Geosci. Model Dev., 15, 4689–4708, 2022
https://doi.org/10.5194/gmd-15-4689-2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Geoscientific Model Development
Open Access
EGU

loopUI-0.1: indicators to support needs and practices in 3D geological modelling uncertainty quantification

Guillaume Pirot[1,2], **Ranee Joshi**[1,2], **Jérémie Giraud**[1,2,3], **Mark Douglas Lindsay**[1,2,4,5], **and Mark Walter Jessell**[1,2,4]

[1]The Centre for Exploration Targeting, School of Earth Sciences, The University of Western Australia, Perth, Australia
[2]Mineral Exploration Cooperative Research Centre (MinEx CRC), School of Earth Sciences, University of Western Australia, Perth, Australia
[3]GeoRessources Lab, University of Lorraine, Nancy, France
[4]ARC Industrial Transformation and Training Centre in Data Analytics for Resources and the Environment (DARE), Sydney, Australia
[5]CSIRO Mineral Resources, Perth, Australia

Correspondence: Guillaume Pirot (guillaume.pirot@uwa.edu.au)

Received: 10 November 2021 – Discussion started: 7 January 2022
Revised: 16 May 2022 – Accepted: 18 May 2022 – Published: 20 June 2022

Abstract. To support the needs of practitioners regarding 3D geological modelling and uncertainty quantification in the field, in particular from the mining industry, we propose a *Python* package called `loopUI-0.1` that provides a set of local and global indicators to measure uncertainty and features dissimilarities among an ensemble of voxet models. Results are presented of a survey launched among practitioners in the mineral industry, enquiring about their modelling and uncertainty quantification practice and needs. It reveals that practitioners acknowledge the importance of uncertainty quantification even if they do not perform it. A total of four main factors preventing practitioners performing uncertainty quantification were identified: a lack of data uncertainty quantification, (computing) time requirement to generate one model, poor tracking of assumptions and interpretations and relative complexity of uncertainty quantification. The paper reviews and proposes solutions to alleviate these issues. Elements of an answer to these problems are already provided in the special issue hosting this paper and more are expected to come.

1 Introduction

An objective of researchers who develop open-source 3D geological modelling algorithms (Loop, 2019; de la Varga et al., 2019) is to make them findable, accessible, interoperable and reusable (FAIR) for practitioners. As for any software, these algorithms should satisfy the needs and expectations of users (Franke and Von Hippel, 2003; Kujala, 2008). Thus, new developments should rely on a good understanding of modelling purposes, processes and limitations, following a philosophy of continuous improvement. In a general context, this becomes even more important given the increasing number of open-source algorithms in the fields of earth and planetary sciences (see Fig. 1). However, to the best of our knowledge, the needs and uses of 3D geological modelling practitioners with respect to uncertainty quantification are only partially described in the literature, as it constitutes an emerging field that only recently gained traction in both academia and industry.

An essential purpose of modelling is to support decision makers by offering a simplified representation of nature that also provides a corresponding quantitative assessment of uncertainty, communicating what we know, what remains unknown and what is ambiguous (Ferré, 2017). Uncertainty quantification is essential, because it allows mitigation of predictive uncertainty (Jessell et al., 2018) by expanding our knowledge, rejecting hypotheses (Wilcox, 2011) or falsifying scenarios (Raftery, 1993). Questions related to 3D geological modelling and uncertainty quantification are not just limited to the minerals industry, but also concern the fields of CO_2 sequestration (Mo et al., 2019), petroleum (Scheidt and Caers, 2009) and geothermal (Witter et al., 2019) en-

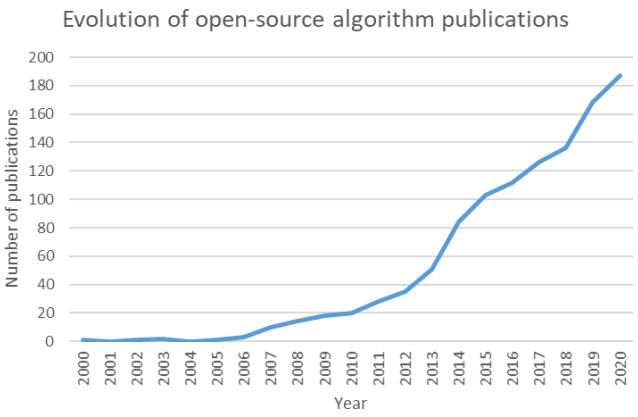


Figure 1. Evolution of yearly open-source algorithm publications between 2000 and 2020; data from Web Of Knowledge (last access: 1 September 2021).

ergy resources as well as hydrogeology (Linde et al., 2017) or civil engineering in urban environments (Osenbrück et al., 2007; Tubau et al., 2017). Here, we are interested in the uses and practices of the minerals industry, that is dealing with both sedimentary basin and hard rock and/or cratonic settings across regional to mine scales.

The three main pillars of uncertainty quantification are the characterisation of uncertainty sources, their propagation and mitigation throughout the modelling workflow (see Fig. 2). The different sources of uncertainty, often overlooked, are related to measurement errors, interpretations, assumptions, modelling approximations and limited knowledge (sample size or unknown process). Measurement or data errors can be estimated by repetitive independent sampling or from instrument characteristics; they can be propagated through the modelling workflow by Monte Carlo data perturbation (Wellmann and Regenauer-Lieb, 2012; Lindsay et al., 2012; Pakyuz-Charrier et al., 2018). Combined with expert knowledge, initial datasets can be used to shape some assumptions and define plausible conceptual models but despite the importance of conceptual uncertainty on predictions (Pirot et al., 2015), it is too often limited to the definition of a unique scenario (Ferré, 2017). From a perspective on algorithms, some assumptions such as how to set parameter ranges are needed and this can greatly impact the definition of geological parameters (Lajaunie et al., 1997) prior to running predictive numerical simulations. Another aspect that is not always considered is the uncertainty related to a spatially limited sampling. Unsampled locations suggest a high uncertainty about the spatial distribution of the model parameters (or values of the property field); this is why it is preferable to resort to spatial stochastic simulations (e.g. sequential Gaussian simulations in a multi-Gaussian world) rather than interpolations (e.g. kriging) to generate models that are parameter fields (Journel and Huijbregts, 1978).

While the main objective of uncertainty quantification and data integration might be to improve the confidence level of predictions for decision making, it usually involves the generation of model ensembles via a Monte Carlo algorithm and it is rarely a straightforward step. Indeed, because of the high dimensionality and non-linearity of earth processes or the lack of data, history matching might prove difficult to achieve and predicted outcomes might present multiple modes (e.g. Suzuki and Caers, 2008; Sambridge, 2014; Pirot et al., 2017). In such cases, geological uncertainty analysis allows to improve our understanding of geological model (dis)similarities and how specific or shared features can be related to upstream parameters and downstream predictions. Model comparison requires a shared discrete support; one possibility is to project each geological model onto a discrete mesh such as a regular grid voxet composed of volumetric cells, also called voxels. Local uncertainty indicators such as voxel-wise entropy or cardinality (Lindsay et al., 2012; Schweizer et al., 2017), computed over an ensemble of geological voxets will inform about property field variability at specific locations (voxels) of the model mesh. Global indicators or summary metrics might be useful to identify how the statistics of specific patterns (e.g. fault or fracture network density, anisotropy, connectivity, etc.) evolve between different models and might also be a way to perform model or scenario selection (e.g. Pirot et al., 2019) or to reduce the dimensionality of the sampling space, from a high dimensional geological space to a low dimensional latent space (Lochbühler et al., 2013). Although some indicators have been used or developed for specific studies or softwares (Li et al., 2014), to the best of our knowledge, no independent uncertainty analysis tool applicable to both discrete and continuous property fields, combining local and global indicators is available to practitioners.

To investigate the uses and practices of the minerals industry regarding modelling and uncertainty quantification, we recently conducted a survey among numerical modelling practitioners from industry, government and academia in the sector of exploration and production of economic minerals. In this paper, first we present the main results and interpretations from this survey, the questions of which are listed in Appendix A. Second, to answer some of these needs, we propose a set of indicators to quantify geological uncertainty over an ensemble of geological models, represented as regular grid voxets and characterised by lithological units in the discrete domain and their underlying scalar field derived from implicit modelling in the continuous domain. The various indicators are illustrated with a synthetic case derived from a simplified Precambrian dataset of the Hamersley, Western Australia. The *Python* code called `loopUI-0.1` (Pirot, 2021a) and the notebooks used to compute and illustrate these indicators are available at https://doi.org/10.5281/zenodo.5656151.

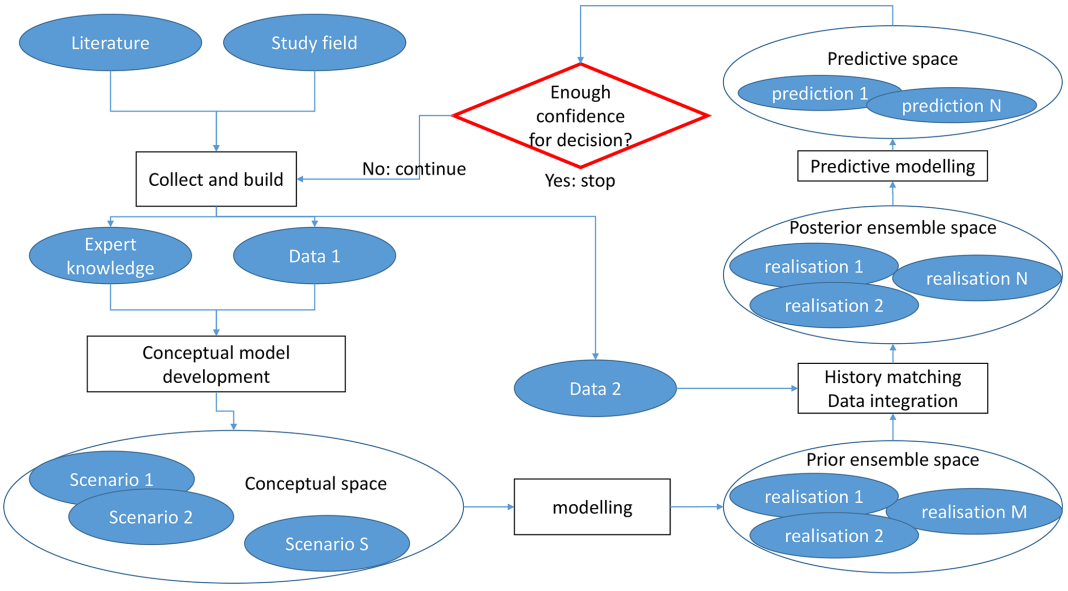


Figure 2. Schematic representation of a geomodelling workflow; each ellipse is associated with some uncertainty; Data 1 (e.g. geological observations) is used only to build a prior ensemble of models; Data 2 (e.g. geophysical measurements) is used to reduce the number of models that can be considered. It is achieved by exploring the ensemble of prior models and selecting those that are more likely to reproduce Data 2 within a prescribed level of error.

2 Survey

2.1 Material and method

The survey was designed to be concise to encourage participation but with several open-ended questions to maximise our chance to learn about different uses and practices, as well as to minimise induced bias whenever possible. The survey was in two parts. The first part was general and enquired about the scales and dimensions (questions 1 and 2) of geological models. The second part of the survey was more specific at a fixed modelling scale. It enquired about outputs or objectives (questions 3 and 4), about data input (question 5), current workflows (questions 6–9) and limitations and expected improvements (question 10).

It was distributed between October 2019 and January 2020 among the 3D Interest Group (3DIG, last access: 8 June 2022; https://www.linkedin.com/groups/6804787/members/, last access: 8 June 2022), Centre for Exploration Targeting (CET (last access: 8 June 2022) members (https://www.cet.edu.au/personnel/, last access: 8 June 2022; https://www.cet.edu.au/members/), Loop, last access: 8 June 2022; researchers (https://loop3d.github.io/loopers.html, last access: 8 June 2022) and related networks. About 150 persons were given the opportunity to participate. The solicited participants were essentially based in Australia but from a number of nationalities, with interests in geological modelling for mining applications. They were either from the industry or academia, from junior to senior profiles. Respondents had the opportunity to complete the survey on paper, in a text file or online.

A total of 35 responses were collected and anonymised. Of the responses seven concerned models at different scales, and when the second part of the survey was not clearly duplicated for each scale, the answers were considered with caution for each modelling scale. Although 35 responses is a relatively small number disallowing computation of robust statistics, statistical analysis was not the intended outcome. We endeavoured to ascertain opinions from practitioners in a specialised and important domain in the geosciences. The value is in the opinions we gathered which enhanced the survey results, aided interpretation and indicated which practices are commonly employed in geomodelling.

2.2 Results

This section summarises the answers provided by the survey respondents. Table 1 provides the general statistics of answers to the survey. The high answer rate (mostly over 80 %) indicates that questions are meaningful for the respondents and indicates that the reader can have confidence in the presented results. The most answered question was Q1, about the modelling scale. The least answered question was Q9, about data integration and upscaling . The second least answered question was Q8, about the modelling workflow used. All other questions had an answer rate above 80 % on a global average. Note that the global number of survey answers was smaller than the sum of each survey answer

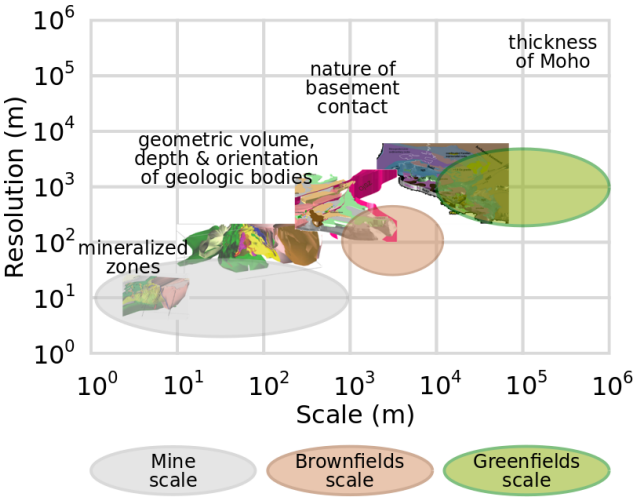


Figure 3. Modelling scales and resolutions in the mineral industry with examples of scale-specific scientific enquiry; model illustrations of the Vihanti-Pyhäsalmi Area, Finland, adapted from Laine et al. (2015).

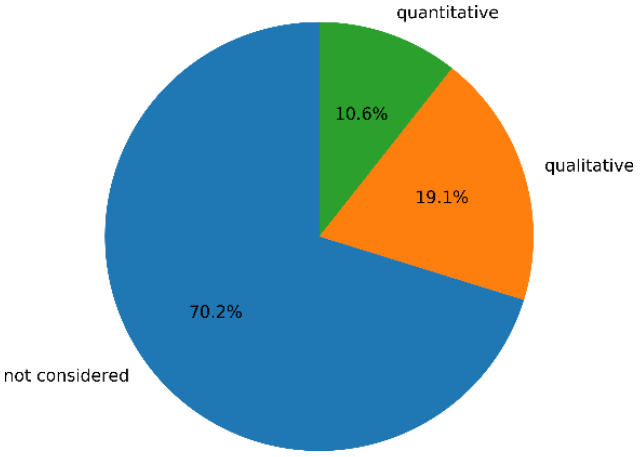


Figure 4. Consideration of input data uncertainty in modelling.

grouped by scale, as some survey answers were returned only once for multiple scales.

Due to existing overlap of collected answers on different questions, the results presented hereafter summarise and group the collected answers by theme (output or objectives, input data, current modelling workflow and limitations), as outlined in the survey (see Appendix). Each theme is treated by modelling scale (see Fig. 3) when it involves different answers.

Q3–4. The main modelling objectives depend on the scale of investigation. At the largest scales, investigation and modelling are particularly useful to assist exploration and prospectivity mapping. At the Greenfields or regional scale ($>$ 10 km dimension, $\sim$ 1 km resolution) , the main objective is to obtain a contextual and conceptual understanding of the regional geology, in particular regarding stratigraphy, topology and geochronology. At the Brownfields scale (1–10 km dimension, $\sim$ 100 m resolution), modelling aims to estimate deep structures beyond the depths reached by drilling, such as the depth of a particular interface or the delineation of potential mineral system components (e.g. fluid alteration pathways and ore deposition environments). At the Mine scale ($<$ 1 km dimension, $\sim$ 10 m resolution), the objectives are to assist with near mine exploration, resource estimation, ore body localisation, drill targeting, operation scheduling and efficient mining, by characterising local structures and geometries of stratigraphy and mineralisation. In addition to fulfilling these various objectives, 3D models are useful to improve hydrogeological characterisation, to identify critical zones where more knowledge has to be gained, to allow for comparison between datasets and test internal consistency, scenarios and last but not least, for visualisation and communication.

Q5. Data types used as input for geoscientific modelling are similar and complementary across scales. For instance, geological mapping, geophysical data (gravity, magnetics, electromagnetics, seismic), and geochemistry are useful at all scales even though the related measurements might inform about a regional trend. Drill hole-derived data are usually much more abundant at the Mine scale and can be useful to understand the regional context if analysed with this alternative use in mind. For example, drill logs that only record the presence or absence of mineralisation are not useful for the regional context; however, those that record rock properties, lithology and structure can be extrapolated to larger scales and integrated in regional models (e.g. Lindsay et al., 2020).

Q6. Regarding current modelling practices, although the survey respondents acknowledge the importance of uncertainty quantification and propagation into modelling to estimate the confidence around predictions, only about 10 % perform quantitative uncertainty characterisation and propagation from input data; about 20 % do some qualitative characterisation and a vast majority of 70 % recognise that it is ignored (see Fig. 4).

Q7–9. Modelling steps, assumptions or geological interpretations are not recorded in one third of the cases. In most cases, assumptions and elements of the modelling process are described in metadata or in separate reports. In very few cases, specific procedures and tools are available to keep track of these. The respondents use a variety of software, platforms or programming tools to produce 3D geological models and further data integration. The lack of a coherent workflow introduces the potential for uncertainty, error and loss of precision when forced to translate data formats across multiple software platforms, in addition to the time lost that could otherwise be dedicated to solving the problem under question or exploring alternative hypotheses.

Q10. Overall, the main limitations involved in 3D geological modelling are uncertainty underestimation due to

Table 1. Global number and rate of answers per question and detailed by modelling scale.

	Total	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
All scales	Answer rate	100 %	97 %	94 %	83 %	91 %	89 %	86 %	77 %	60 %	83 %
Nb. answers	35	35	34	33	29	32	31	30	27	21	29
Mine scale	Answer rate	100 %	93 %	100 %	71 %	93 %	86 %	86 %	64 %	43 %	79 %
Nb. answers	14	14	13	14	10	13	12	12	9	6	11
Brownfields scale	Answer rate	100 %	94 %	100 %	82 %	94 %	94 %	94 %	82 %	76 %	82 %
Nb. answers	17	17	16	17	14	16	16	16	14	13	14
Greenfields scale	Answer rate	100 %	100 %	88 %	75 %	81 %	81 %	88 %	81 %	63 %	81 %
Nb. answers	16	16	16	14	12	13	13	14	13	10	13

strong constraining and underlying assumptions (e.g. lack of consideration of alternative conceptual models, lack of uncertainty around interpreted horizons), navigation between scales, the amount of work required to process and prepare input data (including necessary artificial data), the time required to generate one model, the difficulty to integrate 2D data into a 3D framework, the difficulty to visualise and manage huge amounts of drill hole data in 3D, the advent of geological inconsistencies, workflow and model reproducibility given the same inputs, joint integration of geological and geophysical data and lack of tools to visualise uncertainty.

In what follows we address one of the needs identified in the survey: the lack of tools to quantify and visualise uncertainty from an ensemble of 3D voxets (prior or posterior ensembles). Indeed, it is of utmost importance for practitioners as it allows reinterpretation of data or scenario importance with respect to geological uncertainty or predictive uncertainty.

3 Uncertainty indicators for categorical or continuous property fields

The estimation of prediction confidence in geosciences relies heavily on numerical simulations and requires generation of an ensemble of models. As indicated by the survey results, an important aspect of uncertainty quantification is its representation. It helps identifying specific features or zones of interest. Moreover, respondents estimate that such visualisation tools are missing. The uncertainty indicators presented hereafter provide a way to identify zones of greater or smaller uncertainty as well as (dis)similarities between geomodel realisations.

Geomodels can be used to convey very different discrete or continuous properties. Discrete or categorical properties, such as lithological formations or classification codes should be compared carefully. Indeed, while it is straightforward to state if two values are identical or different, additional information is needed to rank dissimilarities. Continuous properties, such as potential fields or physical properties (e.g. porosity, conductivity) do not present this ambiguity to compare range of values. One can note that depending on how physical properties are assigned during modelling, their value spectrum might be discrete. Some modelling platforms may produce a discrete physical property value spectrum depending on how physical properties are assigned or what input constraints are enforced during model construction.

Local measures of uncertainty provide indicator voxets of the same dimensions as the voxets of a model ensemble. For a given voxel, uncertainty indicators are computed from the distributions of values taken by a given property at the corresponding voxel (same location) across the ensemble of model realisations. Such local indicators are very convenient for visualisation: by sharing the same voxet as the model realisations, it is relatively easy to spot zones of low or high uncertainty. However, to be useful it requires advanced modelling that integrates some spatial constraints and possibly computationally expensive, particularly if they are produced by inversion algorithms. Here, we propose to compute cardinality and Shannon's entropy for discrete properties (e.g. Wellmann and Regenauer-Lieb, 2012; Pakyuz-Charrier et al., 2018), and similarly, range, standard deviation and continuous entropy for continuous properties (e.g. Marsh, 2013; Pirot et al., 2017).

Global measures rely on the computation of summary statistics or on the identification of feature characteristics, independently from their locations. The dissimilarities between summary statistics or characteristics can be estimated via appropriate metrics, such as the Wasserstein distance (Vallender, 1974) or the Jensen-Shannon divergence (Dagan et al., 1997). The resulting global measures allow comparison of models with voxets or meshes of different dimensions. However, the computation of summary statistics might be more time consuming than local measures. Their main advantage is that it enables a focus on pattern similarity between models which might be particularly useful to explore the selection of alternative scenarios or prior realisations, before data integration and history matching. In what follows, we propose a series of dissimilarity measures, applicable to categorical and continuous property fields, based on 1-point statistics (histogram, Dagan et al., 1997), 2-point (geo)statistics

(semi-variogram, Matheron, 1963), multiple-point statistics (multiple-point histogram, Boisvert et al., 2010), connectivity (Renard and Allard, 2013), wavelet decomposition (e.g. Scheidt et al., 2015) and topology (e.g. Thiele et al., 2016). Some indicators are similar to exploratory data analysis techniques but here they are applied to model ensemble analysis.

To illustrate the different indicators, inspired by a Precambrian geological setting, that is a simplified dataset from the Hamersley region, Western Australia, we generated synthetic ensembles of 10 model sets for 3 different scenarios. Each model set is composed of a lithocode voxet describing the lithological units (categorical variable), and of its underlying scalar field voxet (continuous variable). The underlying scalar field is obtained by composition of the different scalar fields for each unconformable stratigraphic group. Scenario 1 considers all synthetic input data, while scenario 2 keeps only 50 % of the data within a north-south limited band and scenario 3 retains input data with a 50 % probability (see Fig. 5). For each scenario, the positions and orientations of the input data are perturbed to provide 10 stochastic realisations. Positions are perturbed with a Gaussian error of zero mean and 3 m standard deviation. Orientations are perturbed with a von Mises-Fisher error of $\kappa = 150$ (corresponding to about $\pm 5°$ of error). Each 3D model was generated using *LoopStructural* (Grose et al., 2021).

3.1 Cardinality

The cardinality of a set in mathematics, is the number of elements of the set. Here, we define the cardinality for a given voxel as the number of unique elements over the ensemble of models for the corresponding voxel (Lindsay et al., 2012). Computed for all voxels, it provides a cardinality voxet of the same dimensions as the model voxets. It assumes that all voxets of the model ensemble have the same dimensions. By definition, this indicator can be computed on discrete or categorical property fields. For continuous property fields, we propose to use the range between the minimum and maximum value, the standard deviation. Eventually, these continuous indicators can be normalised over the voxet and then averaged or weighted to provide another indicator. Figure 6 shows a cardinality voxet computed from an ensemble of lithocode voxets as well as the range, standard deviation and their normalised average from an ensemble of density voxets.

3.2 Entropy

Shannon's entropy (see Eq. 1) is a specific case of the Rényi entropy when its parameter α converges to 1 (Rényi, 1961). It has been already been applied to geomodels for a few decades (Journel and Deutsch, 1993; Wellmann and Regenauer-Lieb, 2012) and is a finer way than cardinality to describe uncertainty over an ensemble of models, as it takes into account the histogram proportions of the unique values encountered. Let us consider the categorical or discrete variable X that represents a voxel property and assume that it can take n distinct values among an ensemble of voxets. By denoting the probability of observing the ith possible value as p_i, the entropy H of X is computed as follows:

$$H(X) = -\sum_{i=1}^{n} p_i \ln p_i. \tag{1}$$

For continuous property fields, one needs to discretise the continuous domain and integrate along the width of the bins (Marsh, 2013). Here, for the considered example, we choose 50 regular bins. Figure 7 displays Shannon's entropy for the lithocode voxets and the continuous entropy for the other ensemble of property fields: magnetic field, gravity field, density and magnetic susceptibility.

3.3 Histogram dissimilarity

Given a pair of voxets V_P and V_Q, we measure the dissimilarities between their histograms by computing a symmetrised and smooth version of the Kullback-Leibler divergence (Kullback and Leibler, 1951) known as the Jensen-Shannon divergence or total divergence to the average (Dagan et al., 1997). Given two random variables P and Q, the Jensen-Shannon divergence is computed as $\mathrm{JSD}(P||Q) = \frac{1}{2}\mathrm{KLD}(P||M) + \frac{1}{2}\mathrm{KLD}(Q||M)$, where $M = \frac{P+Q}{2}$ and KLD is the Kullback-Leibler divergence. It requires P and Q to share the same support $\mathcal{X}$ and can be computed for continuous or discrete variables. Here, we assume that for our pair of voxets V_P and V_Q, the support of our random variables P and Q, respectively, is discrete and of size n (possibly n bins for discretised continuous variables).

Denoting the support of P and Q by $(x_i)_{i=1\ldots n}$, $p_i = \mathrm{Prob}(P = x_i)$, $q_i = \mathrm{Prob}(Q = x_i)$ and $m_i = \frac{p_i+q_i}{2}$, the Jensen-Shannon divergence is computed as in Eq. 2.

$$\mathrm{JSD}(P||Q) = \frac{1}{2}\sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{m_i}\right) + \frac{1}{2}\sum_{i=1}^{n} q_i \ln\left(\frac{q_i}{m_i}\right). \tag{2}$$

3.4 Semi-variogram dissimilarity

Introduced by Matheron (1963), the semi-variogram measures the dissimilarity of values taken by random variables at different spatial locations as a function of distance. Assuming stationarity, it can be written as

$$\gamma(\boldsymbol{h}) = \frac{1}{2} E\left[(Z(\boldsymbol{s}) - Z(\boldsymbol{s}+\boldsymbol{h}))^2\right],$$

where $\boldsymbol{s}$ denotes a spatial location vector, $\boldsymbol{h}$ denotes a vector and Z is the random field of interest.

Using spatial samples of a random variable, and segmenting the $\boldsymbol{h}$ space into radial distance lags, it is then possible to compute an experimental or empirical semi-variogram over n lags of width δ as follows:

$$\hat{\gamma}(h_i) = \frac{1}{2N_i}\sum_{(j,k)} |Z(\boldsymbol{s}_j) - Z(\boldsymbol{s}_k)|^2,$$

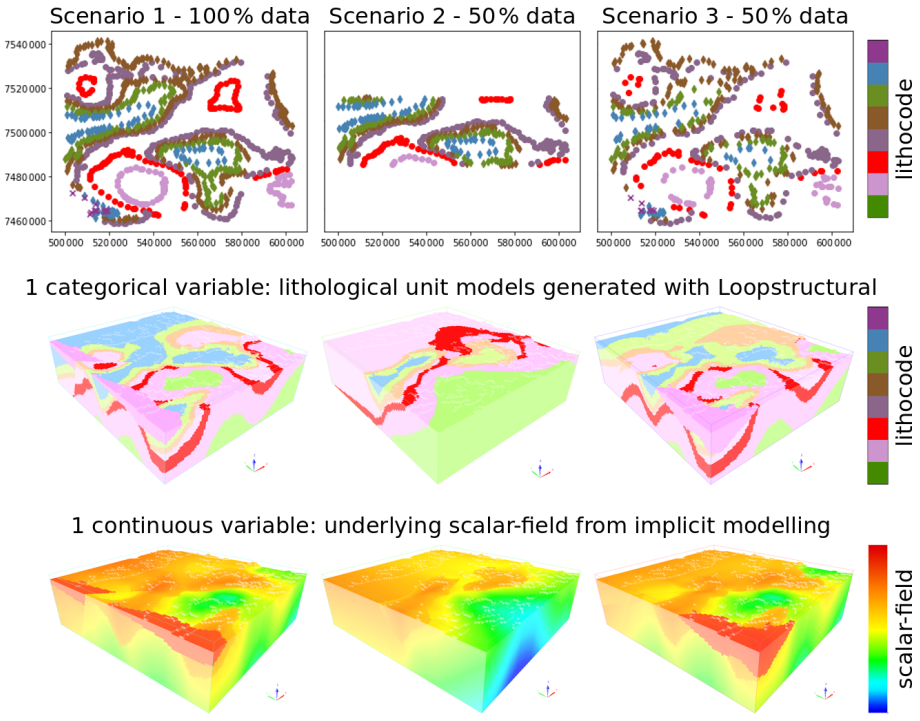


Figure 5. Example model set of lithocode and scalar fields for each of the three scenarios; the left column illustrates input data and model set for scenario 1 when all data are considered; the middle column illustrates input data and model set for scenario 2 when 50 % of the data within a band are considered; the right column illustrates input data and model set for scenario 3 where each input data are decimated with a probability of 50 %.

where $h_i = (i - \frac{1}{2})\delta$ is the centre of the ith lag, $1 \leq i \leq n$ and N_i is the number of pairs (j,k) of points such that $(i-1)\delta \leq ||\boldsymbol{s}_j - \boldsymbol{s}_k|| \leq i\delta$. Note, that we obtain here an omnidirectional semi-variogram; it will be less sensitive to anisotropy but simpler to assess than many directional semi-variograms as anisotropic directions.

Given two empirical semi-variograms $\hat{\gamma}_1$ and $\hat{\gamma}_2$ (see e.g. Fig. 8), we propose to use a weighted l_p norm as defined in Eq. (3).

$$||\hat{\gamma}_1 - \hat{\gamma}_2||_p = \left(\frac{1}{\sum\limits_{1\leq i\leq n} \frac{1}{h_i}} \sum_{1\leq i\leq n} \frac{1}{h_i} |\hat{\gamma}_1(h_i) - \hat{\gamma}_2(h_i)|^p \right)^{\frac{1}{p}}, \quad (3)$$

where $p = 2$ in the following examples. Note that the weight is inversely proportional to the lag distance, giving more importance to dissimilarities of the semi-variogram for small distances. This allows to account for the structural noise and (dis)continuity of the property fields. Thus, the experimental semi-variogram is not always well-behaved at short distances, in particular when dealing with sparse data. However, this is not a concern when considering fully populated voxets, such as in our case.

3.5 Connectivity dissimilarity

The existence of preferential flow-paths or barriers in the subsurface often has a strong impact in many geoapplications. Their characterisation can improve the management of groundwater quality, the extraction of geothermal energy, and help mitigate the environmental impact related to either the production of non-renewable and renewable resources from the subsurface or the sequestration of carbon dioxide and waste (e.g nuclear waste). Renard and Allard (2013) have shown that connectivity cannot be captured by topological indicators, such as the Euler characteristic, or by 1-point or 2-point statistics (e.g. by histogram or semi-variogram analysis, respectively). However, they have shown how a global percolation metric $\Gamma(p)$ and a lag distance connectivity function $\tau(\boldsymbol{h})$ are useful to characterise the connectivity of binary, categorical or continuous property fields. Connectivity indicators have also been used in multiple-point statistics applications to characterise the quality of stochastic simulations with respect to a training image (Meerschman et al., 2013) and in hydrogeophysical application for model selection (Pirot et al., 2019).

Let us consider a binary spatial variable $X \in 0, 1$, and a distance $\boldsymbol{h}$. Then, the lag distance connectivity function $\tau(\boldsymbol{h})$ is defined as the probability that two $\boldsymbol{h}$-distant points $\boldsymbol{s}$ and $\boldsymbol{s} + \boldsymbol{h}$, whose value of $X = 1$, are connected. For a binary

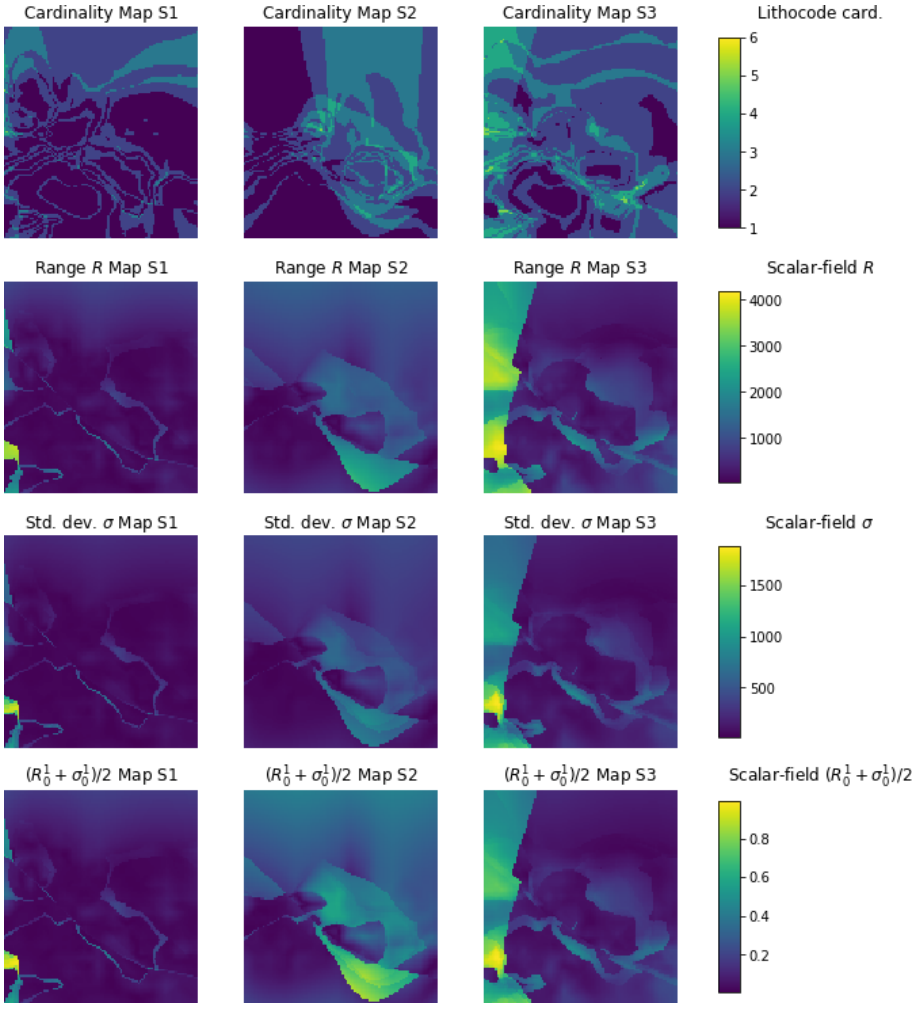


Figure 6. Horizontal sections of cardinality voxets for a categorical property field (first row) or similar indicators for a continuous property field (second–fourth row); the first row shows the cardinality over the ensemble of lithocode voxets, the second row displays the (max-min range) over the ensemble of scalar field voxets, the third row presents the standard deviation over the ensemble of scalar field voxets, the fourth row displays the averaged normalised range and standard deviation over the ensemble of scalar field voxets; the left column refers to scenario 1, the middle column to scenario 2 and the right column to scenario 3.

voxet, two voxels are connected if a path through the face of successive neighbour voxels with the same property exists. The lag distance connectivity function (see Fig. 9) can be written as

$$\tau(\boldsymbol{h}) = \mathrm{Prob}\left(\boldsymbol{s} \overset{\text{connected}}{\leftrightarrow} \boldsymbol{s}+\boldsymbol{h} \;\middle|\; X(\boldsymbol{s})=1, X(\boldsymbol{s}+\boldsymbol{h})=1\right).$$

Now let us assume that the percolation threshold p produces a binary voxet characterised by the binary spatial variable $X \in 0,1$. The global percolation metric $\Gamma(p)$ (see Fig. 10) is the proportion of the pairs of voxels that are connected amongst all the pairs of voxels for which $X=1$:

$$\Gamma(p) = \frac{1}{n_p^2}\sum_{i=1}^{N(X_p)} n_i^2 = \sum_{i=1}^{N(X_p)} p_i^2,$$

where $N(X_p)$ is the number of distinct connected components formed by voxels of value $X=1$ and $p_i = n_i/n_p$ is the proportion of voxels forming the ith distinct connected component, n_i being the size in voxels of the ith connected component and n_p being the total number of voxels of value $X=1$ in the voxet.

Conversely, for the complementary set of voxels for which $X=0$, we can compute $\Gamma(p)^{\mathrm{c}}$ (see Fig. 10). One can note that the two connectivity metrics are related as $\sum_h \tau(\boldsymbol{h}) = n_p\Gamma(p)$ (Renard and Allard, 2013).

We propose two measures of connectivity dissimilarity between voxets, based either on $\tau(\boldsymbol{h})$ or $\Gamma(p)$ and $\Gamma(p)^{\mathrm{c}}$. Let us denote by N_{c} the number of considered classes of values. Note that for a categorical variable voxet, the classes are defined by the category values, while for continuous variable

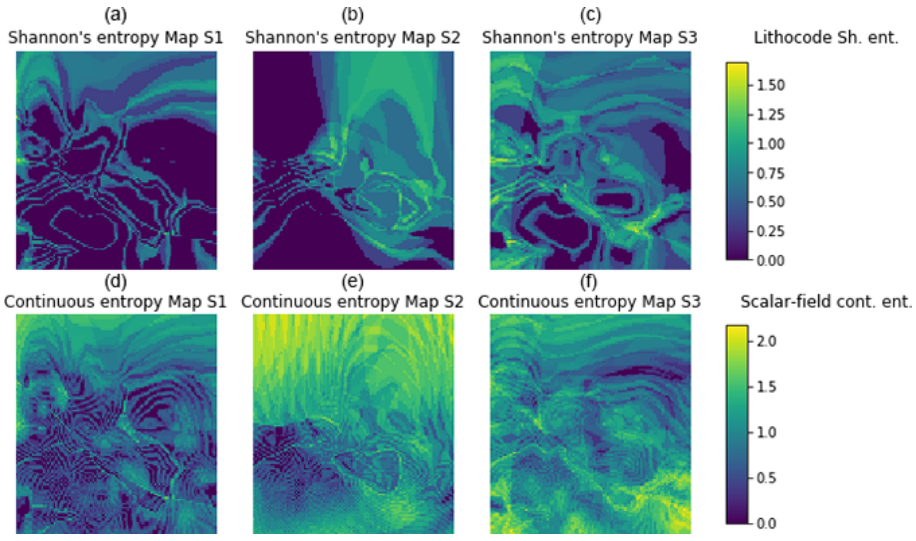


Figure 7. Horizontal sections entropy voxets; **(a–c)** show Shannon's entropy over the ensemble of lithocode voxets, **(d–f)** display the continuous entropy over the ensemble of scalar field voxets; **(a)** and **(d)** refer to scenario 1, **(b)** and **(e)** to scenario 2 and **(c)** and **(f)** to scenario 3.

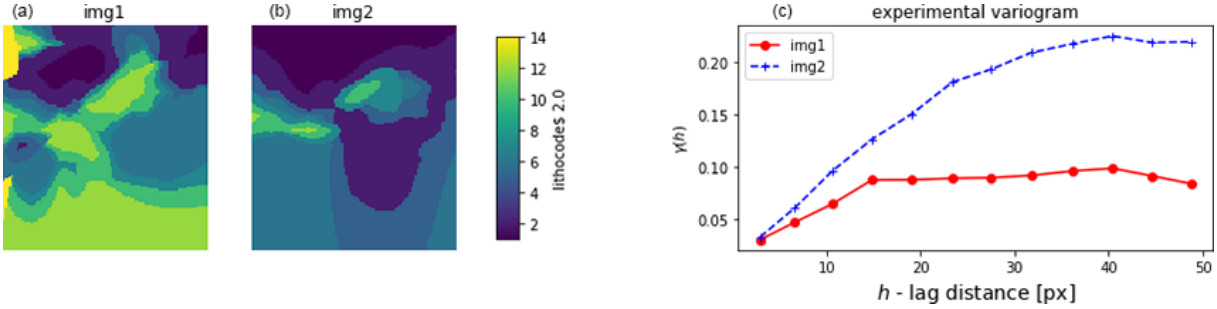


Figure 8. Example of experimental semi-variogram for two lithocode voxets, computed for the lithocode value of 2; **(a)** and **(b)** show a horizontal section for each voxet; **(c)** shows the two corresponding experimental semi-variogram.

voxet, they can be obtained by thresholding with N_c percolation thresholds p. Let us consider N_{lag} the number of lags (defined similarly as for the experimental semi-variogram in the previous subsection), $l_p = 2$ the distance norm and $\mathbb{1}_\tau = 1 - \mathbb{1}_\Gamma$ the indicator allowing to choose between τ or Γ connectivity. Then we can compute the connectivity dissimilarity between two voxets as follows in Eq. (4):

$$d_{\mathrm{CTY}}(\mathrm{Voxet}_1, \mathrm{Voxet}_2) = \sum_{i=1}^{N_c} \frac{1}{N_c} \left[\mathbb{1}_\tau \left(\sum_{\boldsymbol{h}=1}^{N_{lag}} \frac{|\tau_1(\boldsymbol{h}) - \tau_2(\boldsymbol{h})|^{l_p}}{N_{lag}} \right) + (1 - \mathbb{1}_\tau) |\Gamma_1(i) - \Gamma_2(i)|^{l_p} \right]. \quad (4)$$

3.6 Multiple-point histogram dissimilarity

Multiple-point histograms (MPH, Boisvert et al., 2010) are based on pattern recognition and have been primarily used in the field of geostatistics (Guardiano and Srivastava, 1993) to quantify the quality of multiple-point statistics simulations. Patterns are delimited by a search window whose dimensionality matches the one of the dataset. One can count unique patterns; however, the number of unique patterns might be relatively important, in particular for continuous property fields. In that case, it might require that the analysis be restrained to the most frequent patterns (Meerschman et al., 2013). An alternative is to base the analysis on pattern cluster representatives (see Fig. 11). Here, using an l_2 norm distance between patterns and k-means clustering (Pedregosa et al., 2011, scikit-learn implementation), we classify all patterns into $N_c = 10$ clusters. Each cluster centroid or barycentre defines its representative.

In addition, voxets can be easily upscaled, which allows MPH analysis of potentially large scale features with a small search window at a high level of upscaling. Note that at a given level l of upscaling, the size of the dataset is divided by 2^l along each dimension. Here, to avoid property values smoothing, we perform a stochastic upscaling, i.e. in a 2D case, the upscaled value of a 2×2 subset of pixels is achieved

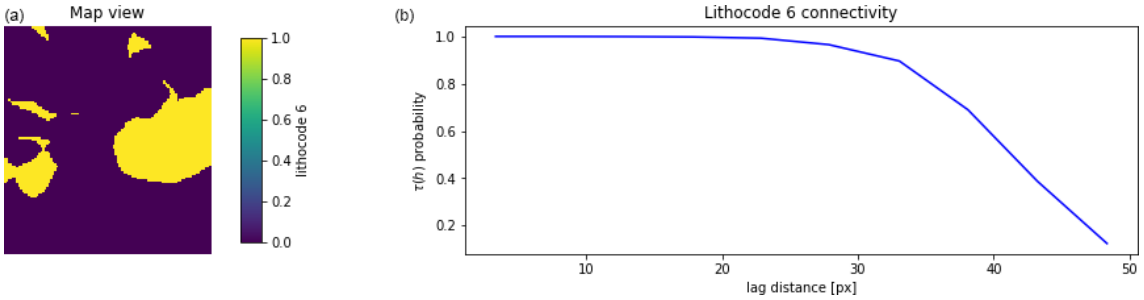


Figure 9. Illustration of the $\tau(\boldsymbol{h})$ connectivity function **(b)** computed on a 2D horizontal section from a binary 3D voxet **(a)**.

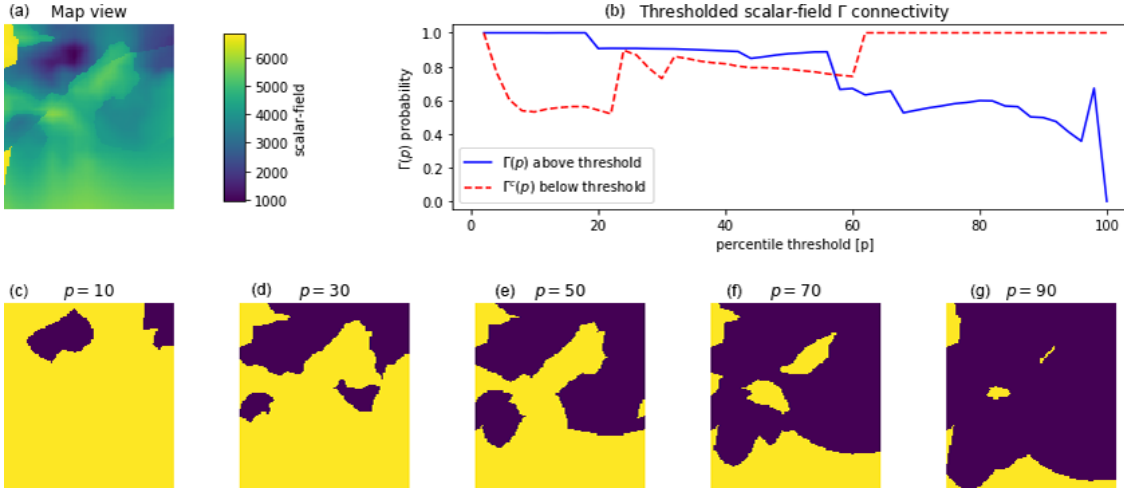


Figure 10. Illustration of a scalar field voxet horizontal section **(a)** and its global percolation metrics $\Gamma(p)$ and $\Gamma^{\mathrm{c}}(p)$ **(b)** computed over the 2D sections; **(c–g)** display horizontal sections of binary fields obtained by applying different percolation threshold $p[\%]$ to the scalar field 3D voxet; blue areas, above the percolation threshold, contribute to $\Gamma(p)$; yellow areas, below the percolation threshold, are used to compute $\Gamma^{\mathrm{c}}(p)$.

by a uniform random draw among the values of the 4 pixels. Cluster pattern identification is performed at the initial resolution level ($l = 0$) and at all possible upscaled levels.

For a given upscaling level, once k-means clustering of patterns has been performed on two voxets or datasets, distances between cluster representatives of two images can be computed

$$d(C_1^i, C_2^j) = \left(\sum_{w=1}^{N_{\mathrm{w}}} \left(C_1^i(w) - C_2^j(w) \right)^2 \right)^{\frac{1}{2}},$$

where C_1^i is the ith cluster representative for voxet 1, C_2^j is the jth cluster representative for voxet 2 and w denotes the index of the window search elements.

The cluster representatives between two datasets are paired by similarity (smallest distance), and reordered such that $\forall i,\ 1 \leq i \leq N_{\mathrm{c}}$, and C_1^i is paired with C_2^i. To account for cluster size differences, the distance between paired cluster representative are weighted by proportion dissimilarities. It results in an MPH cluster-based distance between voxets or datasets 1 & 2 defined as in Eq. (5).

$$d_{\mathrm{MPH}}(\mathrm{Voxet}_1, \mathrm{Voxet}_2) = \sum_{i=1}^{N_{\mathrm{c}}} \frac{1}{N_{\mathrm{c}}} \left[\left(1 + d(C_1^i, C_2^i) \right) \times \left(1 + \frac{|p_1^i - p_2^i|}{p_1^i + p_2^i} \right) - 1 \right], \quad (5)$$

where p_1^i and p_2^i are the proportions of the paired clusters C_1^i and C_2^i with respect to voxets 1 & 2, respectively.

One advantage of selecting cluster representatives independently between two voxets is to lower computing requirements over large ensembles of voxets, performing the analysis for N_{v} voxets instead of $\frac{N_{\mathrm{v}}(N_{\mathrm{v}}-1)}{2}$ pairs. However, performing R pattern clustering on two datasets might provide a more accurate and precise way to compute a distance between histograms with the same support of cluster representatives, thus allowing the use of Jensen-Shannon divergence for instance. One can note that we accounted for the size of the clusters but we could also consider the density spread or concentration around cluster representatives.

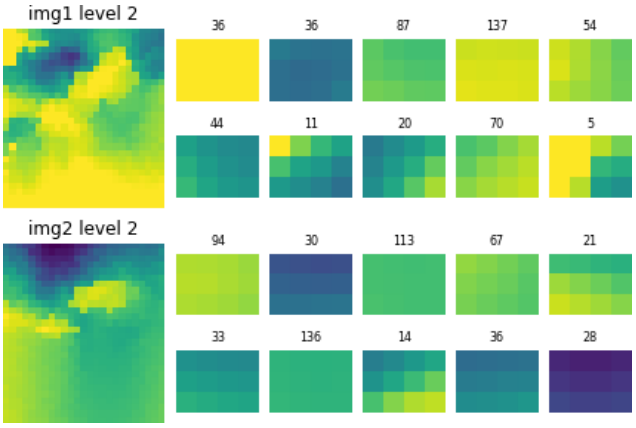


Figure 11. Illustration of multiple-point histogram cluster representatives and sizes for two scalar field horizontal sections, at the 3rd upscaling level; the left column shows the two 2D voxets; for the other columns the first and second rows, and the third and fourth rows, display the 10 cluster representatives and their size (number of counted patterns attached to the cluster representative) for the first voxet and for the second voxet, respectively; their order reflects the best similarity between the cluster representatives for both voxets.

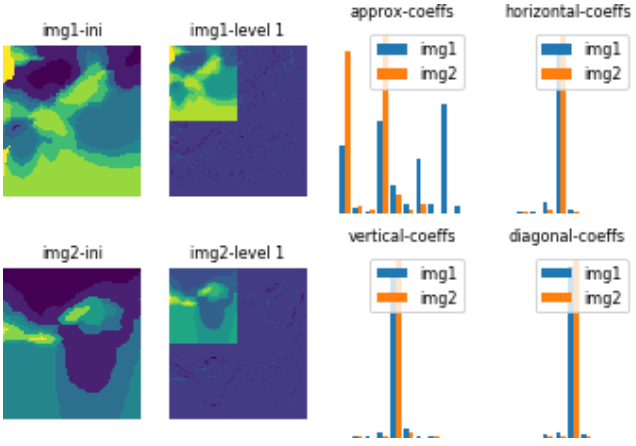


Figure 12. Illustration of a first level of Haar-wavelet decomposition and the resulting coefficient histograms for two lithocode voxet horizontal sections.

3.7 Wavelet decomposition coefficient dissimilarity

Wavelet decomposition is a way to compress images. Each level of decomposition produces a series of coefficients. If computed for images to be compared, the dissimilarity of histograms of coefficients can be computed with the Jensen-Shannon divergence (Eq. 2). Here, wavelet decomposition (Fig. 12) is performed with the *PyWavelets Python* package (Lee et al., 2019) at all possible levels of decomposition and using the "Haar" wavelet. Other wavelets could be used; however, tests have shown that such dissimilarity measures are not very sensitive to the choice of the wavelet (Pirot et al., 2019).

Then a wavelet-based dissimilarity measure between two voxet can be computed as in Eq. (6) by summing the Jensen-Shannon divergences computed for all pairs of approximation and decomposition coefficients at all possible levels:

$$d_{\mathrm{WVT}}(\mathrm{Voxet}_1, \mathrm{Voxet}_2) = \sum_{i,j} \frac{\mathrm{JSD}(C_1^{i,j} || C_2^{i,j})}{\sum_{i,j} 1}, \tag{6}$$

where $C_1^{i,j}$ and $C_2^{i,j}$ denote the distributions of the ith coefficients at upscaling level j for Voxet_1 and Voxet_2, respectively.

3.8 Topological dissimilarity

Thiele et al. (2016) give an overview of possible representations of the topology in the context of 3D geological modelling. Different levels of complexity (e.g. 1st or 2nd orders…) can be used. Nonetheless, any topological indicator is a graph that can take the form of an adjacency matrix. Therefore, to compute a topological distance between two 3D geological models (for instance as in Giraud et al., 2019), it seems natural to look at distances defined between graphs. Donnat and Holmes (2018) provided a comprehensive review of graph distances used in the study of graph dynamics or temporal evolution. Although, here, in a geological context, we aim at comparing the topological diversity of an ensemble of geological models, we can use similar distances. Donnat and Holmes (2018) classified graph distances into three main categories as summarised below.

Low-scale distances capture local changes in the graph structure. The Hamming (structural) distance is the sum of the absolute value of differences between two adjacency matrices and requires the same number of vertices (nodes) between the graphs – note that it is a specific case of the more general graph edit distance. The Jaccard distance is defined as the difference between the union and intersection of two graphs. The graph edit distance belongs to the NP-complete class of problems, and is not computed here. More information is available in (Gao et al., 2010) or in part IV Chapter 15 of the *Encyclopedia of Distances* (Deza and Deza, 2009, p. 301). Note that some packages and implementations to compute the graph-edit distance exist but have not been tested here (GMatch4py, last access: 8 June 2022; graphkit-learn, last access: 8 June 2022; other proposed heuristic, last access: 8 June 2022).

High level or spectral distances are global measures and reflect the smoothness of the overall graph structure changes by measuring dissimilarities in the eigenvalues of the graph Laplacian or its adjacency matrix. Some examples are the IM distance (Ipsen and Mikhailov, 2003), l_p distances on eigenvalues and the Kruglov distance on eigenvector coordinates (Shimada et al., 2016).

Meso-scale distances are a compromise or combination of low-scale and spectral distances: Hamming-IM (HIM) com-

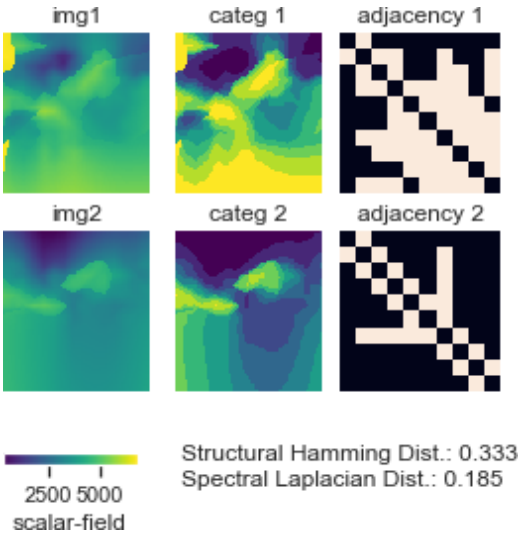


Figure 13. Illustration of topological distances and adjacency matrices in the right column for two categorised voxets in the middle column, derived from two scalar field voxet horizontal sections in the left column.

bination, heat-wave distance, polynomial distance, neighbourhood level distances and connectivity-based distances.

Here, we propose to build first order adjacency matrices (see Fig. 13) from 2D or 3D voxet models. For continuous property fields, the voxet is discretised into $N = 10$ classes of values defined by N equipercentile thresholds over the distribution of the combined voxets. We compute two topological distances: the structural Hamming distance and the Laplacian spectral distance (Shimada et al., 2016).

Note that graphs characterising geological model topology could be defined as attributed graphs, to contain more information (edges properties, such as age constraint, type of contact; vertices properties, such as formation type, geophysical properties). Thus, more specific measures could be developed to take such characteristics into account; however, it would rely on the ability of geomodelling engines to provide these topology graphs with each model, and there is no guarantee that it would be meaningful for the inference of geochronology from geophysics.

3.9 Results: indicator comparison

Local measures of uncertainty (see Figs. 14 and 15) and global indicators (see Figs. 17 and 18) have been computed for 2D, 3D, categorical and continuous variable voxets (lithocode, scalar field) for an ensemble of 30 model sets. We also provide a comparison of the required computing time for the different indicators and highlight the contributing factors or parameters (see Table 2).

One can see from Fig. 14 that Shannon's entropy and cardinality computed from lithocode voxets can have a good correlation. However, the following equivalent indicators: continuous entropy, averaged normalised range and standard deviation computed from scalar field voxets, (Fig. 15) have very little in common. Indeed, the standard deviation or the range of values is sensitive to extremely different values, while the continuous entropy is sensitive to the proportion of categories of values.

Figure 16 shows that histogram dissimilarities computed from lithocode or scalar field voxets have a good correlation. Multidimensional scaling (MDS) plots reveal that dissimilarities are smallest within scenario 1 and then within scenario 3 while they are greatest within scenario 2. The MDS plots show that scenarios 1 and 3 model sets overlap while scenario 2 model sets are characterised by less similar histograms and thus the scenario 2 sample cloud of points form a distinct cluster.

Figure 17 shows some correlation between histogram, semi-variogram, wavelet and structural Hamming-based measures from the lithocode voxets. Figure 18 shows a good correlation between histogram dissimilarity, wavelet-based dissimilarity and structural Hamming distance from the scalar field voxets but not as strong as when computed from the lithocode voxets.

4 Discussion

Several factors might explain why the majority of practitioners do not consider input data uncertainty but all are related to the limited resources available to practitioners (knowledge, algorithms, computing time, project timeframe or funding). One of them is that data uncertainty is not quantified at the time of data acquisition or not available for some measurements, which is the case when only one measurement or observation of geological data is made at a given location (e.g. for azimuth, dip, and lithology). Reasonable metadata standards may help to enforce error quantification, or at the very least provide some information about the nature of data collection as to infer where and what magnitude of error may be present but even these are poorly or not recorded. Although repeated independent measurements would provide uncertainty estimates, procedures and limited time or budget resources are a hindrance. Sometimes, knowing the survey set-up and the instrumentation characteristics, such as their precision and accuracy might avoid repeating field measures and allow an estimation of measurement uncertainty. Inversion algorithms used for geophysical data integration can also provide estimates of geophysical data errors.

A second reason is related to the required time and associated costs of modelling. Indeed, the process of data integration only uses a very limited amount of automation, thus the generation of a single model already consumes most of the practitioners' resources. In addition, the complexity of real-world data often leads to a substantial number of parameters. Thus for high-dimensional problems, uncertainty propagation requires sufficiently large model ensembles to be rep-

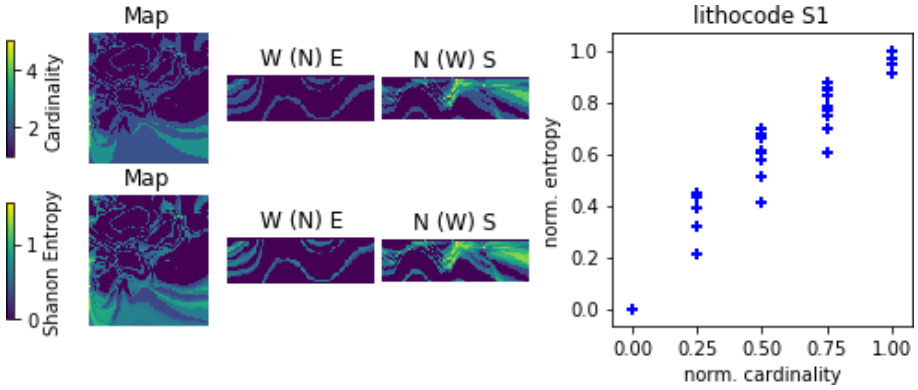


Figure 14. Comparison of local uncertainty measures for an ensemble of 10 lithocode 3D voxets for scenario 1; 3D visualisation looking from the NW of the voxet, the top surface of the voxet an EW section at the northern face of the model looking from the south, a NS section on the western face of the voxet looking from the east.

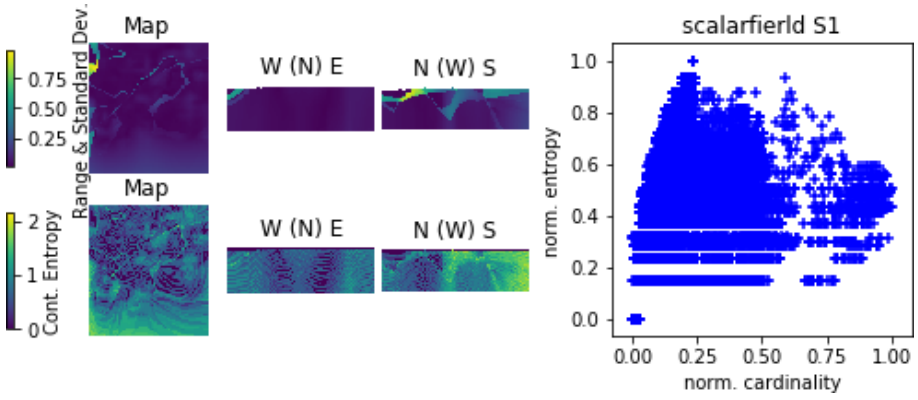


Figure 15. Comparison of local uncertainty measures for an ensemble of 10 scalar-field 3D voxets for scenario 1; 3D visualisation looking from the NW of the voxet, the top surface of the voxet an EW section at the northern face of the model looking from the south, a NS section on the western face of the voxet looking from the east.

resentative, which might not be compatible with the limited resources available to the practitioners.

A third reason is due to the fact that assumptions, such as choices in geological interpretations made during the modelling process, are not always tracked and when they are, they are often "forgotten" at the next stage of the modelling workflow (Jessell et al., 2018). When these assumptions or justifications are recorded, they are described in metadata or in distinct reports. Consequently, conceptual uncertainty, which describes alternative yet plausible stratigraphy, tectonic and geodynamic settings, is also ignored.

Another possible reason is that uncertainty is ignored out of convenience (Ferré, 2017) or by ignorance, lack of knowledge or education about the importance of uncertainty quantification. However, the formulation of the collected answers suggests that this is not the case for the surveyed practitioners, who tend to acknowledge the importance and need for tools or algorithms to integrate uncertainty quantification in their modelling workflow.

While about 11 % of the respondents indicated that they perform a quantitative uncertainty quantification, it is limited to aleatoric uncertainty, i.e. data measurement errors. However, the lack of spatial data samples contributes to epistemic uncertainty and our limited contextual knowledge adds up to conceptual uncertainty. In addition, it is generally expected that these sources of uncertainty have a bigger impact on predictive uncertainty (Pirot et al., 2015). Thus, in addition to developing tools to facilitate aleatoric uncertainty quantification for practitioners, accessible tools integrating epistemic and conceptual uncertainty quantification need to be developed and promoted in the minerals industry.

Restricting the survey to a few open questions encouraged participants to take the survey and to express their uses, needs and opinion with limited perception bias; however, it did not allow quantification of the gathered answers, and often required some interpretation. Nevertheless, it is a first step in acknowledging the practices and needs of 3D geological modellers in the minerals industry. Another limitation of the survey is that it did not look at the practice and needs of

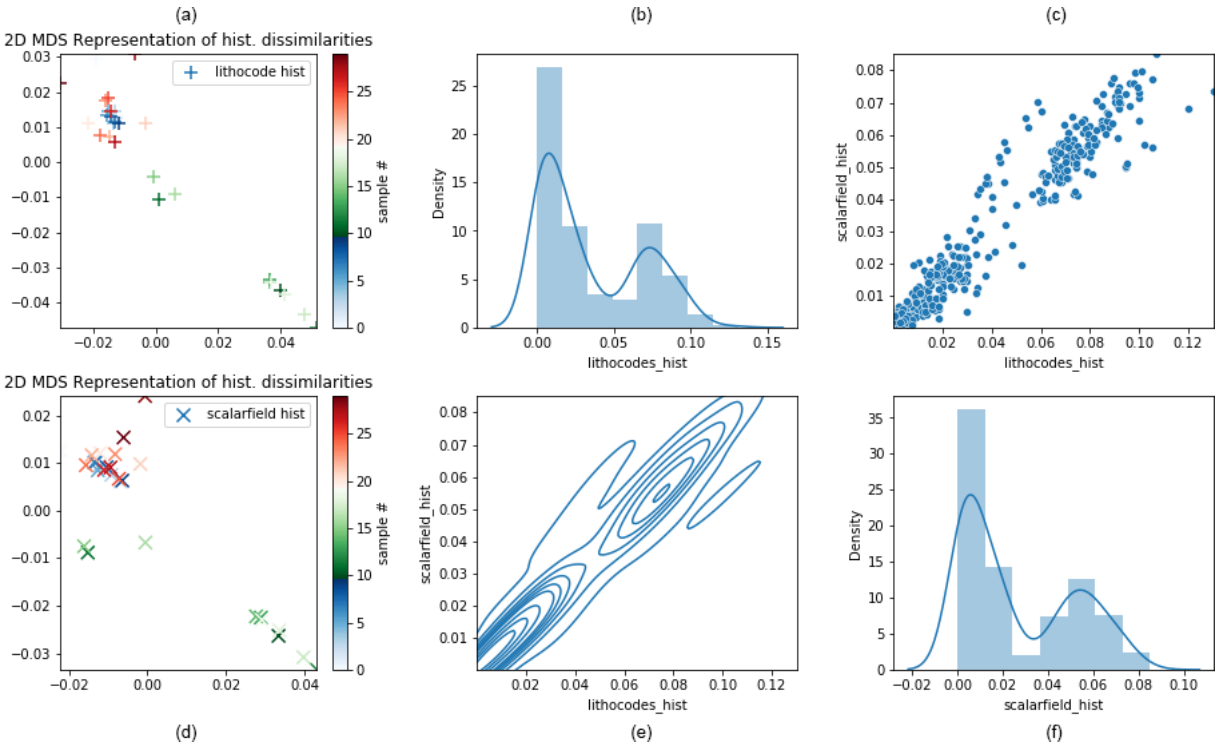


Figure 16. Comparison of model set histogram dissimilarities across the three scenarios; **(a)** and **(d)** display a 2D MDS representation of histogram dissimilarities computed from lithocode voxets **(a–c)** and from scalar-field voxets **(d–f)**; samples 0–9 belong to scenario 1, samples 10–19 belong to scenario 2 and samples 20–29 belong to scenario 3; the subplots of **(b)**, **(c)**, **(e)**, and **(f)** show histogram and density, joint density and cross-plot between histogram dissimilarities computed from lithocode voxets or from scalar field voxets.

Table 2. Complexity and computing time for local and global measures of uncertainty using a single Intel®Core™ i7-8550 1.80 GHz processor, based on an ensemble size of $N = 10$ geological models.

Measures	Number of evaluations	Total CPU time (HH:MM:SS)	Influential parameters
Cardinality	2	00:00:02	n_{voxels}, n_{voxets}
Entropy	2	00:00:05	n_{voxels}, n_{voxets}, n_{bins}
Histogram dissimilarity	60	00:00:38	n_{voxels}, n_{voxets}
Semi-variogram dissimilarity	60	00:09:42	n_{voxels}, n_{voxets}, n_{categ}, $\text{rate}_{\text{sub-sampling}}$
Connectivity dissimilarity	60	00:07:02	n_{voxels}, n_{voxets}, n_{categ}, $\text{rate}_{\text{sub-sampling}}$
Multiple-point histogram dissimilarity	870	00:37:45	n_{voxels}, n_{voxets}, $\text{rate}_{\text{sub-sampling}}$
Wavelet decomposition coefficients dissimilarity	870	00:00:38	n_{voxels}, n_{voxets}
Topological distances	870	00:02:49	n_{voxels}, n_{voxets}, n_{bins}

other fields, such as the petroleum industry (Scheidt et al., 2018), geothermal industry (Chen et al., 2015) or hydrogeology (Pirot et al., 2019), although they share similar scientific problems and also propose interesting solutions to deal with uncertainty quantification.

For model ensembles of size n, the calculation of local uncertainty indicators, such as cardinality or entropy voxets seems to be much faster $O(n)$ to compute than global indicators $O(n^2)$, in particular when dealing with discrete or categorical variables. In addition, local indicators are most convenient to visualise uncertainty in 2D or 3D spaces. However, they cannot provide information about the variability of important specific features, such as connectivity or topology, that can be estimated with global uncertainty indicators. Thus, depending on the modelling objectives and relevant features or characteristics, both local and global uncertainty indicators should be considered.

Nevertheless, a few recommendations can be made regarding the selection of uncertainty indicators. When looking at the most shallow learning curve for a typical modeller or

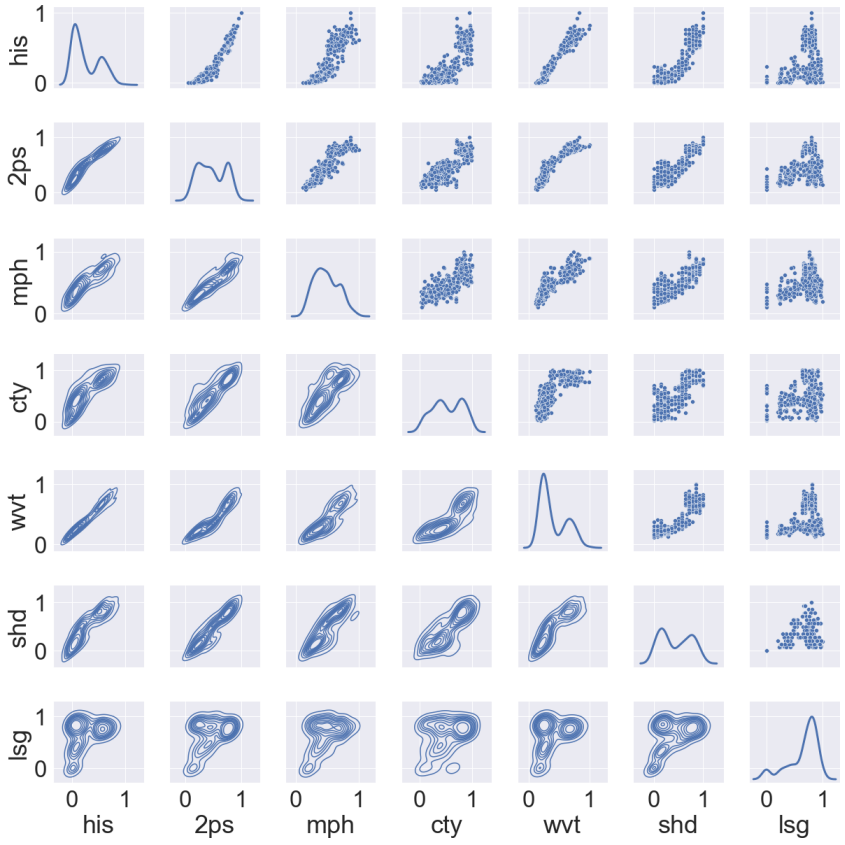


Figure 17. Comparison of global normalised dissimilarity measures for an ensemble of 30 lithocode 3D voxets across the 3 scenarios; cross-plots and density plots by pair of normalised dissimilarity measures; his – histogram, 2ps – semi-variogram, mph – multiple-point histogram, cty – connectivity, wvt – wavelet decomposition coefficients, shd – topological structural Hamming distance, lsg – topological Laplacian spectral distance.

geologist, a good practice might be to start with cardinality, which simply states how many different lithologies are present at a given location. Then entropy will become more appropriate when the geologist starts to compare ensembles of models with different stratigraphies (in which case the total number of lithological units will change, making cardinality inappropriate). A following step in the learning curve might then be the use of topological or connectivity-based indicators to better understand the impact of stratigraphic uncertainty. Alternatively, a good strategy might be to compute as many indicators as can be afforded in the allocated computing budget (see Table 2 for indicative computing requirements). Another possibility would be to select a subset of voxets from the whole ensemble on which all indicators are computed and then perform a selection of the most informative or suitable ones prior to computing them on the whole set.

Presumably, the presented indicators are not exhaustive and remain a subjective choice (Wellmann and Caumon, 2018), even though all of them are already used rather individually in the geomodelling community. Pellerin et al. (2015), for instance, proposed other specific global geometric indicators. Here, we have focused on indicators that have shown some usefulness and practicality. Local indicators could be extended to higher moments of the voxel-valued probability distributions, such as to consider asymmetry. Global indicators could be extended to dissimilarity measures of transition probabilities, transiograms or cross-to-direct indicator variogram ratios as well as measures derived from computer graphics or deep learning techniques. One could also consider summary metrics of lower dimensional model representations that could be obtained from discrete-cosine transform (e.g. Ahmed et al., 1974), (kernel) principal component analysis (e.g. Schölkopf et al., 1997) or (kernel) auto-encoding (e.g. Laforgue et al., 2019). This could be particularly appealing as it could reduce indicator computing costs drastically but might not allow identification of the specific features of interest in a representation space of lower dimensions.

Last but not least, it should be noted that we compared ensembles of models where the underlying characteristics such as various lithological units derived from a shared stratigraphy and scalar fields generated under similar assumptions are consistent. However, we must point out that the various

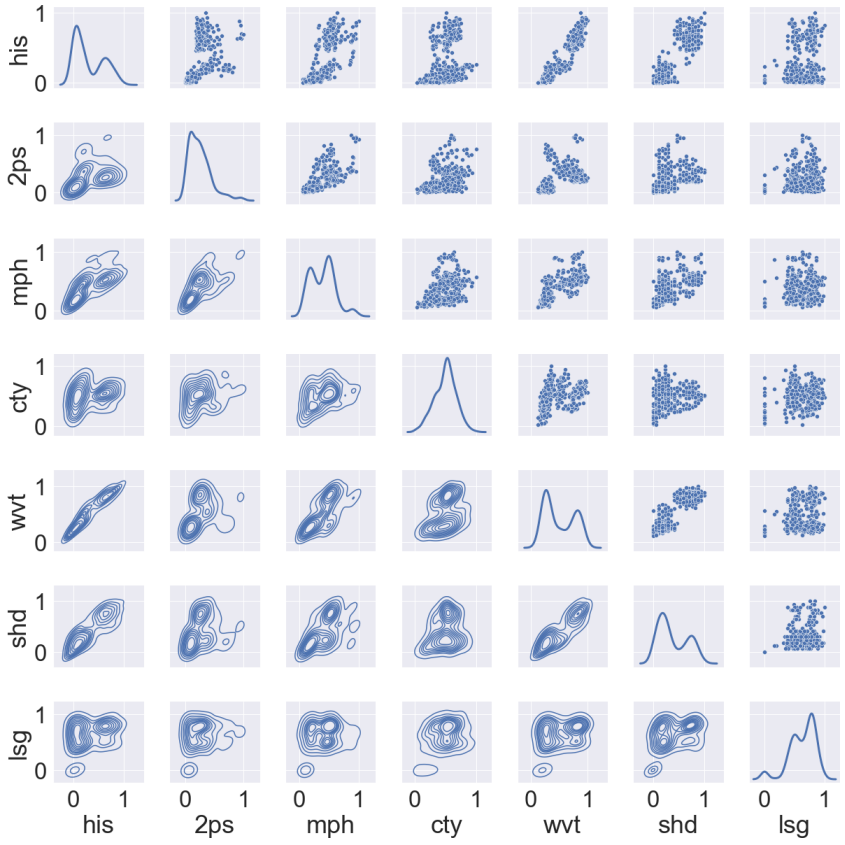


Figure 18. Comparison of global normalised dissimilarity measures for an ensemble of 30 scalar field 3D voxets across 3 scenarios; cross-plots and density plots by pairs of normalised dissimilarity measures; his – histogram, 2ps – semi-variogram, mph – multiple-point histogram, cty – connectivity, wvt – wavelet decomposition coefficients, shd – topological structural Hamming distance, lsg – topological Laplacian spectral distance.

indicators are compatible with differences in property ranges or meanings (for categorical variables), and thus it is the responsibility of the user to ensure the coherence of the model ensemble used as an input for the uncertainty computation.

5 Conclusions

The survey clearly shows that practitioners acknowledge the importance of uncertainty quantification; the majority recognise that they do not perform uncertainty quantification at all and all would like to do better. From this survey we have identified four main factors preventing practitioners from performing uncertainty quantification: lack of data uncertainty quantification, computing requirement to generate one model, poor tracking of assumptions and interpretations and relative complexity of uncertainty quantification. Here, as a first response, we have provided the geomodelling community with `loopUI-0.1`, an open-source *Python* package to compute local and global uncertainty indicators. Then, to increase the confidence in predictions from 3D geological model, efforts should be made to explore conceptual uncertainty (Laurent and Grose, 2020) as well as towards the implementation of systematic geological data uncertainty quantification, and the exploration of parametric and epistemic uncertainty (Pirot et al., 2020). It should be performed appropriately at all scales, across all geoscientific methods, such as the extraction of additional lithological data from drill hole databases (Joshi et al., 2021). To encourage uncertainty propagation among practitioners, accessible and compatible algorithms should be offered (1) to automatically extract geological data from open databases (Jessell et al., 2021) and (2) to quickly generate plausible geological models from a given dataset (Grose et al., 2021) in interaction with geophysical data integration (Giraud et al., 2021) and at an appropriate resolution (Scalzo et al., 2021). This special issue (Ailleres, 2020) already provides elements of an answer to these problems and is expected to host future advances on these topics.

Appendix A: Survey

For each model scale that you encounter in your work, please answer all the questions of the survey (1–10):

A1 Part I – scale

1. What are the scale and characteristic dimensions (width, length, depth and resolution) of the models that you build or use?

Table A1. Investigation scale.

Mine ($<$ 1 km dimension, $\sim$ 10 m resolution)
Brownfields (1–10 km dimension, $\sim$ 100 m resolution)
Greenfields/regional ($>$ 10 km dimension, $\sim$ 1 km resolution)

2. What are the dimensions and grid cell resolution of your models?

Table A2. Model dimensions and resolution.

Model width (m)
Model length (m)
Model depth (m)
Horizontal resolution (m)
Vertical resolution (m)
Main purpose (e.g. resource estimation)

A2 Part II

Output

3. Which objectives do geological models help you to achieve?
4. How are they useful to fulfil other needs (and which ones)?

Input

5. What kind of input data, and what quantity and quality metrics (if any) are used to build your geological models?

Current modelling

6. How is the uncertainty of input data assessed and taken into account?
7. How is the geologist/modeller's interpretation recorded into the model (recorded tracks of assumptions, choices and justifications)?
8. What is the usual modelling workflow and which tools or algorithms are involved?
9. How are data integration and upscaling performed? Which tools or algorithms are involved?

Improvements

10. What are the limitations of existing geological models to achieve your current and future objectives? How do you prioritise them and what kind of solution would you imagine?

Code and data availability. The detailed survey questions are available in Appendix A and the gathered anonymised answers are available on request. The code to compute the uncertainty indicators and a set of illustrative notebooks are available at https://doi.org/10.5281/zenodo.5656151 (Pirot, 2021b).

Author contributions. The survey was initiated by GP, developed jointly with RJ and with the support of all other co-authors. All co-authors distributed the survey and collected answers. Answers were processed and analysed by GP. The manuscript was mainly written by GP, with contributions from all co-authors.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. This work is supported by the ARC-funded Loop: Enabling Stochastic 3D Geological Modelling consortia (LP170100985) and DECRA (DE190100431) and by the Mineral Exploration Cooperative Research Centre whose activities are funded by the Australian Government's Cooperative Research Centre Programme. This is MinEx CRC Document 2021/5.

Financial support. This research has been supported by the Australian Research Council (grant nos. LP170100985 and DE190100431) and by the Mineral Exploration Cooperative Research Centre.

Review statement. This paper was edited by Thomas Poulet and reviewed by three anonymous referees.

References

Ahmed, N., Natarajan, T., and Rao, K. R.: Discrete cosine transform, IEEE T. Comput., 100, 90–93, 1974.

Ailleres, L.: The Loop 3D stochastic geological modelling platform – development and applications, GMD Special Issue, https://gmd.copernicus.org/articles/special_issue1142.html (last access: 8 June 2022), data available at: https://loop3d.github.io/ (last access: 8 June 2022), 2020.

Boisvert, J. B., Pyrcz, M. J., and Deutsch, C. V.: Multiple point metrics to assess categorical variable models, Nat. Resour. Res., 19, 165–175, 2010.

Chen, M., Tompson, A. F., Mellors, R. J., and Abdalla, O.: An efficient optimization of well placement and control for a geothermal prospect under geological uncertainty, Appl. Energ., 137, 352–363, 2015.

Dagan, I., Lee, L., and Pereira, F.: Similarity-based methods for word sense disambiguation, in: Proceedings of the 35th ACL/8th EACL, arXiv preprint, 56–63, https://doi.org/10.48550/arXiv.cmp-lg/9708010, 1997.

de la Varga, M., Schaaf, A., and Wellmann, F.: GemPy 1.0: open-source stochastic geological modeling and inversion, Geosci. Model Dev., 12, 1–32, https://doi.org/10.5194/gmd-12-1-2019, 2019.

Deza, M. M. and Deza, E.: Encyclopedia of distances, in: Encyclopedia of distances, Springer, Berlin, Heidelberg, 1–583, https://doi.org/10.1007/978-3-642-00234-2_1, 2009.

Donnat, C. and Holmes, S.: Tracking network dynamics: A survey using graph distances, Ann. Appl. Stat., 12, 971–1012, https://doi.org/10.1214/18-AOAS1176, 2018.

Ferré, T.: Revisiting the Relationship Between Data, Models, and Decision-Making, Groundwater, 55, 604–614, 2017.

Franke, N. and Von Hippel, E.: Satisfying heterogeneous user needs via innovation toolkits: the case of Apache security software, Res. Policy, 32, 1199–1215, https://doi.org/10.1080/01449290601111051, 2003.

Gao, X., Xiao, B., Tao, D., and Li, X.: A survey of graph edit distance, Pattern Anal. Appl., 13, 113–129, https://doi.org/10.1007/s10044-008-0141-y, 2010.

Giraud, J., Ogarko, V., Lindsay, M., Pakyuz-Charrier, E., Jessell, M., and Martin, R.: Sensitivity of constrained joint inversions to geological and petrophysical input data uncertainties with posterior geological analysis, Geophys. J. Int., 218, 666–688, 2019.

Giraud, J., Ogarko, V., Martin, R., Jessell, M., and Lindsay, M.: Structural, petrophysical, and geological constraints in potential field inversion using the Tomofast-x v1.0 open-source code, Geosci. Model Dev., 14, 6681–6709, https://doi.org/10.5194/gmd-14-6681-2021, 2021.

Grose, L., Ailleres, L., Laurent, G., and Jessell, M.: LoopStructural 1.0: time-aware geological modelling, Geosci. Model Dev., 14, 3915–3937, https://doi.org/10.5194/gmd-14-3915-2021, 2021.

Guardiano, F. and Srivastava, R.: Multivariate geostatistics: beyond bivariate moments, in: Geostatistics Troia 1992, edited by: Soares, A., 133–144, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-011-1739-5_12, 1993.

Ipsen, M. and Mikhailov, A. S.: Erratum: Evolutionary reconstruction of networks, Physical Review E, 67, 039901, https://doi.org/10.1103/PhysRevE.67.039901, 2003.

Jessell, M., Pakyuz-Charrier, E., Lindsay, M., Giraud, J., and de Kemp, E.: Assessing and mitigating uncertainty in three-dimensional geologic models in contrasting geologic scenarios, Metals, Minerals, and Society, 21, 63–74, 2018.

Jessell, M., Ogarko, V., de Rose, Y., Lindsay, M., Joshi, R., Piechocka, A., Grose, L., de la Varga, M., Ailleres, L., and Pirot, G.: Automated geological map deconstruction for 3D model construction using map2loop 1.0 and map2model 1.0, Geosci. Model Dev., 14, 5063–5092, https://doi.org/10.5194/gmd-14-5063-2021, 2021.

Joshi, R., Madaiah, K., Jessell, M., Lindsay, M., and Pirot, G.: dh2loop 1.0: an open-source Python library for automated processing and classification of geological logs, Geosci. Model Dev., 14, 6711–6740, https://doi.org/10.5194/gmd-14-6711-2021, 2021.

Journel, A. G. and Huijbregts, C. J.: Mining geostatistics, Academic Press, New York, 600 p., ISBN 0123910501, 1978.

Journel, A. G. and Deutsch, C. V.: Entropy and spatial disorder, Math. Geol., 25, 329–355, 1993.

Kujala, S.: Effective user involvement in product development by improving the analysis of user needs, Behav. Inform. Technol., 27, 457–473, 2008.

Kullback, S. and Leibler, R. A.: On information and sufficiency, Ann. Math. Stat., 22, 79–86, 1951.

Laforgue, P., Clémençon, S., and d'Alché Buc, F.: Autoencoding any data through kernel autoencoders, in: The 22nd International Conference on Artificial Intelligence and Statistics, PMLR, Naha, Okinawa, Japan, 1061–1069, http://proceedings.mlr.press/v89/laforgue19a/laforgue19a.pdf (last access: 9 June 2022), 2019.

Laine, E., Luukas, J., Mäki, T., Kousa, J., Ruotsalainen, A., Suppala, I., Imaña, M., Heinonen, S., and Häkkinen, T.: The Vihanti-Pyhäsalmi Area, in: 3D, 4D and Predictive Modelling of Major Mineral Belts in Europe, edited by: Weihed, P., Springer International Publishing, Cham, 123–144, https://doi.org/10.1007/978-3-319-17428-0_6, 2015.

Lajaunie, C., Courrioux, G., and Manuel, L.: Foliation fields and 3D cartography in geology: principles of a method based on potential interpolation, Math. Geol., 29, 571–584, 1997.

Laurent, G. and Grose, L.: A Hitchhiking Foray into the Structural Uncertainty Space, 2020 Ring meeting, Nancy, France, https://www.ring-team.org/research-publications/ring-meeting-papers?view=pub&id=4971 (last access: 9 June 2022), 2020.

Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K., and O'Leary, A.: PyWavelets: A Python package for wavelet analysis, Journal of Open Source Software, 4, 1237, https://doi.org/10.21105/joss.01237, 2019.

Li, L., Boucher, A., and Caers, J.: SGEMS-UQ: An uncertainty quantification toolkit for SGEMS, Comput. Geosci., 62, 12–24, 2014.

Linde, N., Ginsbourger, D., Irving, J., Nobile, F., and Doucet, A.: On uncertainty quantification in hydrogeology and hydrogeophysics, Adv. Water Resour., 110, 166–181, 2017.

Lindsay, M. D., Aillères, L., Jessell, M. W., de Kemp, E. A., and Betts, P. G.: Locating and quantifying geological uncertainty in three-dimensional models: Analysis of the Gippsland Basin, southeastern Australia, Tectonophysics, 546, 10–27, 2012.

Lindsay, M. D., Occhipinti, S., Laflamme, C., Aitken, A., and Ramos, L.: Mapping undercover: integrated geoscientific interpretation and 3D modelling of a Proterozoic basin, Solid Earth, 11, 1053–1077, https://doi.org/10.5194/se-11-1053-2020, 2020.

Lochbühler, T., Pirot, G., Straubhaar, J., and Linde, N.: Conditioning of Multiple-Point Statistics Facies Simulations to Tomographic Images, Math. Geosci., 46, 625–645, 2013.

Loop: An open source 3D probabilistic geological and geophysical modelling platform, https://loop3d.org/ (last access: 9 June 2022), 2019.

Marsh, C.: Introduction to continuous entropy, Tech. rep., Department of Computer Science, Princeton University, 1–17, https://www.crmarsh.com/pdf/Charles_Marsh_Continuous_Entropy.pdf (last access: 9 June 2022), 2013.

Matheron, G.: Principles of geostatistics, Econ. Geol., 58, 1246–1266, 1963.

Meerschman, E., Pirot, G., Mariethoz, G., Straubhaar, J., Van Meirvenne, M., and Renard, P.: A practical guide to performing multiple-point statistical simulations with the direct sampling algorithm, Comput. Geosci., 52, 307–324, 2013.

Mo, S., Shi, X., Lu, D., Ye, M., and Wu, J.: An adaptive Kriging surrogate method for efficient uncertainty quantification with an application to geological carbon sequestration modeling, Comput. Geosci., 125, 69–77, 2019.

Osenbrück, K., Gläser, H.-R., Knöller, K., Weise, S. M., Möder, M., Wennrich, R., Schirmer, M., Reinstorf, F., Busch, W., and Strauch, G.: Sources and transport of selected organic micropollutants in urban groundwater underlying the city of Halle (Saale), Germany, Water Res., 41, 3259–3270, 2007.

Pakyuz-Charrier, E., Lindsay, M., Ogarko, V., Giraud, J., and Jessell, M.: 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, 385–402, https://doi.org/10.5194/se-9-385-2018, 2018.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Pellerin, J., Caumon, G., Julio, C., Mejia-Herrera, P., and Botella, A.: Elements for measuring the complexity of 3D structural models: Connectivity and geometry, Comput. Geosci., 76, 130–140, 2015.

Pirot, G.: loopUI-0.1: uncertainty indicators for voxet ensembles, https://doi.org/10.5281/zenodo.5118418, https://github.com/Loop3D/uncertaintyIndicators/, 2021a.

Pirot, G.: Loop3D/loopUI: v0.1 (v0.1-gamma), Zenodo [code] [data set], https://doi.org/10.5281/zenodo.5656151, 2021b.

Pirot, G., Renard, P., Huber, E., Straubhaar, J., and Huggenberger, P.: Influence of conceptual model uncertainty on contaminant transport forecasting in braided river aquifers, J. Hydrol., 531, 124–141, 2015.

Pirot, G., Linde, N., Mariethoz, G., and Bradford, J. H.: Probabilistic inversion with graph cuts: Application to the Boise Hydrogeophysical Research Site, Water Resour. Res., 53, 1231–1250, 2017.

Pirot, G., Huber, E., Irving, J., and Linde, N.: Reduction of conceptual model uncertainty using ground-penetrating radar profiles: Field-demonstration for a braided-river aquifer, J. Hydrol., 571, 254–264, 2019.

Pirot, G., Lindsay, M., Grose, L., and Jessell, M.: A sensitivity analysis of geological uncertainty to data and algorithmic uncertainty, AGU Fall Meeting 2020, AGU, https://agu.confex.com/agu/fm20/meetingapp.cgi/Paper/673735 (last access: 9 June 2022), 2020.

Raftery, A. E.: Bayesian model selection in structural equation models, Sage Foc. Ed., 154, 163–163, 1993.

Renard, P. and Allard, D.: Connectivity metrics for subsurface flow and transport, Adv. Water Resour., 51, 168–196, 2013.

Rényi, A.: On measures of information and entropy, in: Proceedings of the 4th Berkeley symposium on mathematics, Statistics And Probability, https://projecteuclid.org/proceedings/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fourth-Berkeley-Symposium-on-Mathematical-Statistics-and/Chapter/On-Measures-of-Entropy-and-Information/bsmsp/1200512181?tab=ChapterArticleLink (last access: 9 June 2022), 1961.

Sambridge, M.: A Parallel Tempering algorithm for probabilistic sampling and multimodal optimization, Geophys. J. Int., 196, 357–374, https://doi.org/10.1093/gji/ggt342, 2014.

Scalzo, R., Lindsay, M., Jessell, M., Pirot, G., Giraud, J., Cripps, E., and Cripps, S.: Blockworlds 0.1.0: a demonstration of anti-aliased geophysics for probabilistic inversions of implicit and kinematic geological models, Geosci. Model Dev., 15, 3641–3662, https://doi.org/10.5194/gmd-15-3641-2022, 2022.

Scheidt, C. and Caers, J.: Uncertainty quantification in reservoir performance using distances and kernel methods–application to a west africa deepwater turbidite reservoir, SPE Journal, 14, 680–692, 2009.

Scheidt, C., Jeong, C., Mukerji, T., and Caers, J.: Probabilistic falsification of prior geologic uncertainty with seismic amplitude data: Application to a turbidite reservoir case, Geophysics, 80, M89–M12, 2015.

Scheidt, C., Li, L., and Caers, J.: Quantifying uncertainty in subsurface systems, John Wiley & Sons, 304 pp., ISBN 978-1-119-32586-4, https://www.wiley.com/en-au/Quantifying+Uncertainty+in+Subsurface+Systems-p-9781119325864 (last access: 9 June 2022), 2018.

Schölkopf, B., Smola, A., and Müller, K.-R.: Kernel principal component analysis, in: International conference on artificial neural networks, Springer, 583–588, https://doi.org/10.1007/BFb0020217, 1997.

Schweizer, D., Blum, P., and Butscher, C.: Uncertainty assessment in 3-D geological models of increasing complexity, Solid Earth, 8, 515–530, https://doi.org/10.5194/se-8-515-2017, 2017.

Shimada, Y., Hirata, Y., Ikeguchi, T., and Aihara, K.: Graph distance for complex networks, Sci. Rep., 6, 34944, https://doi.org/10.1038/srep34944, 2016.

Suzuki, S. and Caers, J.: A distance-based prior model parameterization for constraining solutions of spatial inverse problems, Math. Geosci., 40, 445–469, 2008.

Thiele, S. T., Jessell, M. W., Lindsay, M., Ogarko, V., Wellmann, J. F., and Pakyuz-Charrier, E.: The topology of geology 1: Topological analysis, J. Struct. Geol., 91, 27–38, https://doi.org/10.1016/j.jsg.2016.08.009, 2016.

Tubau, I., Vázquez-Suñé, E., Carrera, J., Valhondo, C., and Criollo, R.: Quantification of groundwater recharge in urban environments, Sci. Total Environ., 592, 391–402, 2017.

Vallender, S.: Calculation of the Wasserstein distance between probability distributions on the line, Theor. Probab. Appl., 18, 784–786, 1974.

Wellmann, F. and Caumon, G.: 3-D Structural geological models: Concepts, methods, and uncertainties, Adv. Geophys., 59, 1–121, 2018.

Wellmann, J. F. and Regenauer-Lieb, K.: Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models, Tectonophysics, 526, 207–216, 2012.

Wilcox, R. R.: Introduction to robust estimation and hypothesis testing, Academic press, 690 pp., https://doi.org/10.1016/C2010-0-67044-1, 2011.

Witter, J. B., Trainor-Guitton, W. J., and Siler, D. L.: Uncertainty and risk evaluation during the exploration stage of geothermal development: A review, Geothermics, 78, 233–242, 2019.