satisfy the condition of single $Z(\boldsymbol{u})$. For instance, we can separate the different limbs of an overturned fold into distinct interfaces and divide strata into different parts on either side of fault surface.

In BME, the prior distribution reflects the understanding of a geological structure based on knowledge and experience that exists before the sample data are considered. In addition, the BME PDF represents the new understanding after the field measurement data have been considered. BME provides a tool that can be used to update our understanding of geological structures.
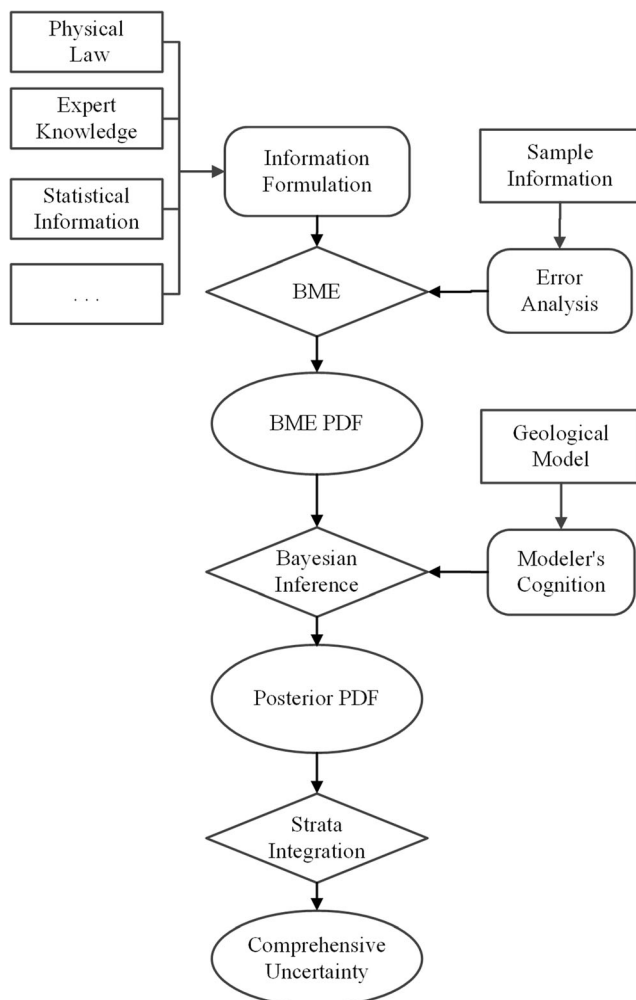
## Uncertainty update considering cognition

The uncertainty of a modeler's cognition is affected by many subjective factors that are difficult to describe quantitatively. We consider the established model as a deterministic instance based on the combination of sample information and modeler's empirical knowledge. The model that incorporates empirical knowledge reflects the modeler's biased cognition. To integrate the cognitive uncertainty into the comprehensive uncertainty, we adopt a Bayesian Inference (BI) method to update the BME PDF with information from modeler's empirical knowledge.

We reference Tacher's assumption (Tacher et al. 2006) about the established model and provide further development. In a practical geological modeling procedure, most of the modeling software and modelers ignore the measurement error in observations. Usually, they just utilize the best estimate of the observations to build the model. In Tacher's methodology, ignoring the data error, the uncertainty in a geological variable can be estimated by substituting the expectation $m_k(\boldsymbol{u})$ of random variables $Z(\boldsymbol{u})$ with the best guess $m(\boldsymbol{u})$, which is the model established by a modeler with a background in geology. In the uncertainty integration, before and after the modeler's empirical knowledge is considered, the uncertainty of $Z(\boldsymbol{u})$ varies as follows:

$$Z_{prior}(\boldsymbol{u}) = m_k(\boldsymbol{u}) + \sigma_k(\boldsymbol{u}) \cdot \varepsilon_k(\boldsymbol{u}) \overset{update}{\rightarrow} Z_{post}(\boldsymbol{u})$$
$$= m(\boldsymbol{u}) + \sigma_k(\boldsymbol{u}) \cdot \varepsilon_k(\boldsymbol{u}). \tag{2}$$

Without the influence of measurement errors, the distribution of $\varepsilon_k(\boldsymbol{u})$ is standard normal distribution, the prior distribution of the $Z(\boldsymbol{u})$ can be expressed as a Gaussian distribution (Abrahamsen 1993; Gunning and Glinsky 2004), in which the expectation is the kriging prediction values $m_k(\boldsymbol{u}) = z_k(\boldsymbol{u})$ and the variance equals the prediction variance $\sigma_k^2(\boldsymbol{u})$, namely $Z_{prior}(\boldsymbol{u}) \sim N\left(z_k(\boldsymbol{u}), \sigma_k^2(\boldsymbol{u})\right)$. The kriging prediction variance $\sigma_k^2(\boldsymbol{u})$ represents the possible random fluctuation of the interface elevation $Z(\boldsymbol{u})$. Suppose $m(\boldsymbol{u}) = z_m(\boldsymbol{u})$ is the most likely elevation value of the interface model at the location $\boldsymbol{u} = (u, v)$, the posterior distribution of $Z(\boldsymbol{u})$ can be expressed as the Gaussian distribution composed of the expectation $z_m(\boldsymbol{u})$ and variance $\sigma_k^2(\boldsymbol{u})$, namely $Z_{post}(\boldsymbol{u}) \sim N\left(z_m(\boldsymbol{u}), \sigma_k^2(\boldsymbol{u})\right)$. The difference between the expectations $z_m(\boldsymbol{u})$ and $z_k(\boldsymbol{u})$ reflects the influence of modeler's cognition. As a systematic error, the cognitive bias will not influence the level of the random fluctuation (i.e. $\sigma_k^2(\boldsymbol{u})$ ).

As mentioned in Section 2.3.1, without data error, the probability distribution of $Z(\boldsymbol{u})$ calculated by the BME method will attain the same Gaussian distribution $N\left(z_k(\boldsymbol{u}), \sigma_k^2(\boldsymbol{u})\right)$. However, considering the various errors in the measurement data, BME may attain a skewed distribution or even a complex distribution with multiple peak values. $m(\boldsymbol{u})$ and $\varepsilon(\boldsymbol{u})$ in the BME estimation may not have explicit expressions. Therefore, it is difficult to determine an optimal solution of $Z(\boldsymbol{u})$. In this situation, it is not appropriate to replace the expectation with the established model. To update the probability distribution of $Z(\boldsymbol{u})$, we choose the Bayesian inference method. To avoid confusion and awkward phrasing, some comments on notation are needed here. In this paper, $p(\cdot)$ represents the marginal distribution. $p(\cdot | \cdot)$ represents the conditional probability density of the parameters mentioned in the context. $P(\cdot)$ represents the probability of an event.

Since the measurement and modeling happen in different stages and are operated by different people. We assume the sample information from measurement and the modeler's subjective empirical knowledge are conditional independent given $Z(\boldsymbol{u}) = z$. The Bayesian inference of the uncertainty of $Z(\boldsymbol{u})$ is as follows:

$$p(Z = z | D, K) \propto p(Z = z | D) \cdot L(Z = z | K) \tag{3}$$

Where $z$ means the possible value in the domain of $Z(\boldsymbol{u})$. $D$ denotes the dataset from sample observations. $K$ denotes the modeler's empirical knowledge which implies the cognitive bias of modeler. $p(Z = z | D)$ is the prior probability density of $Z(\boldsymbol{u}) = z$ before the Bayesian update, which is estimated by geostatistical method based on the observation $D$. The prior distribution $p(Z | D)$ reflects the understanding of the geological structure summarized from statistical information and the modeling algorithm. $p(Z = z | D, K)$ is the posterior probability density of $Z(\boldsymbol{u}) = z$ after considering the modeler's empirical knowledge $K$. The posterior distribution $p(Z | D, K)$ reflects the new understanding after integrating all of the observation information and the subjective cognition in the modeler's mind. Likelihood function $L(Z = z | K) = p(K | Z = z)$ is a function of $Z(\boldsymbol{u})$. $p(K | Z = z)$ denotes the probability density of the empirical knowledge $K$ in the condition of given $Z(\boldsymbol{u}) = z$. $L(Z = z | K)$ represents the likelihood that $Z(\boldsymbol{u}) = z$ is the real interface position value under the given empirical knowledge $K$.

Without considering data error (i.e., Tacher's study), for each location $\boldsymbol{u}$, using the sample data $D$ and the empirical knowledge $K$, the prior probability density $p(Z = z | D)$ and the posterior probability density $p(Z = z | D, K)$ would be calculated as follows:

$$p(Z = z | D) = \frac{1}{\sqrt{2\pi}\sigma_k(u)} e^{-\frac{[z - z_k(u)]^2}{2\sigma_k^2(u)}}, \tag{4}$$

$$p(Z = z | D, K) = \frac{1}{\sqrt{2\pi}\sigma_k(u)} e^{-\frac{[z - z_m(u)]^2}{2\sigma_k^2(u)}} \tag{5}$$

After the update, the expectation of $Z(\boldsymbol{u})$ changes from the kriging prediction value $z_k(\boldsymbol{u})$ to the model value $z_m(\boldsymbol{u})$. According to Eq. (3), we have the likelihood function:

$$L(Z = z|K) = \alpha \cdot \frac{p(Z = z|D, K)}{p(Z = z|D)} \tag{6}$$

Where $\alpha$ denotes a constant as scaling factor.

To assess the uncertainty from measurement error, we estimate the error distribution for each sample and convert inaccurate sample data to soft data. We use $D_e$ to denote the dataset after error assessment. Considering the measurement error, the posterior probability density $p(Z = z|D_e, K)$ can be obtained by:

$$p(Z = z|D_e, K) \propto p(Z = z|D_e) \cdot L(Z = z|K) \tag{7}$$

According to Eq. (6),

$$p(Z = z|D_e, K) \propto p(Z = z|D_e) \cdot \frac{p(Z = z|D, K)}{p(Z = z|D)} \tag{8}$$

As mentioned above, the probability distribution $p(Z|D)$ and $p(Z|D_e)$ can be calculated with geostatistical methods; here we use kriging and the BME method to calculate $p(Z|D)$ (i.e., ignore data errors) and $p(Z|D_e)$ (i.e., consider data errors), respectively. In addition, we adopt the methodology of Tacher et al. (2006), i.e., select the Gaussian distribution composed of model $z_m(\boldsymbol{u})$ and the kriging variance $\sigma_k^2(\boldsymbol{u})$ as $p(Z|D, K)$. After calculation by the proportion in Eq. (A2), a normalization process should be taken to ensure the integral of $p(Z|D_e, K)$ on all $z$ equals to 1.

### Comprehensive uncertainty field calculation

The posterior distribution of $Z(\boldsymbol{u})$ can only describe the spatial uncertainty of the interface between two adjacent strata. However, the occurrences of different interfaces are mutually exclusive in certain locations, and we should consider the impact from the other strata when calculating the uncertainty in the present stratum. We utilize the conditional probability of a stratigraphic type to assess the uncertainty in a certain stratum existing in a given location under the influence of other strata. The occurrence probability of each stratigraphic type is calculated with the CDF (cumulative distribution function) of stratigraphic interface $Z(\boldsymbol{u})$. According to the contact relationship of strata, we update the stratigraphic type probability in an iterative way (Pomian-Srzednicki 2001). Using this method (See comments to modeling method above in Appendix), we obtain a multi-stratigraphic type probability field $P(\boldsymbol{X}, L_i)$ that in each location $\boldsymbol{X} = (x, y, z)$, $P(\boldsymbol{X}, L_i)$ represents the conditional probability of the stratigraphic type $L_i$.

To guide the further adjustment of the geological model, we should know the spatial distribution of the uncertainty of the model. The stratigraphic type probability field $P(\boldsymbol{X}, L_i)$ can quantify the spatial uncertainty in each stratum. However, sometimes we need to not only reveal the uncertainty of each

stratum but also reveal the uncertainty in the integral structure. For this purpose, the probability field of each stratum should be merged into a 3D uncertainty field. Following Wellmann and Regenauer-Lieb (2012), information entropy is applied as the measurement of the integral structure uncertainty in this study. We can calculate the information entropy of the stratigraphic type based on the multi-type probability field $P(\boldsymbol{X}, L_i)$. In any location $\boldsymbol{X} = (x, y, z)$, the information entropy $H(\boldsymbol{X})$ of $n$ stratigraphic types is defined as:

$$H(\boldsymbol{X}) = -\sum_{i=1}^{n} P(\boldsymbol{X}, L_i) \log_b P(\boldsymbol{X}, L_i). \tag{9}$$

Using different values of base $b$, information entropy has different units; if $b = 2$, the unit is a bit. In this paper, 'bit' is adopted as the unit of information entropy. When the probability of each event is equal, the information entropy will be at its maximum, which means that the dispersion degree of the possible result is maximized; therefore, the uncertainty in the prediction will be at its maximum in this situation. If the sample space has only one result, the entropy of this definite event will be 0.

It should be noted that in the probability density functions $p(Z|D)$, $p(Z|D_e)$, $p(Z|D, K)$ and $p(Z = z|D_e, K)$ mentioned in Section 2.3.2, only $p(Z|D, K)$ and $p(Z = z|D_e, K)$ imply the information from the established model. $p(Z|D)$ and $p(Z|D_e)$ are improper to be used as distribution of $Z(\boldsymbol{u})$ in the calculation of probability field of geological model or else the uncertainty fields just express geological uncertainty and have no connection with established model. In other words, the consideration of cognitive bias is inherent in the uncertainty assessment of geological model.

The MATLAB codes of comprehensive uncertainty calculation can be obtained freely (https://github.com/bomer2000/UAGM).

## Results

This section shows the results of an experiment used to illustrate the approach to comprehensive uncertainty assessment of the Huangtupo geological model. The result of the uncertainty assessment is shown below.

### Probability distribution of stratigraphic interface

To analyze the spatial uncertainty of the structural model in each location, we discretized the study area into a 3D field with a resolution of $15 \times 15 \times 1$ m. By analyzing the main tectonic directions, the coordinate system $(U, V, Z)$ was set up.

We analyzed the sample data to evaluate the measurement errors of boreholes and sections. Some of the contact data with few measurement errors were treated as hard data, while other data were converted into probability distributions as soft data.

We considered the quality standard of the data we used, and set up the error distribution for the data by referring to the previous research (Wellmann et al. 2010; Pakyuz-Charrier et al. 2018; Hou et al. 2019). For example, according to the quality standard of the geological survey data (e.g., the measurement error of the borehole should not be higher than 1‰ of the drilling length; in geological profile, the root mean square error of the geological boundary should correspond with the scale of profile), some contact positions from the boreholes and profiles were described with Gaussian distributions (estimated values as mean and their root mean square errors as standard deviation). For the data with direct contact missing, uniform distribution was used to represent the uncertainty of the missing part. For the virtual borehole given by geologist, we made educated guesses based on expert knowledge and used probability distribution to represent the contact point position. The rest of data assumed to contain no errors were treated as hard data.

The experimental covariance function that expresses the spatial variation in the interface elevation was computed using the contact data. Using the experimental covariance fitting and the assistance of expertise, the parameters of the covariance model were determined. We used samples (both soft data and hard data) and their covariance function to calculate the PDF of the interface elevation at each location $(u, v)$ by the BME method. Considering the data error, the randomness pattern of the interface elevation became a non-Gaussian distribution. With the help of the established model and the kriging prediction, the likelihood function was calculated to update the BME PDF. Finally, the posterior PDFs of all of the interfaces were calculated. The comprehensive uncertainty that is expressed in the posterior PDFs integrates the data error, the randomness of spatial variation, and the cognitive uncertainty.

We chose the interface between strata $T_2b^2$ and $T_2b^1$ (as shown in Fig. 6) as an example of the uncertainty integration

of data error and spatial variation. At the interface, the interface elevation $Z(\boldsymbol{u})$ is anisotropic. The covariance functions of the two directions, U and V, are different. The experimental covariance functions and the fitted spherical models of the two directions were calculated, as shown in Fig. 8. For the case in which the sills were observed to be equal in the directions U and V, we set up an anisotropic nested covariance function. The nested structural model is as follows:

$$\begin{cases} C(0) = C_0 + C_1, h = 0 \\ C(h) = C_1 \cdot \left(1 - \frac{3}{2} \cdot \frac{h}{a_u} + \frac{1}{2} \cdot \frac{h^3}{a_u^3}\right), 0 < h \leq a_u \\ C(h) = 0, h > a_u \end{cases} \quad (10)$$

In the nested spherical model, the ranges of two directions are $a_u = 1000$ m and $a_v = 500$ m. The separation distance of isotropic is expressed as $h = \sqrt{h_u^2 + (K \cdot h_v)^2}$, with the anisotropic ratio $K = \frac{a_u}{a_v}$. The sill is the sum of the nugget $C_0 = 0.25$ and the partial sill $C_1 = 25000$. This covariance model reflects structural information and the randomness of the elevation $Z(\boldsymbol{u})$ of the interface between strata $T_2b^2$ and $T_2b^1$.

A location $\boldsymbol{u} = (u_0, v_0)$ (as shown in Fig. 6) is chosen in the study area to show the uncertainty update of the interface between strata $T_2b^2$ and $T_2b^1$. At the location $\boldsymbol{u} = (u_0, v_0)$, the elevation of the interface in the established model is $z_m = 561.2$ m, the kriging prediction of this interface is $z_k = 574.7$ m, and the prediction variance of the kriging is $\sigma_k^2 = 210.3$. In Fig. 9, the red solid curve is the probability distribution $p(Z | D)$ from the kriging prediction. This probability distribution expresses the random uncertainty of the geological variable $Z(\boldsymbol{u})$. The red vertical line is the kriging estimated value $z_k$. The green vertical line represents the elevation $z_m$ of the interface model. The green solid curve is the
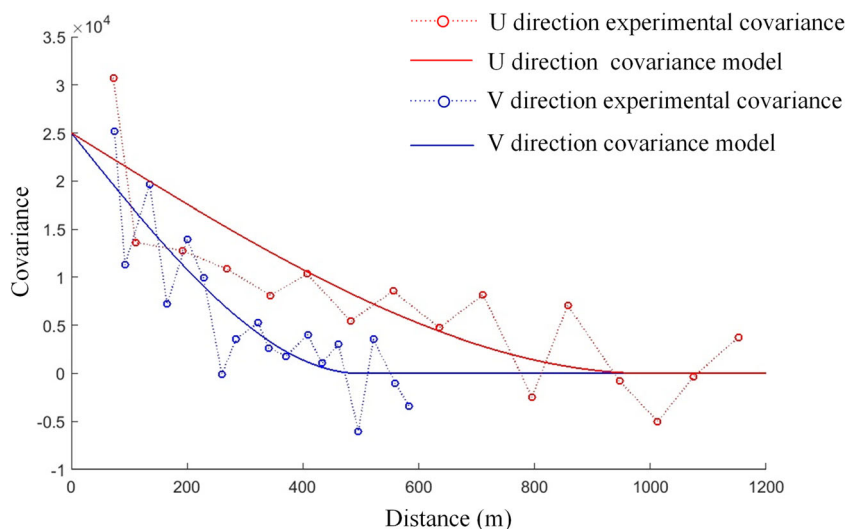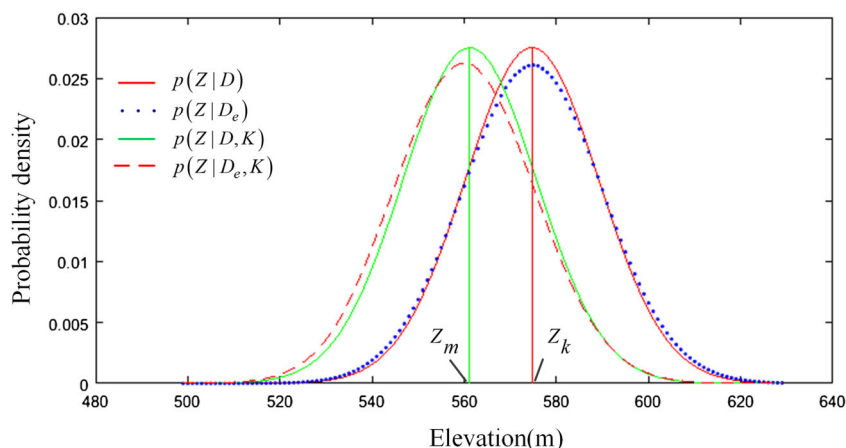
**Fig. 8** Covariance functions in two directions

uncertainty $p(Z|D, K)$ before considering data error. The blue
dotted curve is the BME prediction $p(Z|D_e)$ that integrates data
error and the randomness of the spatial variation. The BME
PDF is a non-Gaussian distribution. The red dashed curve is
the comprehensive uncertainty $p(Z|D_e, K)$ that integrates the
data error, the randomness of spatial variation, and the cognitive
uncertainty. From a variance perspective, the conditional vari-
ances $Var(Z|D) = \sigma_k^2 = 210.3$, $Var(Z|D_e) = 242.5$ and $Var(Z|D_e, K) = 446.8$ reflect that the uncertainty of geological inter-
face increases with considering more uncertainty factors.

Likewise, the probability distributions of the interface be-
tween the strata $T_2b^2$ and $T_2b^1$ at all of the locations were
calculated and shown in Fig. 10. The probability density at
each location means the occurrence probability that the inter-
face extends through this grid cell. It is observed that, in the
$p(Z|D)$ stage, a high probability appears around the observa-
tions. The occurrence probability of the interface gathers
around the prediction of the kriging estimation. In the $p(Z|D_e)$ stage, considering the impact of the data error, the occur-
rence probability in the observation area decreases and the
confidence interval width of interface increases. In addition,
the occurrence probability in the area far away from observa-
tions increases slightly. In the process of integration of data
error and spatial variation ($p(Z|D) \rightarrow p(Z|D_e)$), the changes
(increase or decrease) of the probability distribution of inter-
face are shown in Fig. 11a. In the $p(Z|D_e, K)$ stage, the prob-
ability density is condensed around the interface model. In
contrast, the occurrence probability decreases rapidly outside
of the model coverage. In the process of uncertainty update
considering cognition ($p(Z|D_e) \rightarrow p(Z|D_e, K)$), the changes
(increase or decrease) of the probability distribution of inter-
face are shown in Fig. 11b.

### Probability field of geological models

After the uncertainties of each interface have been integrated
separately, the probability of each stratigraphic type was cal-
culated with the Eq. (A1). The contact relationships are

depositional or erosional between adjacent strata in this study
area. According to the contact relationships, a conditional
probability field of all the stratigraphic types was obtained
by the iterative update Eqs. (A2) and (A3). Each location in
this conditional probability field stores a vector of 22 condi-
tional probabilities of all of the stratigraphic types. This multi-
stratigraphic type probability field represents the comprehen-
sive uncertainty of every stratum.

Similar to what is shown in Fig. 12, for each stratigraphic
type, high probabilities mainly occur in the outcrop and in the
interior of the stratum. In addition, the probability gradually
decreases from the interior to the boundary between the pres-
ent stratum and its neighbor. At the observation location, con-
sidering the measurement error, the probability of each strat-
igraphic type in the sample is close to, but not exactly, 0 or 1.
Around the section area, the interpretation error of the section
is much higher than the boreholes' error; therefore, the reli-
ability at this location is lower than that at the drilling area, and
the probability is slightly farther from 0 or 1.

### Comprehensive uncertainty field

We calculated the information entropy based on the condition-
al probability of each stratigraphic type to quantify the struc-
tural uncertainty of the entire model. The entropy value rep-
resents the quantity of the comprehensive uncertainty of the
model at any given location. In order to demonstrate the in-
fluence of data error on the uncertainty assessment, we also
calculated the model uncertainty only considering spatial var-
iation and cognitive bias (without data error). The 3D struc-
tural model and its uncertainty fields are shown in Fig. 13.

In the comprehensive uncertainty field (Fig. 13c), the high
entropy mainly appears in the area where more than one stra-
tum may exist. The low entropy mostly occurs in the area near
the observations or where there is only one single stratum.
Sample information from the observation reduces the uncer-
tainty of the spatial variation. It is easy to understand why low
entropy usually occurs at a location that only has one single

**Fig. 10** Probability density distribution of the interface between strata $T_2b^2$ and $T_2b^1$ in different stages of integration: (**a**) the interface model; (**b**) perspective view of probability density distribution; (**c**) side view of probability density distribution



perspective view    side view

(a)

$p(Z|D)$    $p(Z|D)$

Probability density
>0.1
0.075
0.05
0.025
0

$p(Z|D_e)$    $p(Z|D_e)$
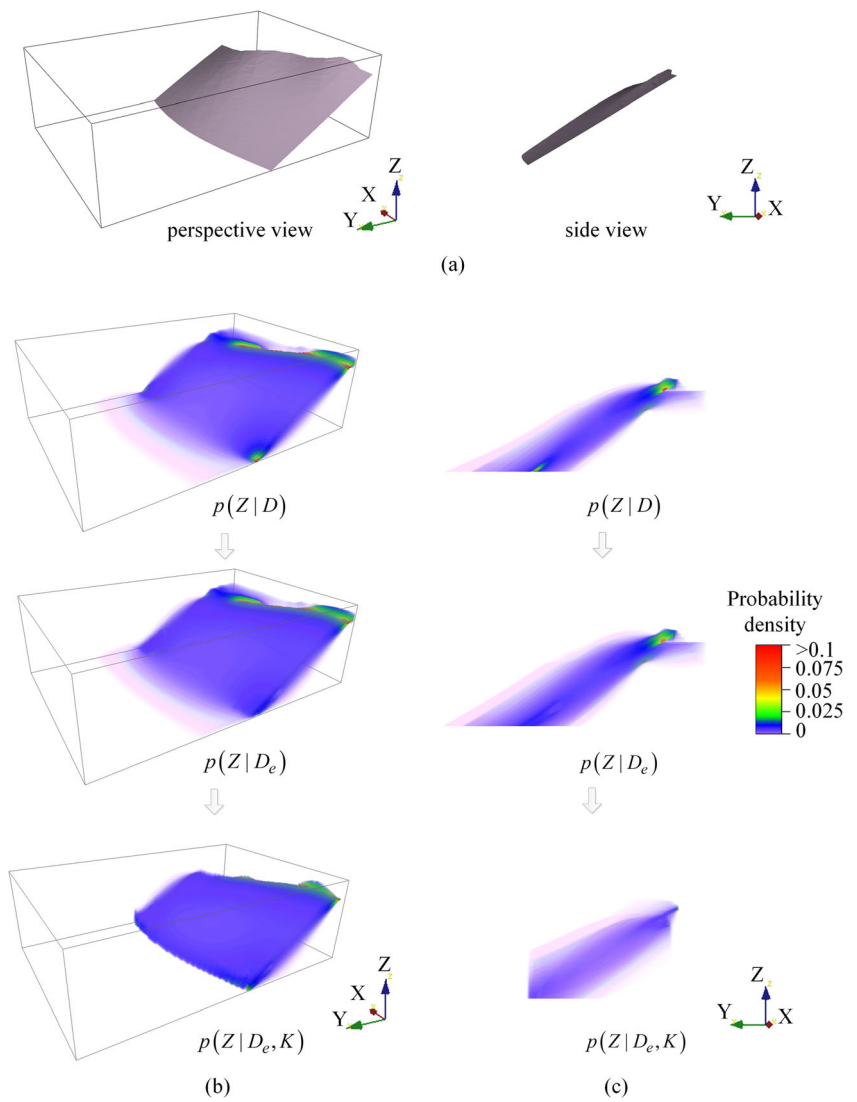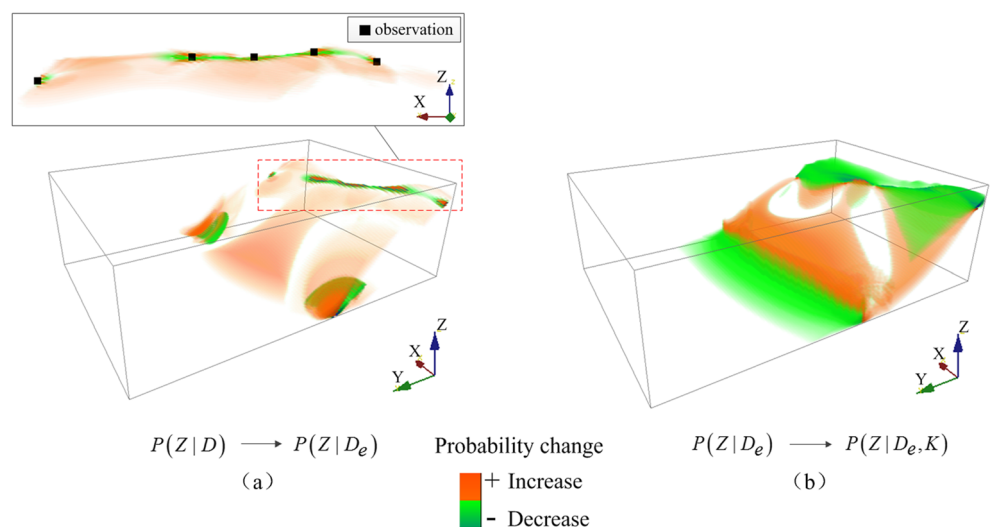
$p(Z|D_e,K)$    $p(Z|D_e,K)$

(b)    (c)

**Fig. 11** Probability changes in the uncertainty integration: (**a**) Probability changes after integrating data errors. The area near observations is enlarged and displayed. (**b**) Probability changes after considering the cognitive uncertainty



observation

$P(Z|D) \longrightarrow P(Z|D_e)$    Probability change    $P(Z|D_e) \longrightarrow P(Z|D_e,K)$
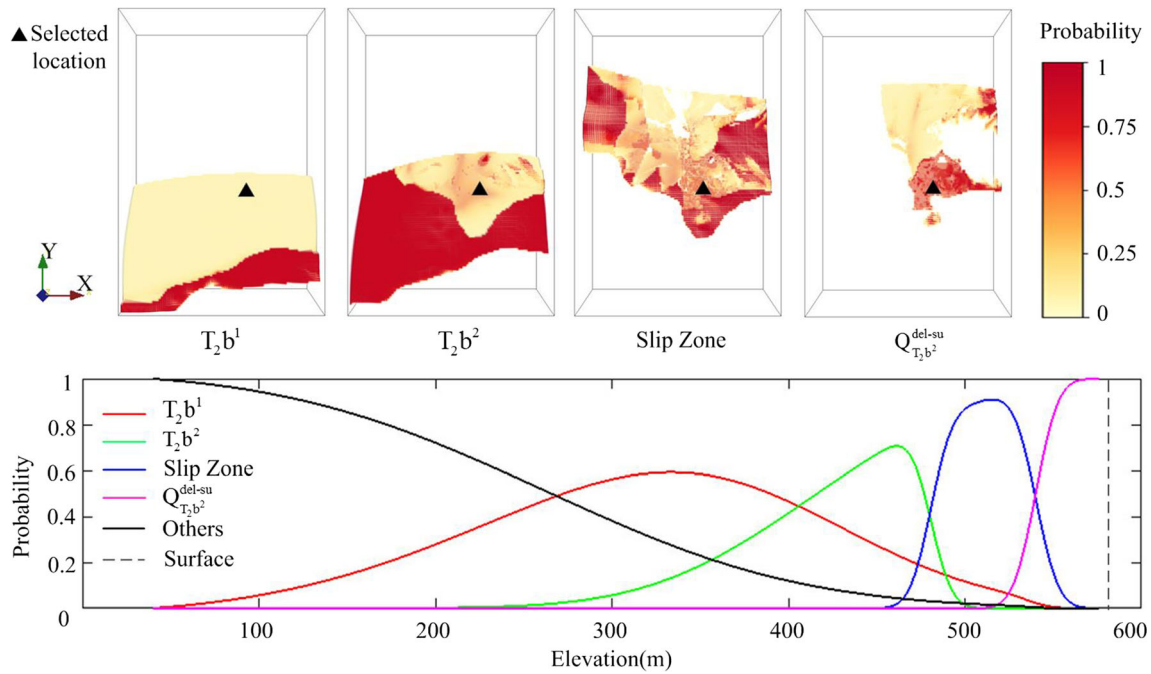
（a）    + Increase    （b）

− Decrease

Fig. 12 Probability distribution of the four stratigraphic types ($T_2b^1$, $T_2b^2$, Slip Zone and $Q_{T_2b^2}^{del-su}$). In the overlay area of the strata, a location is selected to show the profile of the probability from the base to top of each stratum

stratum because there are few other possibilities. The maximum entropy in the study area is approximately 2.359 bit, which occurs in the aggregation area of the six different strata. In the interlayer of the adjacent interfaces, if the thickness of
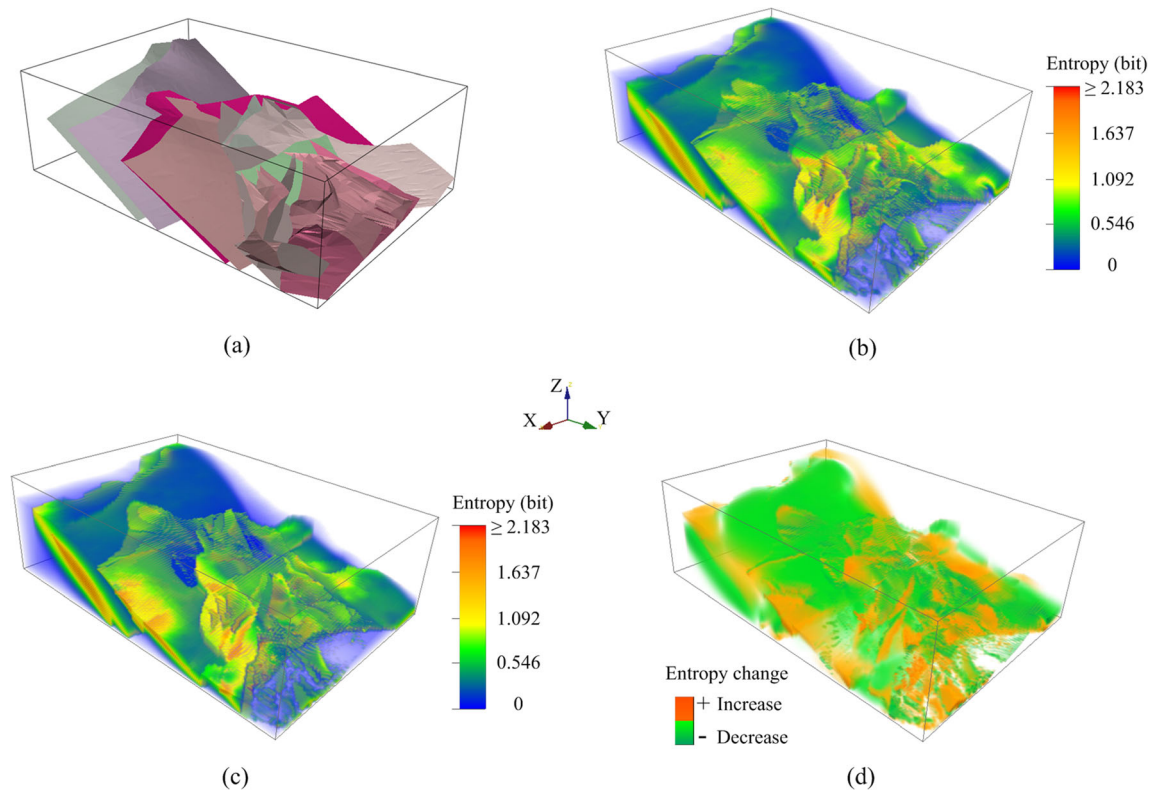


Fig. 13 3D structural model of Huangtupo slope and its uncertainty field (unit of entropy: bit): (a) subsurface model of Huangtupo slope; (b) uncertainty field considering spatial variation, and cognitive bias; (c) comprehensive uncertainty field considering data error, spatial variation, and cognitive bias; (d) influence of data error on model uncertainty

the stratum is small, then more types of possible stratigraphic types appear in this complex area. The entropy range of $1 \sim 1.58$ bit is common in the interlayer area. The low entropy usually appears on the surface because the field data are much easier to collect on the ground. Most of the high entropy regions near the ground congregate around the stratigraphic boundary. Some other high entropy areas are caused by a lack of observations and a reliance on the extrapolations from the modeler's experience in the modeling software. As a contrast, Fig. 13b shows the uncertainty field without considering data error. In Fig. 13d, we can see the influence of data error on model uncertainty clearly: entropy increases at the locations of geological interface in the model and decreases at the areas that far away (in height) from geological interface. The complex errors from diverse data affect the distribution of uncertainty of subsurface structure. With the consideration of data error, the uncertainty of geological interface increased.

## Discussion

In the prediction of a geological variable, the PDF calculated by BME is more credible than that calculated by kriging (Christakos and Li 1998). However, to estimate the likelihood function $L(Z|K)$, we still need the help of the kriging variance. In the result of $p(Z|D)$ (as shown in Fig. 10), it is observed that the local maximum probabilities from the kriging prediction vary around the observations, and the model is deterministic in the observation positions. As the distance departs from the observation values, the variance of $Z(\boldsymbol{u})$ increases, which indicates that the randomness of the spatial variation is increasing and the occurrence probability of the interface is decreasing. After the data error is integrated by BME, the uncertainties in the observations increase. The observed deterministic value becomes randomly distributed with the observation error. As the uncertainty is propagated, the occurrence probabilities of the interfaces decrease near the observations and increase far away from the observations (as shown in Fig. 11a). Considering the modeler's subjective cognition, the occurrence probabilities of the interface increase in the corresponding model region and decrease outside of this region (as shown in Fig. 11b). In the area where the model approximates the BME prediction, the probability remains approximately unchanged. These results are consistent with our expectations for the propagation of multi-source uncertainties in the modeling process. Our approach further considers the data error and the modeler's cognitive bias relative to previous work.

The uncertainty analysis is essentially a model-based spatial analysis. This spatial analysis process may introduce, propagate, or even amplify the uncertainty of the original model (Shi 2009). Usually, the modeler and the analyst of a model's uncertainty are not the same person. In the

uncertainty analysis of an established model, the original parameter settings and interactive modifications employed by the modeler are difficult to obtain and to reproduce. A discrepancy between the uncertainty assessment and the theoretical accuracy of modeling will inevitably appear when different methods and parameters are adopted in the analysis versus the modeling process. Different uncertainty analysis methods will lead to different assessment results for the same model (Bárdossy and Fodor 2004). The selection of methods depends on the subjective cognition and the empirical knowledge of the analyst. Similarly, the established model contains the modeler's cognitive bias. Therefore, the uncertainty assessment of the established model is not only affected by the data error and the randomness of spatial variation, but is also influenced by the cognitive uncertainty of both the modeler and the analyst (Caers 2011). Whether in modeling or in uncertainty analysis, the influence of a human's subjective cognition is hard to avoid (Shi 2009).

In this research, modeler's empirical knowledge is assumed to be conditional independent with sample data (no matter whether considering data error or not). In fact, the Bayesian inference method in the Section 2.3.2 is Naive Bayes model. In practical application, it does not always confirm to the assumption on conditional independence. This is a problem of estimation in condition of incomplete information, because we have no further information about the correlation of the variables. Some studies indicate that (Rish 2001; Zhang 2004; Kupervasser 2014), for the case that the correlation is unknown, the Naive Bayes model maybe not exact, but optimal solution. In our case, the Naive Bayes model is applicable. In addition, we assume that our object of analysis is the best guess model using mathematical or empirical criteria. If misconceptions or inaccurate knowledge exists in the modeler's cognition, the "best guess" assumption is perhaps no longer valid. In this situation, some other methods should be considered to quantify cognitive uncertainties.

## Conclusions

We developed a new method for performing uncertainty assessments of geological models. This method applies to the comprehensive evaluation of spatial uncertainty in structural models. The multi-source uncertainties derived from data errors, spatial variations, and cognitive bias are considered. Under the constraints of observations and geological rules, the various uncertainties of a geological model are quantified as probabilities. Using this method, multi-source uncertainties are integrated into the posterior probability of each interface and assembled according to the contact relationships. To determine the uncertainty of the integral structure, a comprehensive uncertainty field is calculated based on the posterior probability of the interfaces and their contact relationships. The

assessment result can be treated as a representation of the modeler's confidence in the model after integrating all available information.

In this paper, we employ BME to integrate multi-source uncertainties. The BME method can integrate information that contains a variety of error distributions and prior knowledge, however the information sources for BME need to satisfy a conditional independence assumption (Allard et al. 2012). The information integration methods suitable for more general condition will be considered in future studies. Expert knowledge is useful for evaluating the possibility of geological phenomena (Guillen et al. 2008; Howard et al. 2009; Wellmann et al. 2014). More examples of expert knowledge, such as the stratum thickness and conceptual graphs of the geological structure, will be considered in the next stage.

## Appendix: Combining uncertainties according to stratigraphic relations

The contact relationships of strata are divided into two types: depositional and erosional. By assembling the deposition and erosion interfaces on the basis of stratigraphic sequences, the stratigraphic type probability fields can be consistent with the actual geological structure.

First, we define the occurrence probability of each stratigraphic type. Since the $(U, V, Z)$ coordinate system of different interfaces are not the same, to consider multiple interfaces, the calculation of probability field is in $(X, Y, Z)$ coordinate system. Assuming that at location $(x, y, z)$ in the study area, there are n strata sequentially numbered from bottom to top, the occurrence probability of the $i$th stratum, $L_i$ $(i = 1, \ldots n)$, is $P(L_i)$. $p_i(z)$ is the PDF of the interface of stratum $L_i$ and stratum $L_{i+1}$. $P(L_i)$ can be calculated using the integral of $p_i(z)$:

$$P(L_i) = F(x, y, z) = \int_{L_i} p_i(z)dz. \tag{A1}$$

In the case of a depositional contact, the stratigraphic type probabilities are calculated from bottom to top, and the probability fields are updated iteratively on the basis of stratigraphic sequence. The iterations depend on the total number of strata. The iterative process in the case of a depositional contact is:

$$\begin{cases} P_N(L_i) = P_{N-1}(L_i), i = 1, 2, \ldots, N-2 \\ P_N(L_{N-1}) = P_{N-1}(L_{N-1})[1-P(L_N)] \\ P_N(L_N) = P(L_N) - \sum_{j=1}^{N-2} P(L_N) \cdot P_{N-1}(L_j) \end{cases} \tag{A2}$$

In the formula, $L_i$ is the $i$th stratum from the bottom up. $P_N(L_i)$ is the conditional probability of the stratigraphic type, $L_i$, at the $N$th iteration; $P(L_N)$ is the probability of stratigraphic type, $L_N$, without considering the impact of other strata. In the $N$th iteration, the probabilities of all the strata under $L_N$ in the stratigraphic sequence are considered and updated. The iterative update ends after the iteration $N = n$.

Faults and discordance interfaces appear as discontinuities in the stratigraphic sequence and can be treated as erosion interfaces. Because erosion does not belong to the depositional process, it needs to be treated separately. When the interface of $L_i$ and $L_{i+1}$ is an erosional interface, the probability field is updated as follows:

$$\begin{cases} P_N(L_i) = P_{N-1}(L_i)[1-P(L_N)], i = 1, 2, \ldots, N-1 \\ P_N(L_N) = P(L_N) \end{cases} \tag{A3}$$

Based on the principle of mutual exclusion of stratigraphic type and the full probability formula, the cumulative probability of all possible strata at each location is equal to 1.

## References

Abrahamsen P (1993) Bayesian kriging for seismic depth conversion of a multi-layer reservoir. In: Soares A (ed) Geostatistics Tróia '92. Quantitative geology and geostatistics, vol 5. Springer, Dordrecht, pp 3859–398. https://doi.org/10.1007/978-94-011-1739-5_31

Abrahamsen P, Omre H (1994) Random functions and geological subsurfaces. Ecmor IV-European Conference on the Mathematics of Oil Recovery, Røros, p 21

Abrahamsen P, Omre H, Lia O (1991) Stochastic models for seismic depth conversion of geological horizons. Soc Petrol Eng. https://doi.org/10.2118/23138-MS

Allard D, Comunian A, Renard P (2012) Probability aggregation methods in geoscience. Math Geosci 44(5):545–581

Ariely D (2008) Predictably irrational: the hidden forces that shape our decisions. HarperCollins, New York

Arnold D, Demyanov V, Rojas T, Christie M (2019) Uncertainty quantification in reservoir prediction: part 1—model realism in history matching using geological prior definitions. Math Geosci 51(2):209–240

Bárdossy G, Fodor J (2001) Traditional and new ways to handle uncertainty in geology. Nat Resour Res 10(3):179–187

Bárdossy G, Fodor J (2004) Evaluation of uncertainties and risks in geology-new mathematical approaches for their handling. Springer Press, New York, pp 3–11

Bistacchi A, Massironi M, Piaz GVD, Piaz GD, Monopoli B, Schiavo A et al (2008) 3d fold and fault reconstruction with an uncertainty model: an example from an alpine tunnel case study. Comput Geosci 34(4):351–372

Bond CE (2015) Uncertainty in structural interpretation: lessons to be learnt. J Struct Geol 74:185–200

Bond CE, Philo C, Shipton ZK (2011) When there isn't a right answer: interpretation and reasoning, key skills for twenty-first century geoscience. Int J Sci Educ 33(5):629–652

Bordwell P (2002) Spatial prediction of categorical variables: the Bayesian maximum entropy approach. Stoch environ res risk assess. Stoch Env Res Risk A 16(6):425–448

Caers J (2011) Modeling uncertainty: concepts and philosophies. In modeling uncertainty in the earth sciences, J. Caers (Ed.). doi: https://doi.org/10.1002/9781119995920.ch3

Caumon G, Journel AG (2005) Early uncertainty assessment: application to a hydrocarbon reservoir appraisal. In: Geostatistics Banff 2004. Springer, Dordrecht, pp 551–557

Chilès JP, Aug C, Guillen A, Lees T (2004) Modelling the geometry of geological units and its uncertainty in 3D from structural data: the potential-field mdethod, vol 2004. Orebody Modelling and Strategic Mine Planning, Perth, WA, pp 313–320

Christakos G (1990) A Bayesian maximum-entropy view to the spatial estimation problem. Math Geol 22(7):763–777

Christakos G (2017) Uncertainty, modeling with spatial and temporal. In: Shekhar S, Xiong H, Zhou X (eds) Encyclopedia of GIS. Springer, Cham, pp 1189–1193

Christakos G, Li X (1998) Bayesian maximum entropy analysis and mapping: a farewell to Kriging estimators? Math Geol 30(4):435–462

Christakos G, Bogaert P, Serre M (2002) Temporal GIS. In: advanced functions for field-based applications. Springer-Verlag Berlin Heidelberg, pp 143–188. https://doi.org/10.1007/978-3-642-56540-3

de la Varga M, Wellmann JF (2016) Structural geologic modeling as an inference problem: a Bayesian perspective. Interpretation 4(3): SM1–SM16

Demyanov V, Arnold D, Rojas T, Christie M (2019) Uncertainty quantification in reservoir prediction: part 2—handling uncertainty in the geological scenario. Math Geosci 51(2):241–264

Dowd P (2018) Quantifying the impacts of uncertainty. In: Daya SB, Cheng Q, Agterberg F (eds) Handbook of mathematical geosciences. Springer, Cham, pp 359–360

Doyen PM, Den Boer LD (1996) Bayesian sequential Gaussian simulation of lithology with non-linear data, no. 5539704. Western Atlas International, Inc., Houston. https://www.freepatentsonline.com/5539704.html

Edwards J, Lallier F, Caumon G, Carpentier C (2017) Uncertainty management in stratigraphic well correlation and stratigraphic architectures: a training-based method. Comput Geosci 111:1–17

Fazekas I, Kukush AG (2005) Kriging and measurement errors. Discuss Math Probab Stat 25(2):139–159

Guillen A, Calcagno P, Courrioux G, Joly A, Ledru P (2008) Geological modelling from field data and geological knowledge. Phys Earth Planet Inter 171(1–4):158–169

Gunning J (2000) Constraining random field models to seismic data: getting the scale and the physics right. ECMOR VII: proceedings, seventh European conference on the mathematics of oil recovery, Baveno, Italy, p M-20

Gunning J, Glinsky ME (2004) Delivery: an open-source model-based Bayesian seismic inversion program. Comput Geosci 30(6):619–636

Haselton MG, Nettle D, Andrews PW (2005) The evolution of cognitive bias. In: Buss DM (ed) The handbook of evolutionary psychology: Hoboken. NJ, US, John Wiley & Sons Inc, pp 724–746

Hou W, Yang Q, Yang L, Cui C (2017) Uncertainty analysis of geological boundaries in 3D cross-section based on Monte Carlo simulation. J Jilin Univ (Earth Sci Ed) 47(3):925–932

Hou W, Cui C, Yang L, Yang Q, Clarke K (2019) Entropy-based weighting in one-dimensional multiple errors analysis of geological contacts to model geological structure. Math Geosci 51(1):29–51

Howard AS, Hatton B, Reitsma F, Lawrie KIG (2009) Developing a geoscience knowledge framework for a national geological survey organisation. Comput Geosci 35(4):820–835

Hu X, Tang H, Li C, Sun R (2012) Stability of Huangtupo riverside slumping mass II# under water level fluctuation of three gorges reservoir. J Earth Sci 23(3):326–334

Jessell MW, Ailleres L, Kemp EAD (2010) Towards an integrated inversion of geoscientific data: what price of geology? Tectonophysics 490(3):294–306

Kang J, Jin R, Li X, Zhang Y (2017) Block kriging with measurement errors: a case study of the spatial prediction of soil moisture in the middle reaches of heihe river basin. IEEE Geosci Remote Sens Lett PP(99):1–5

Kupervasser O (2014) The mysterious optimality of naive Bayes: estimation of the probability in the system of "classifiers". Pattern Recognit Image Anal 24:1–10

Lark RM, Mathers SJ, Thorpe S, Arkley SLB, Morgan DJ, Lawrence DJD (2013) A statistical assessment of the uncertainty in a 3-d geological framework model. Proc Geol Assoc 124(6):946–958

Lelliott MR, Cave MR, Wealthall GP (2009) A structured approach to the measurement of uncertainty in 3D geological models. Q J Eng Geol Hydrogeol 42(1):95–105

Li X, Li P, Zhu H (2013) Coal seam surface modeling and updating with multi-source data integration using Bayesian Geostatistics. Eng Geol 164:208–221

Lindsay MD, Aillères L, Jessell MW, Kemp EA, Betts PG (2012) Locating and quantifying geological uncertainty in three-dimensional models: analysis of the Gippsland Basin, southeastern Australia. Tectonophysics 546-547(2012):10–27

Mann CJ (1993) Uncertainty in geology. In: Davis JC, Herzfeld UC (eds) . Computers in geology – 25 years of progress: Oxford Univ. Press, Oxford, pp 241–254

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):385–402

Pakyuz-Charrier E, Jessell M, Giraud J, Lindsay M, Ogarko V (2019) Topological analysis in Monte Carlo simulation for uncertainty propagation. Solid Earth 10(5):1663–1684

Pomian-Srzednicki I (2001) Calculation of geological uncertainties associated with 3-D geological models. Dissertation, Ecole Polytechnique Fédérale de Lausanne, Lausanne, pp 15–17 59pp

Riddick A, Laxton J, Cave M, Wood B, Duffy T, Bell P et al (2005) Digital geoscience spatial model project final report. BGS Occasional Publication No. 9. British Geological Survey, Keyworth, p 56

Rish I (2001) An empirical study of the naive bayes classifier. J Univ Comput Sci 1(2):127

Røe P, Georgsen F, Abrahamsen P (2014) An uncertainty model for fault shape and location. Math Geosci 46(8):957–969

Schneeberger R, Varga MDL, Egli D, Berger A, Kober F, Wellmann JF, Herwegh M (2017) Methods and uncertainty estimations of 3-D structural modelling in crystalline rocks: a case study. Solid Earth 8(5):987–1002

Schweizer D, Blum P, Butscher C (2017) Uncertainty assessment in 3-D geological models of increasing complexity. Solid Earth 8(2):515–530

Shi W (2009) Principles of modeling uncertainties in spatial data and spatial analyses (1st ed). CRC Press, Boca Raton, pp 15–26

Soares A, Nunes R, Azevedo L (2017) Integration of uncertain data in geostatistical modelling. Math Geosci 49(2):253–273

Suzuki S, Caumon G, Caers J (2008) Dynamic data integration for structural modeling: model screening approach using a distance-based model parameterization. Comput Geosci 12(1):105–119

Tacher L, Pomian-Srzednicki I, Parriaux A (2006) Geological uncertainties associated with 3-D subsurface models. Comput Geosci 32(2):212–221

Thiele ST, Jessell MW, Lindsay M, Wellmann JF, Pakyuz-Charrier E (2016) The topology of geology 2: topological uncertainty. J Struct Geol 91:74–87

Thore P, Shtuka A, Lecour M, Ait-Ettajer T, Cognot R (2002) Structural uncertainties: determination, management, and applications. Geophysics 67(3):840–852

Wellmann JF, Regenauer-Lieb K (2012) Uncertainties have a meaning: information entropy as a quality measure for 3-D geological models. Tectonophysics 526-529(2):207–216

Wellmann JF, Horowitz FG, Schill E, Regenauer-Lieb K (2010) Towards incorporating uncertainty of structural data in 3D geological inversion. Tectonophysics 490(3–4):141–151

Wellmann JF, Lindsay M, Poh J, Jessell MW (2014) Validating 3-D structural models with geological knowledge for improved uncertainty evaluations. Energy Procedia 59:374–381

Wellmann JF, de la Varga M, Murdie RE, Gessner K, Jessell M (2017) Uncertainty estimation for a geological model of the Sandstone greenstone belt, Western Australia – insights from integrated geological and geophysical inversion in a Bayesian inference framework. Geological Society. Special Publications, London, p SP453.12

Wu Q, Xu H, Zou X (2005) An effective method for 3D geological modeling with multi-source data integration. Comput Geosci 31(1):35–43

Yang SS, Chen HY (2001) Investigation report of Huangtupo landslide area in Badong County. Hubei Survey and Design Institute for Geohazard Engineering, Hubei Province

Zhang H (2004) The Optimality of Naive Bayes, In FLAIRS Conference. http://www.cs.unb.ca/profs/hzhang/

Zhu L, Zhuang Z (2010) Framework system and research flow of uncertainty in 3D geological structure models. Min Sci Technol (China) 20(2):306–311