Making the link between geological and geophysical uncertainty: geodiversity in the Ashanti Greenstone Belt

Mark D. Lindsay, Stéphane Perrouty, Mark W. Jessell, Laurent Aillères

HAL Id: insu-03620487
https://insu.hal.science/insu-03620487

Submitted on 26 Mar 2022

Making the link between geological and geophysical uncertainty: geodiversity in the Ashanti Greenstone Belt

Mark D. Lindsay,[1,2] Stéphane Perrouty,[2,3] Mark W. Jessell[2,3] and Laurent Aillères[1]

[1] *School of Geosciences, Monash University, P.O. Box* 28E, *Victoria* 3800, *Australia. E-mail: marklindsay@gmail.com*
[2] *Université de Toulouse, UPS (SVT-OMP), GET,* 14 *Av. Edouard Belin, F-31400 Toulouse, France*
[3] *Institut de Recherche pour le Développement, UR* 234, *GET,* 14 *Av. Edouard Belin, F-31400 Toulouse, France*

GJI Gravity, geodesy and tides

SUMMARY

The process of 3-D modelling forces the operator to consider data collection and processing error, while simultaneously making assumptions about geology during interpretation, to arrive at the most likely or logical geological scenario. These kinds of ambiguities lead to situations where multiple model realizations can be produced from a single input data set. Decisions are typically made during the modelling process with the aim of reducing the number of possible models, preferably to produce a single geological realization. These types of decisions involve how input data are processed and what data are included, and are always made without complete knowledge of the system under study. This regularly, if not always, results in natural geometries being misrepresented by the model, which can be attributed to uncertainty inherent in the modelling process. Uncertainty is unavoidable in geological modelling as complete knowledge of the natural system is impossible, though we use many techniques to reduce the amount introduced during the modelling process. A common technique used to reduce uncertainty is geophysical forward modelling, and the misfit between the calculated and observed response provides a means to gauge whether changes in model architecture improve or degrade the quality of the model. Unfortunately, geophysical data are ambiguous and provide a non-unique solution, with different model geometries able to produce the same geophysical response.

We propose a process whereby multiple models, collectively known as the 'model suite', are produced from a single data set that allows an exploration of geological model space. Various 'geodiversity' metrics have been developed to characterize geometrical and geophysical aspects of each model. Geodiversity measurements are combined into multivariate analysis to reveal relationships between metrics and define the boundaries of the geological possibility. A previous study using geodiversity metrics on the Gippsland Basin is extended here by including geophysical metrics. We use the Ashanti Greenstone Belt, southwestern Ghana in West Africa, as a case study to assess the usefulness of the technique. A critical assessment of the 3-D model is performed and aspects of the model space are identified that could be of interest to gold explorers.

Key words: Image processing; Spatial analysis; Gravity anomalies and Earth structure; Africa.

1 INTRODUCTION

Two-dimensional (2-D) and three-dimensional (3-D) geophysical forward modelling is a useful tool commonly used in geoscientific studies to validate or falsify geological models against the observed geophysical data. Both variations compare an 'observed' geophysical response against a 'calculated' response. The calculated response signal is measured from a representation of geology contained within a 2-D or 3-D model. The shape, depth and size of a stratigraphic unit in combination with the contrast of assigned petrophysical properties produces a signal. Other elements in the model, such as faults or dykes, are also taken into consideration. The combination of signals from modelled geological units and structures are calculated and convolved to produce the calculated field (e.g. Betts *et al.* 2003; Joly *et al.* 2008; Williams *et al.* 2009; Perrouty *et al.* 2012).

Together with petrophysical information, the geometry of the geological model plays a key role in the process of geophysical modelling. The aim of this study is to discover whether relationships exist between particular geometrical parameters exhibited in a

3-D model and the geophysical response. For example, the geophysical response may be heavily influenced by the volume or geological complexity of a particular geological formation within a model. Discovering a relationship of this nature will enable model refinements to be guided towards finding a model that honours both geological and geophysical data. However, determining a relationship is difficult as the interaction between model geometries, petrophysical properties and geophysical response is complicated. A model space exploration is a useful means by which to achieve this goal.

We produce a number of 3-D geological models based on the same input data set to produce a model suite: a collection of models that exhibit similar, but not identical, model architecture. Each one of these models is considered geologically feasible and may, *a priori*, provide the best realization of the natural world possible, given the input data. The model suite can be analysed in its entirety to determine uncertainty, the range of geometrical possibility and geophysical misfit. We perform an exploration of geological and geophysical model space using the Ashanti Greenstone Belt, southwestern Ghana as a case study. A set of 'geodiversity metrics', quantifying both geometrical and geophysical aspects of the 3-D model (Jessell *et al.* 2010; Lindsay *et al.* 2013), are used to determine geological end-members existing within the model space. A geodiverisity metric is a measure that characterizes an aspect of a model, such as volume, surface area or geophysical misfit, and allows comparison against other models within the model suite. A multivariate statistical technique is employed to simultaneously compare all metrics and determine their effect on the model suite. Analysis of results is performed using principal component analysis (PCA) which allows: (1) determination of the model space boundaries, a theoretical limit to geological possibility defined by models exhibiting the most unusual architecture; (2) identification of uncommon and common models in terms of both geometry and geophysics and (3) an understanding of which metrics can contribute most to uncertainty in modelling.

1.1 Geodiversity principles and the link to model uncertainty

Uncertainty is inherent in any modelling process and is particularly evident in 3-D geological modelling (Gershon 1998; Thore *et al.* 2002; Chugunov *et al.* 2008; Cherpeau *et al.* 2010; Jessell *et al.* 2010; Wellmann *et al.* 2010). 3-D modelling suffers from a lack of geological information as there is never complete coverage of data that describes the entire system. Outcrop is usually limited, restricting field observations and can be undersampled when available. Geophysical data may not supply complete coverage nor be available at the required resolution for detailed interpretation. Uncertainty caused by data sparseness is well documented (e.g. Rankey & Mitchell 2003; Bond *et al.* 2007; Suzuki *et al.* 2008; Polson & Curtis 2010) and is exacerbated by input data errors generated during collection, processing or preparation for input to the geological model (Yeten *et al.* 2004; Bond *et al.* 2011). Issues such as upscaling data (where clustered data points are subsampled to a representative point prior to input) are well known but typically tolerated. The effects of upscaling, sampling and data error and resulting uncertainty have been examined by Putz *et al.* (2006). They found that their model remained reasonably robust until 50 per cent of input data had been subsampled prior to model generation. After the 50 per cent threshold, the model progressively degrades with increasing degrees of downsampling until observed geology is barely recognizable. 50 per cent is not a threshold that applies to all models. The level at which models degrade beyond being recognizable depends on the redundancy between data and the interpolation algorithm, which also relates to model complexity.

Recent studies have shown how uncertainty can be located, quantified and possibly reduced using different methods (Thomson *et al.* 2005; Jessell *et al.* 2010; Viard *et al.* 2010; Lindsay *et al.* 2012; Wellmann & Regenauer-Lieb 2012). Geodiversity metrics were developed as a means of measuring uncertainty, as the presence of uncertainty implies that model geometry must be variable (Fig. 1). Varied geometry then suggests that the various elements of each model, be the volume of a granitoid or surface area of a contact, will vary as well. Geodiversity metrics are a method to analyse geometrical variations of model elements and determine the upper and lower bounds for each metric to establish model suite end-members. The volume of a modelled granitoid can serves as an example. The model with the smallest granitoid volume can be identified, and so can the model with the largest granitoid volume. The model that contains the granitoid with the smallest volume becomes the minimum end-member
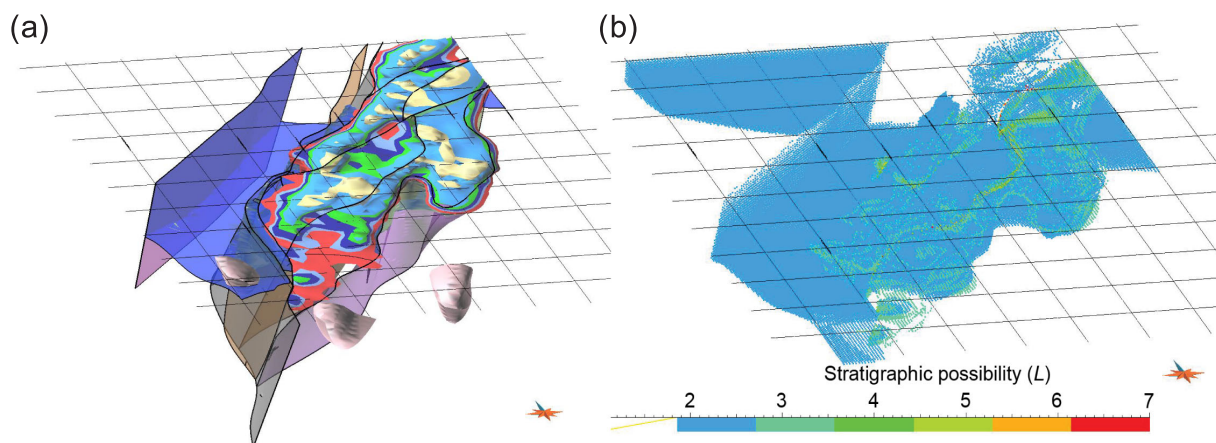
Figure 1. Visualization of uncertainty in the Ashanti Greenstone Belt model. (a) Geological 3-D surface model displaying contacts (colours correspond to the stratigraphic column in Fig. 2) and fault surfaces (thick black borders) from the initial model. (b) Uncertainty within the model displaying stratigraphic possibility colour coded according to uncertainty (blue is lowest uncertainty, red is highest). The stratigraphic possibility (L) is the number of possible stratigraphic units that can be found at a given point for all the models in the model suite (Lindsay *et al.* 2012). The grid lines are located at the surface and spaced at 10 km.

representative for the volume of the granitoid, and the model exhibiting the largest volume becomes the corresponding maximum end-member representative.

Geometrical aspects of a 3-D model are not the only metrics of interest to geoscientists. Geophysics is an integral component of many modern geological studies and is used heavily in the 3-D modelling process. Geophysical metrics can be included in the suite of geodiversity metrics to quantify the capacity of the model suite to match the observed geophysical field. End-members for different geophysical metrics can be identified just like their geometrical counterparts and used to aid further modelling efforts. Finding whether any geometrical metrics can be linked with geo-

physical metrics is of particular interest, providing a guide to which geometrical aspect of the model should be examined to most efficiently decrease geophysical misfit and lead to a model honouring both geophysical and geological data.

2 GEOLOGICAL REVIEW—ASHANTI GREENSTONE BELT, SOUTHWESTERN GHANA

The Leo-Man Craton forms the southern Archaean/ Palaeoproterozoic section of the West African Craton (Fig. 2). Four Palaeoproterozoic greenstone-granitoid belts can be found
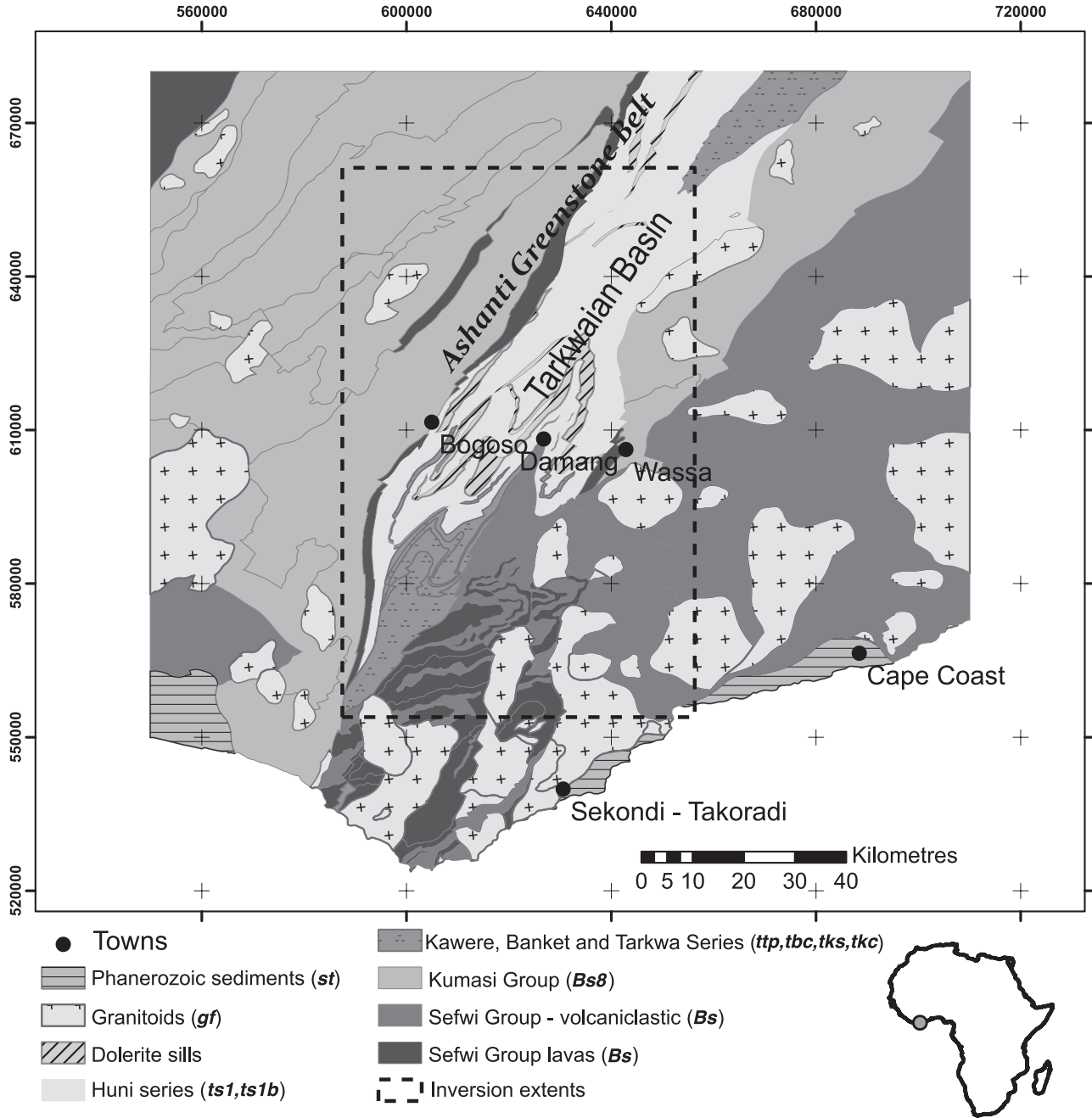

Figure 2. Geological map of Ashanti Greenstone Belt, southwestern Ghana. Note the location of the modelled region over the Ashanti Greenstone Belt and the Tarkwaian Basin. The units codes correspond to the stratigraphic column used in modelling (Fig. 3a).

in the south of the craton (from east to west); the Kibi–Winneba, the Ashanti, the Sefwi and the Bui. Each belt is separated by a sedimentary basin (also from east to west); the Akyem (or Cape Coast Basin), Kumasi and Sunyani basins. The Ashanti Greenstone Belt has economic significance as it hosts a number of large and world-class gold deposits, including Obuasi, Tarkwa, Bogoso/Prestia and Damang (Allibone *et al.* 2002; Pigios *et al.* 2003; Tunks *et al.* 2004; Feybesse *et al.* 2006).

2.1 Mineralization and gold prospectivity

Gold mineralization has been identified either being spatially associated with the Ashanti Fault and other major shear zones or along the contact between the Tarkwaian and Birimian units. Tarkwaian-hosted deposits display two styles of mineralization (Perrouty *et al.* 2012). The first is observed only in quartz lithic conglomerates within the Banket Formation (e.g. Tarkwa Mine). Economically viable deposits are constrained to a few horizons locally named 'Banket Reefs' (Blenkinsop *et al.* 1994) and is thought to be of palaeoplacer origin (Sestini 1973; Hirdes & Nunoo 1994). A second mineralization style, observed within Tarkwaian units, are the hydrothermal deposits that occur along the Birimian/Tarkwaian contact (e.g. Damang Mine). A contrasting mesothermal mineralization style is associated with the Birimian Supergroup and is associated within quartz ± carbonate veins within graphitic-mylonitic shear zones. World-class gold mines such as the Ashanti deposit are hosted within the Obuasi/Main Reef fissure (Allibone *et al.* 2002; Tunks *et al.* 2004).

3 METHOD

3.1 Modelling the Ashanti Greenstone Belt

The purpose of building the Ashanti Greenstone Belt model was to determine the geometry of the Tarkwaian Basin. The depth and morphology of the basin base is of particular economic interest as it plays host to existing and potential placer gold deposits. Existing understanding of Tarkwaian Basin depth is controversial. Hastings (1982) and Barritt & Kuma (1998) predict the basin is between 1500 and 2500 m thick, but these estimates assume an older version of stratigraphy that has low-density Birimian metasediments underlying the higher density Birmian metavolcanics. Depth estimates based on gravity inversion and interpretation may therefore underestimate basin depth and granitoid geometry (Perrouty 2012). Current stratigraphic relationships developed through updated information (Adadey *et al.* 2009) has encouraged construction of this model. The new map proposed by Perrouty *et al.* (2012), structural measurements, stratigraphy (Fig. 3a) and geophysical interpretation have been input to create the 3-D model of the Ashanti Greenstone Belt, southwestern Ghana.

3.2 Data sets obtained from fieldwork

Pre-existing field observations and outcrop maps from Loh *et al.* (1999), BHP Billiton and Golden Star are used in combination with data collected by Perrouty *et al.* (2012). This field data include structural observations and petrophysical measurements calculated from rock samples (Metelka *et al.* 2011; Perrouty *et al.* 2012). Geophysical interpretation was constrained using petrophysical data, and field observations aided geophysical interpretation. A thick lateritic and/or saprolitic layer covers most accessible areas within

Figure 3. Input parameters used in construction and geophysical modelling of the Ashanti Greenstone Belt model. (a) Stratigraphic column based on recently revised tectonic evolution (Perrouty *et al.* 2012). (b) Petrophysical properties and distribution statistics assigned to each formation (Metelka *et al.* 2011; Perrouty *et al.* 2012).

the region that restricts the occurrence of Proterozoic outcrop. Geophysical interpretation was therefore necessary to gain geological understanding between outcrop areas, resulting in much of the region requiring interpretation.

Gravity data have been used to provide a potential field data set to cross-validate magnetic interpretation and better image deeper structures in the region. A number of data sets has been compiled to create the gravity data used in the validation of this model. A pre-processed Free Air anomaly grid data set was obtained through the International Gravimetric Bureau (BGI, http://bgi.omp.obs-mip.fr/) and contains a combination of BGI on- and offshore data, satellite data and Getech ground data (African Gravity Project 1986–1988, http://www.getech.com/history.htm). Spatial resolution is 2.5 arc min, or close to 4.6 km pixel^{-1}. Fig. 4 shows a steep gravity gradient to the west of the Tarkwaian Basin that marks the location of the Ashanti Fault. Other less dramatic anomalies represent Eburnean granites and Eoburnean granitoids.

3.3 3-D modelling and model suite creation

3-D Geomodeller (http://www.geomodeller.com/geo/index.php) was used to integrate field, geophysical and satellite data into a coherent and geologically feasible model. 3-D Geomodeller uses an implicit method to integrate all input data to create scalar potential field of conformable lithological formations (Lajaunie *et al.* 1997). Note that the 'scalar potential field' is different to the geophysical potential field used to generate some of the input for the model. Geological interfaces are interpolated via cokriging of values of the scalar field (located at interfaces) and its partial derivatives dv/dx, dv/dy, dv/dz (located at orientation data). Three types of input data are required for the 3-D Geomodeller implicit method to function: (1) geological contact locations; (2) geological orientation
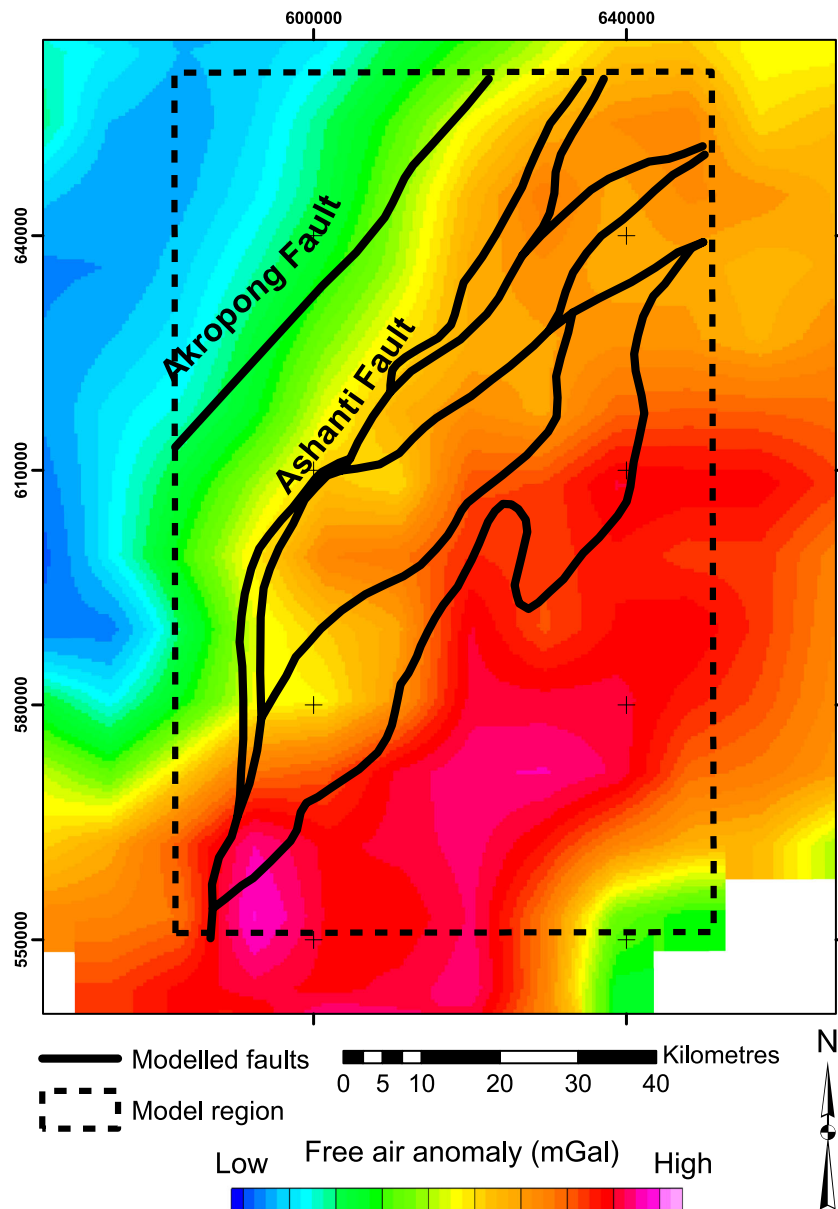
Figure 4. Gravity response of the Ashtanti Greenstone Belt with an overlay showing the location of modelled faults.

measurements and (3) a stratigraphic column with defined geological relationships. The contact locations define where interfaces exist within the model, the orientation measurements give interfaces their geometry and the stratigraphic column defines adjacent geological relationships. Fault relationships can also be defined, allowing complex fault networks and timing relationships between fault and geological units. A complete description of the 3-D Geomodeller method and associated techniques is described by Calcagno *et al.* (2008).

The first model created is considered to be the 'initial model', a model typically created in a normal 3-D geological modelling workflow. It represents the best efforts of the geoscientist to produce a consistent model that attempts to honour all input data. Lindsay *et al.* (2012) describes how the initial model is subjected to uncertainty simulation to create the model suite and is then included in the model suite as a member no more or less likely to exist in nature than the other perturbed models. Uncertainty simulation consists of taking the orientation measurements (including those assigned

to faults) from the input data set and reassigning their values to within $\pm5°$ (both strike and dip) of the original measurement. The reassignment is performed as a pseudo-random Monte Carlo simulation to avoid bias. For example, a measurement of 325/40E could be perturbed so that measurement within model '1' of the model suite would be reassigned 323/35E, model '2' reassigned 320/42E and so on. Each model is then recalculated using the new measurements to create the model suite: a set of geometrically similar, but diverse examples of geological possibility. Voxets used in the following geodiversity analysis were generated by sampling each member of the model suite every 500 m on the x- and y-axes (northings and eastings, respectively) and 200 m on the z-axis (depth). 100 models were produced through uncertainty simulation to generate a model suite totalling 101 members, including the initial model. 101 models are considered to comprise a reasonable sample of model space from which to test the geodiversity techniques. Each member of the model suite was included in end-member and PCA.

Table 1. Summary of geodiversity metrics and their function. Detailed descriptions of the geometrical metrics and their function can be found in Lindsay *et al.* (2013).

Name	Subject	Measurement	Output
Geometrical geodiversity metrics			
Volume	Voxet	Metres^3	Volume for each formation
Depth	Voxet	Metres	Shallowest and deepest occurrence of each formation
Curvature	Surface	k_m: Mean curvature k_g: Gaussian curvature	Average k_m and k_g values for each formation
Contact relationships	Surface	Area (metres^2)	Contact surface area and contact relationships
Geological complexity	Voxet	Number of different lithologies around point-of-interest	Scalar value representing geological complexity
Geophysical geodiversity metrics			
Root mean square (rms)	Residual grid	Global measure of geophysical misfit	Scalar value
Standard deviation	Calculated grid	Global measure of geophysical variability	Scalar value
Entropy	Residual grid	Global measure of geophysical variability	Scalar value
2-D correlation coefficient	Comparison between observed and calculated grids	Global measure of geophysical covariance—recognizes similar patterns	Scalar value
Hausdorff distance	Distance between observed and calculated grids	Global measure of geophysical misfit—accounts for pattern translation, rotation and dilation	Scalar value

3.4 Geometrical geodiversity metrics

A collection of geodiversity metrics have been employed to categorize the Ashanti Greenstone Belt model (Table 1). A short review of the geometrical metric method (volume, depth, surface area, curvature and complexity) is provided in this section, though additional details can be found in Lindsay *et al.* (2013). Being able to quantify a particular aspect of a model allows comparison with other models in the model suite. Comparison then allows identification of end-member representatives for each metric. Quantification also allows each metric to be compared against one another to determine whether there are metrics that can best explain uncertainty contained within the model suite. All the geophysical metrics are used to provide a global measure of the calculated geophysical response or misfit with an objective representation of nature in the observed response. The geophysical metrics have been included in the geodiversity collection to discover if the geophysical response can be associated with any of the geometric geodiversity metrics according to the stated hypothesis.

The following geometrical geodiversity metrics were developed to analyse the geometry of 3-D geological model elements. Note this list of metrics is not exhaustive. Many other metrics may exist that are also effective at measuring the effects of uncertainty on model geometry.

3.4.1 Formation depth and volume

The deepest and shallowest extents of each stratigraphic unit are determined using this metric. Each model in the suite can be analysed to find if any model shows that a stratigraphic unit is significantly deeper or shallower than it is in others. Information like this can be useful to assist in identifying under what conditions a stratigraphic unit may be shallower. Most units in a model that displays either flat-lying geology, low uncertainty or both will share the same deepest and shallowest extents with other models within the model suite which makes identification of end-members difficult. Therefore, the unit that has the most volume at the depth extent under study is considered the end-member. For example, the deepest extent of 'Unit A' is found to be 8500 m and was found in models 3, 6 and 70. In Model

3, Unit A has 3000 m^3 at 8500 m depth, Model 6 has 4000 m^3 and Model 70 has 3500 m^3 at 8500 m depth. Model 6 is determined to be the end-member for the deepest extent of Unit A, as it has the most volume at 8500 m depth. Volume of each unit is calculated by counting the voxels assigned to that unit and multiplying by the voxel volume.

3.4.2 Average mean curvature

Most surfaces within a 3-D geological model are curved as they attempt to represent the natural world. Defining the curvature of a surface can be useful, especially if curvature is conducive to producing an economically viable target. Antiformal traps are highly sought after in oil, gas and minerals exploration and determining where they exist can aid exploration activities. Curvature may also influence the geophysical response of a model. Potential links between a particular style or degree of curvature and high or low geophysical misfit can aid model refinement and improve the modelling workflows.

We use a technique described by Lisle & Robinson (1995) and Lisle & Toimil (2007) that rotates a surface around its normal until the maximum curvature ('k_1') is found. The surface that is perpendicular to k_1, 'k_2' is also recorded. k_1 and k_2 are known as the principal curvatures. The sign of principal curvatures indicates their polarity, negative indicates concave-upwards, positive indicates convex-upwards. Mean curvature (M) is calculated from the arithmetic mean of k_1 and k_2:

$$M = \frac{k_1 + k_2}{2}. \tag{1}$$

$M < 0$ represents a concave (synformal) surface, $M > 0$ represents a convex (antiformal) surface and $M = 0$ either represents a flat plane or a 'perfect saddle' (Lisle & Toimil 2007). The product of k_1 and k_2 is the Gaussian curvature:

$$G = k_1 \cdot k_2, \tag{2}$$

which can be used to identify specific folding interference patterns. Positive G values show that both principal curvatures k_1 and k_2 have the same sign, and represent a dome or basin. Negative G values

Table 2. Voxel count and contact relationships between stratigraphic units generated in the initial model of the Ashanti model suite. Stratigraphic units are labelled in the column and row headings. A voxel count of '0' indicates these units are not in contact within this model volume.

	Bs	*bs*8	*tkc*	*tks*	*tbc*	*ttp*	*ts*1*b*	*ts*1	*gf*
Bs	0								
*bs*8	23 885	0							
tkc	1708	5910	0						
tks	123	2127	2418	0					
tbc	11	880	2220	2083	0				
ttp	1	99	2194	1249	1667	0			
*ts*1*b*	1	59	870	1289	1922	1095	0		
*ts*1	0	12	27	19	32	32	668	0	
gf	1025	134	0	0	0	0	0	0	0

indicate that principal curvatures have different signs, and represent antiformal or synformal saddles (Mallet 2002; Gray *et al.* 2006; Mynatt *et al.* 2007). Use of curvature calculations as a geodiversity metric allows comparison of specific geological interfaces with a model and model suite using quantitative methods.

3.4.3 Contact relationships

A contact relationship metric has been developed to quantify the surface area of modelled stratigraphic units, and to identify which units are in contact with other units within the model. Contact relationships are recorded by determining which stratigraphic units are adjacent to each other and surface area is calculated from a voxel count of the regions where that adjacency exists. Table 2 shows an example of the relationships between stratigraphic units found in the initial model of the Ashanti model suite. This matrix provides a useful guide to geological relationships that exist in the model that may not be immediately evident without thorough visual investigation. Information of this kind can provide a useful means to cross-validate the resulting model with the contact relationships that are described by the input geological data.

3.5 Geophysical forward modelling

The geophysical forward modelling method used in this contribution discretizes the 3-D geological model into a 3-D grid, or 'voxet'. Each cell, or 'voxel', is assigned a stratigraphic unit identifier based on the geological model. Next, each voxel is assigned a petrophysical value representative of the assigned stratigraphic identifier and appropriate to the geophysical response required. In this paper, the gravitational response is being modelled, so density is the assigned petrophysical property. The sum of the contribution of each voxel to the field is calculated using an analytical expression (Plouff 1976; Okabe 1979; Holstein *et al.* 1999; Holstein 2003). The heterogeneity of petrophysical properties predicted to occur within a stratigraphic unit can be simulated by the definition of probability distribution functions (Fig. 3b). A 3-D model of the southern Ashanti Greenstone Belt and corresponding calculated gravitational response is shown in Fig. 5.

3.6 Comparison of geophysical images

The following techniques have been included in this study to obtain a scalar value for each calculated geophysical response. Some of

these values are obtained by calculating the residual between the observed and calculated response and include the 2-D correlation coefficient, root-mean-squared (rms) and the Hausdorff distance. A residual grid is created by subtracting the calculated response grid from the observed response grid. The measure of geophysical misfit is often expressed as a root-mean-squared value, or 'rms',

$$x_{\text{rms}} = \sqrt{\frac{1}{n}\left(x_1^2 + x_2^2 + \cdots + x_n^2\right)}, \tag{3}$$

where $x_1, x_2 \ldots x_n$ equals the difference between the observed and computed signals at a particular location. Fig. 6 shows the observed, calculated and residual grids for the initial model, that is, the model that was calculated using an unperturbed data set. The standard deviation and entropy techniques are other typical image analysis techniques. The standard deviation technique was performed on the calculated grids, the entropy technique was performed on the residual grids. The scalar value obtained through these techniques is global, in other words, a value that represents the entire grid, not local regions within the grid. The global approach was implemented to adhere to the requirements of PCA. A requirement for performing PCA is that a single value represents each geodiversity metric (geometrical and geophysical) for each model in the model suite. While useful, non-global or local image analysis of each of the geophysical grids would produce multiple values for each metric. All values obtained for one model would need to be reduced to single value for use in PCA, making a local analysis averaged.

3.6.1 Standard deviation

The standard deviation of an image is taken as measure of grid value spread

$$s = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{\frac{1}{2}}. \tag{4}$$

The standard deviation of an image is a common technique in image analysis as it represents the underlying intensity probability distribution, and it can be used to measure the degree to which potential field grid values vary across the entire image. Using standard deviation of the calculated response as a geodiversity metric allows the variability of each grid to be compared. We can ask ourselves why would one model have a higher standard deviation than another? Can this be resolved due to a particular geometrical feature of the model? For example, a granitoid with a high petrophysical contrast to the country rock that is larger in a model than others,
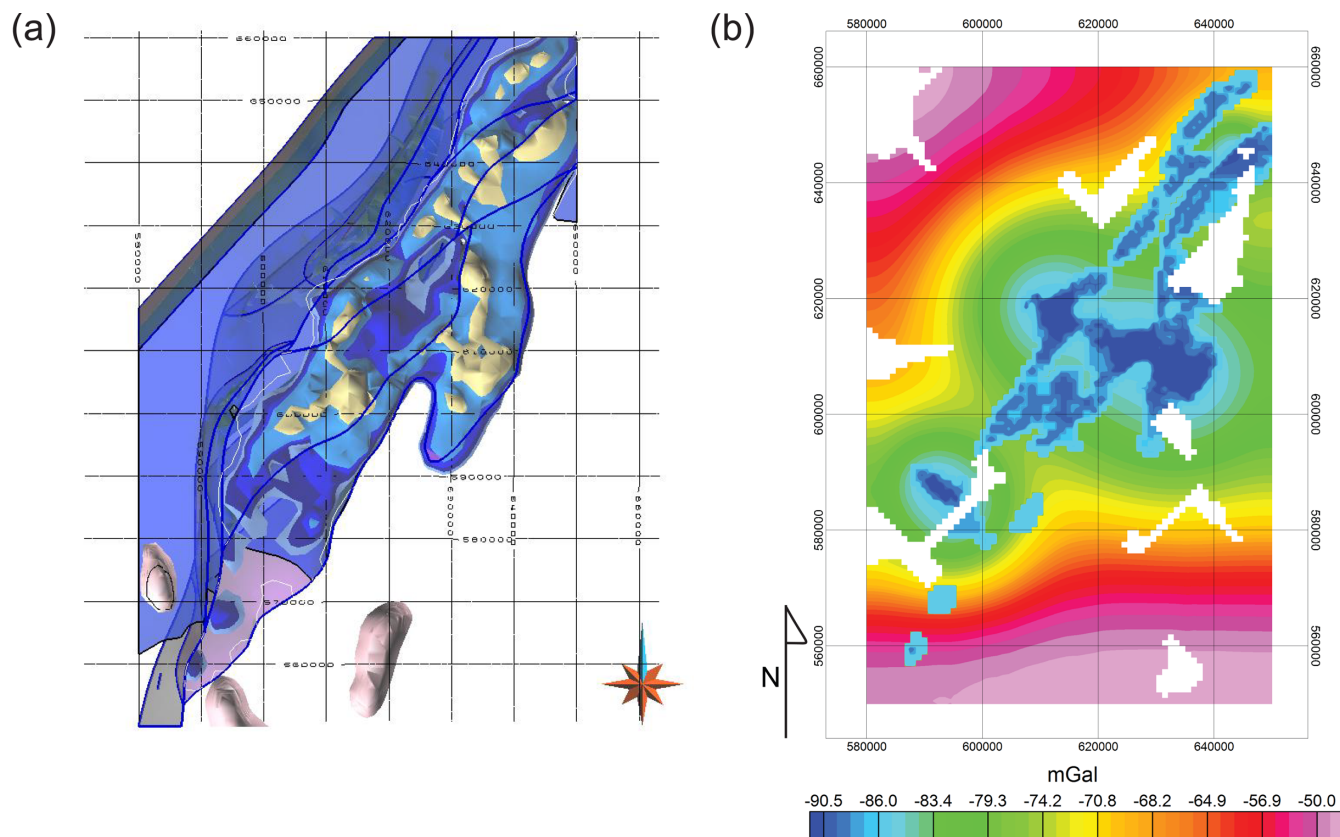
Figure 5. Calculating the geophysical response from a 3-D model. (a) Map view of the Ashanti Greenstone Belt 3-D model. Faults have been shown with blue borders and stratigraphy is shown with opaque surfaces. (b) The calculated geophysical gravity response of (a). Note the 'white spaces' in (b) indicating there is no interpolated gravity response. The non-interpolated regions are the result of undersampled surveying, which can be due to a multitude of reasons including lack of land access (geographic and political), funding shortfalls or poor survey design.
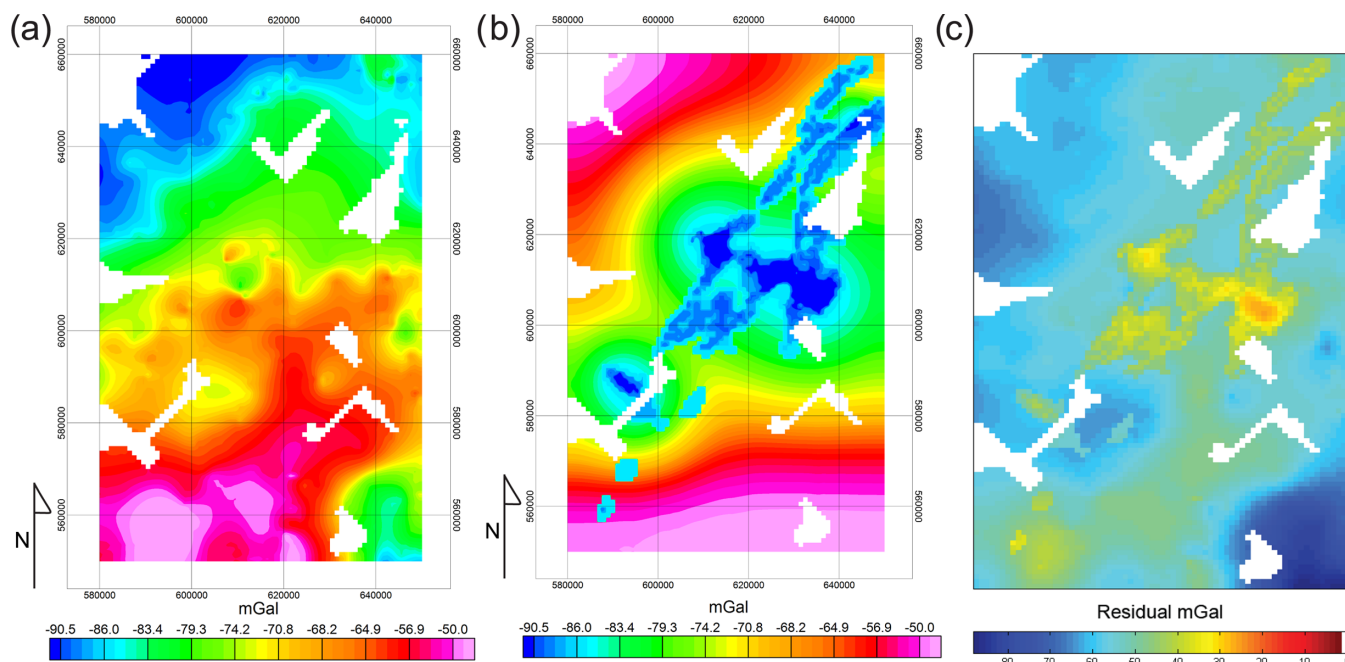

Figure 6. Example of 3-D geophysical forward modelling from the initial Ashanti Greenstone Belt model. (a) The observed gravity grid, which is based on survey data, (b) is the calculated response of the initial model and (c) is the residual, calculated from the difference between (a) and (b) with an rms = 8.4043. 20 km grid line spacing for (a) and (b).
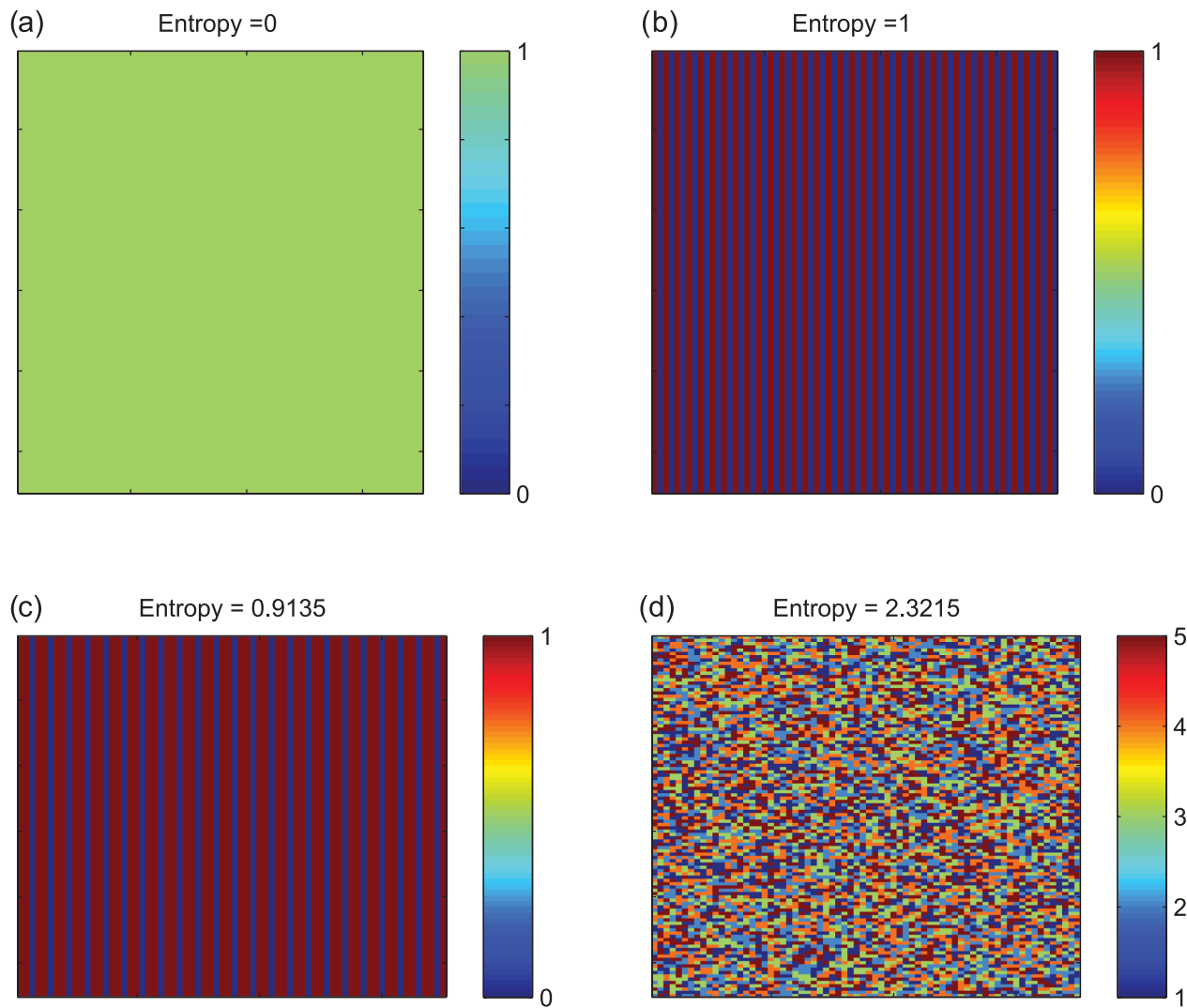
Figure 7. A set of four images showing how information entropy (E) can identify information content in an image. (a) Shows $E = 0$, meaning that entire image contains the same value and all pixels in the images can be predicted to 100 per cent accuracy based on the value of one pixel. (b) An image with two integer possibilities (0 or 1), alternating bands across the image will result in the highest amount of E for a binary system. While instinctively one would assume that the regular bands are easy to predict, in fact as there are equal proportions of 0s and 1s throughout the image, the possibility of picking a 0 or 1 at any given point is equally likely (50 per cent for each value). There is no single value that is more likely to be found, so $E = 1$, the highest value for a binary image. (c) A binary image with a 66 per cent chance to find '1' and 33 per cent chance for '0', E is slightly lower, reflecting less randomness in the image. (d) Shows an image generated using a random function of integers 1–5. E is relatively high as the number of possibilities has increased, and the relative proportions of each integer are similar.

may result in a relatively high standard deviation for the calculated response and assist examination of input petrophysical constraints.

3.6.2 Entropy

Entropy (E) is used to measure the average bits per pixel over an entire image, representing its global information content (O'Gorman *et al.* 2008). The type of entropy used here is 'Information Entropy' which is derived from the Shannon Entropy model (Shannon 1948):

$$E = -\sum_{i}^{N} p_i \log p_i, \tag{5}$$

where E is the sum of all products of p (probability) of each possible outcome (i) out of N total possible outcomes. $E = 0$ indicates that the image is dominated by large regions of the same value (Fig. 7a). In a system of two integer values, 0 and 1 (1-bit system) the $E_{max} = 1$

(Fig. 7b). The image is made of equal proportions of possible values in this case, therefore $E = 1$ reflects that at any point it is equally likely to find a '0' or a '1'. Fig. 7(c) is still a 1-bit system (values 0 or 1), but two of three lines along the y-axis are '1'. The proportions are no longer equal as it is 66.66 per cent likely to find '1' and 33.33 per cent likely to detect '0' at any given point and $E = 0.9135$. E_{max} will increase with the range of values within the data set under study. This is shown in Fig. 7(d) where a range of five integers (one to five) are found in the randomly generated image.

Information entropy is a useful tool in exploring model space uncertainties. Wellmann & Regenauer-Lieb (2012) used Information Entropy as a visualization technique to communicate where uncertainties within 3-D models exist. We use the concept to analyse the residual grids produced by the automatic forward modelling technique to find grids that are smoother and contain less variability in values (Gonzalez *et al.* 2003).

3.6.3 2-D correlation coefficient

2-D correlation coefficients are typically calculated in geophysical and engineering applications to track changes in 2-D and 3-D objects. The subject of 2-D correlations are often images as the algorithm is able to measure how closely an image of an object subjected to deformation resembles the original state over a time period. The correlation is '2-D' as it is performed between matrices, rather than between two vectors (Fig. 8). The 2-D correlation r is calculated using

$$r = \frac{\sum_m \sum_n \left(A_{mn} - \bar{A}\right)\left(B_{mn} - \bar{B}\right)}{\sqrt{\left(\sum_m \sum_n \left(A_{mn} - \bar{A}\right)^2\right)\left(\sum_m \sum_n \left(B_{mn} - \bar{B}\right)^2\right)}}, \tag{6}$$

where $\bar{A}$ is the global mean of image one (observed geophysical response) and $\bar{B}$ is the global mean of image two (the calculated geophysical response). The correlation between the images is not a subtractive comparison, such as performed in the rms method, rather it measures whether patterns in the image resemble each other. The purpose of using this technique to compare observed and calculated
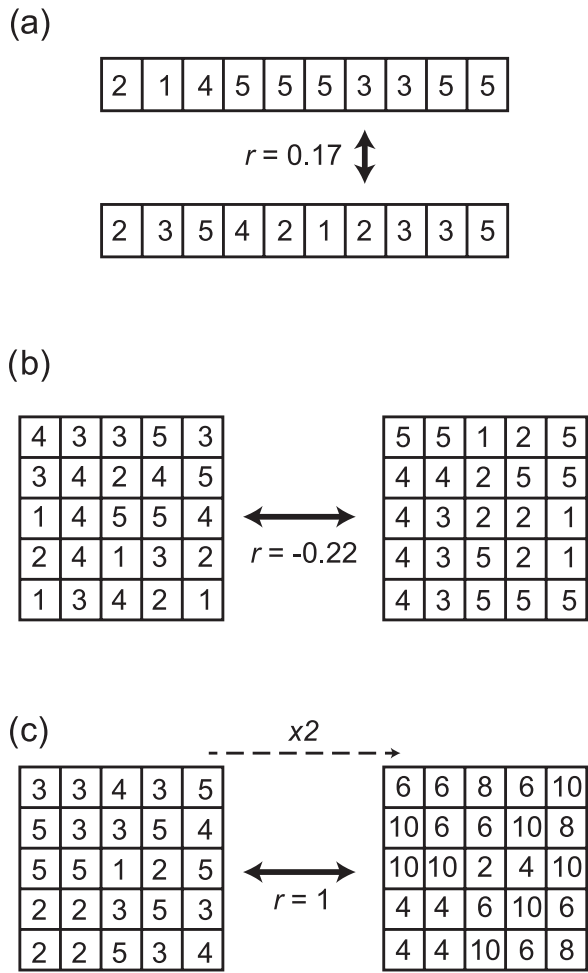

Figure 8. Features of 1-D and 2-D correlation functions. (a) Correlation between two sets of vector data showing little to no linear dependence. (b) 2-D correlation between two matrices showing little linear dependence. (c) The correlation function recognizes that the second matrix has been created simply by multiplying the first matrix by two, returning a score indicating complete linear dependence.

responses is that if the correlation coefficient is high, the spatial variation of values in both the calculated and observed responses is similar.

3.6.4 Hausdorff distance

The Hausdorff distance measures how far points in two different subsets are from each other. The distance can then be used to understand the level of resemblance two superimposed objects have to each other. The Hausdorff distance has been typically used in machine vision (Rucklidge 1997; Wang & Suter 2007) and pattern recognition applications (Olson & Huttenlocher 1997; Sim *et al.* 1999; Gao & Leung 2002) to compare and find patterns in one image that may be present in another (Huttenlocher *et al.* 1993). Geophysical grids can be compared using the Hausdorff distance. If we assume that the geophysical forward response of a given 3-D model is a subset of the model space, then the observed response of the potential field is also a subset of the same model space, but represents an as yet undiscovered 3-D model (Foudil-Bey 2012).

The Hausdorff distance (d_{H}) can account for dilation and limited degrees of rotation and translation of one image with respect to the other (Fig. 9). Geometrical differences between model suite members will be reflected in their respective calculated geophysical responses. The recognition of similar patterns is not performed via standard geophysical misfit algorithms, so it is interesting to determine whether using the Hausdorff distance as a metric for model comparison can be more effective.

The Hausdorff distance is calculated using the following equation. X and Y are two non-empty subsets of a metric space (M, d). The Hausdorff distance between these two sets d_{H} (X,Y) is

$$d_{\mathrm{H}}(X, Y) = \max\left\{\sup_{x \in X} \inf_{y \in Y} d(x, y),\ \sup_{y \in Y} \inf_{x \in X} d(x, y)\right\}, \tag{7}$$

where 'sup' is the supremum, 'inf' is the infimum and x and y are points within sets X and Y, respectively. In the application of the Hausdorff distance used here, X is a grid of the observed response, Y is a calculated forward response grid of some model and x and y are values of a given cell within the grid. The supremum is defined as the least element of subset Y of set X that is greater than or equal to all elements of Y. The infimum is defined as the greatest element of subset Y of set X that is less than or equal to all elements of Y (Fig. 10). In other words, the infimum defines the lower bounds of subset Y within set X, whereas the supremum defines the upper bounds of subset Y within set X. Therefore, the Hausdorff distance finds point x from set X that is farthest from any point in Y and measures the distance from x to the nearest neighbour in Y.

3.7 Using 3-D geophysical forward modelling in combination with geodiversity

The geophysical response of a model will always rely on the 3-D geological architecture it represents. Integration of geophysical misfit and geodiversity metrics allows data exploration to identify which, if any, geodiversity metrics may influence geophysical misfit. Finding a single metric, or combination of metrics, that influence geophysical misfit can help modellers refine their models to reduce geophysical misfit between the observed response and the calculated response.

Caution must be taken with this research direction. It is not intended that any relationship discovered between a geodiversity metric and geophysical response be exploited to manipulate the misfit
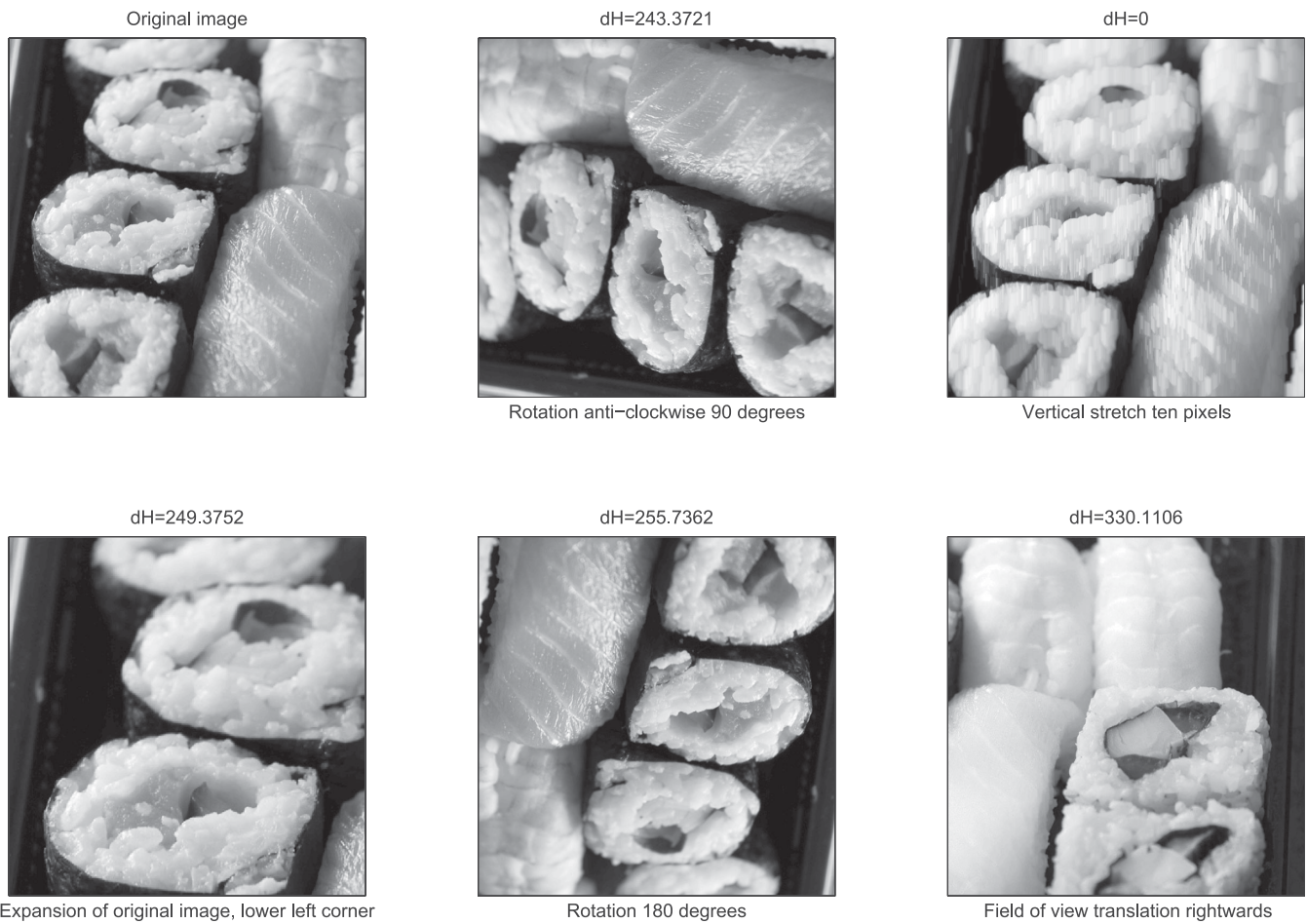
Figure 9. Hausdorff distances (d_H) calculated between an original image (top-left panel) and images subjected to different morphological operations. Note that the stretch example produces zero Hausdorff distance.
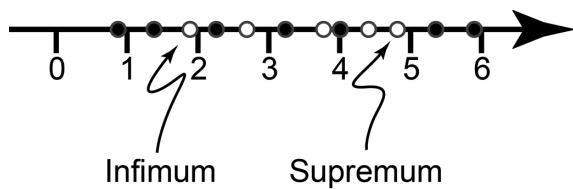

Figure 10. A number line shows a set of numbers (white and black balls) and a subset (white balls). The infimum and supremum of the subset are indicated.

into a more 'agreeable' result. We emphasize the importance of geological feasibility with respect to the geophysical misfit. For example, if a covariant relationship between the volume of a particular unit and geophysical misfit is found, the volume should not be adjusted in isolation simply to decrease the misfit. Changes to the volume of that unit should be considered with respect to geology, so that unrealistic realizations of the geology are avoided. The intent is to guide the modeller towards finding alternative data sources that better resolve the geometry of the geological formation through additional data (Lindsay *et al.* 2012). Adding data is not necessarily going to improve the geophysical misfit, especially if the data are inappropriate to improve the realization of the anomalous modelled geology. For example, if volume was found to be linked to geophysical misfit of the model suite, adding 3-D seismic interpretation to

the input data set may improve the misfit of the model suite and provide a more accurate and reasonable geological realization.

3.8 Data analysis using PCA

Each model can be analysed and then ranked using geodiversity metric results. The models exhibiting the greatest and smallest (i.e. the end-member representatives) volume of a particular geological unit are easily identified, as are the volume ranks of all other models. This information can be very useful, especially for further processing and modeling, such as geophysical inversion. However, potentially more interesting is discovering where a particular model sits in an overall ranking scheme that incorporates all geodiversity metrics. This type of analysis defines where a model resides in relation to other models within model space. For example, finding whether the initial model (the model calculated from non-perturbed data) exhibits typical characteristics is important. The initial model is normally the only model considered in current modelling workflows. If this model does not exhibit typical characteristics, the practice of producing one single model should be seriously questioned.

Simultaneous analysis of results from techniques that measure both geometrical and geophysical phenomena require a specialized set of tools. Typical methods measuring correlation, such as Spearman's Rank or Pearson's *r*, operate pairwise and only two variables

are measured simultaneously. To adequately understand the relationship between different geodiversity metrics, we require a method that can compare different observations, sometimes with different scales of measurement. Furthermore, we want to be able to reduce the number of geodiversity metrics to a select few that best represent the degree of variability observed through the model suite. PCA falls within the 'feature transformation' group of methods that fit these requirements, while also providing visualization techniques that assist in understanding the interaction between variables and model space definition. Formally, PCA is an exploratory data technique that analyses the interaction between geodiversity metrics through orthogonal data transformation, where data are then reorganized in terms of relevance to model suite variability (Jolliffe 2002). PCA identifies which geodiversity metric contributes the most towards model suite variability and defines the model space. 'Outlier' models form the boundary of model space and exhibit geometrical and geophysical characteristics that show the biggest combined difference when compared to other models. 'Barycentre' models form the centre of model space and display characteristics that are similar when considered in combination. Defining the model space is an important exercise as it characterizes what is possible geologically given the modelling method, geodiversity metrics and input data set. Knowledge of the model space parameters assists definition of geological possibility, given the input data set, method of model calculation and geodiversity metrics employed. We performed the PCA in Matlab with the 'princomp' function (http://www.mathworks.com.au/help/toolbox/stats/princomp.html). The multivariate distance of the model from the centre of the data set (i.e. the model's rank within model space) is determined using Hotelling's T^2 statistic, also obtained from the 'princomp' function (Hotelling 1931; Krzanowski 1995).

The PCA is carried out in two stages. The first stage involves analysing the contact relationship, volume, depth and complexity metrics individually. Each metric measures the corresponding observations for each stratigraphic unit within the model suite. PCA is performed on the stratigraphic units to find which units contribute the most to model variability for that metric. For example, if the volume of hypothetical units *XYX* and *ABA* were found to contribute the most to model suite variability, *XYX* and *ABA* are then representatives of the volume metric. They would then be used with other representative units in the second stage 'combined' PCA. A detailed account of the PCA procedure is described in Lindsay *et al.* (2013).

The remaining metrics, including all the geophysical varieties, do not require filtering in a first-stage PCA procedure as a single value representing the entire model (e.g. the rms or Hausdorff distance) is produced. The curvature metric was not subjected to PCA as curvature was not calculated for every stratigraphic unit through the model suite. The computation time required for the curvature procedure is high, so target geological contacts were chosen manually. The decision for which contact should be analysed for curvature was based on which exhibited the highest surface area, and therefore was most likely to vary and have a resulting effect on model geometry. This decision was made possible due to information provided by the contact relationship metric (e.g. Table 2).

4 RESULTS AND DISCUSSION

Results are presented in two parts: (1) as a pure end-member analysis, where end-members for each representative metric (determined in the first-stage PCA where necessary) are presented with the cor-

responding measurement and (2) the results of the combined PCA, where the combined analysis of geodiversity metrics is presented with a depiction of model space.

4.1 End-member analysis

Results from geodiversity analysis are shown in Table 3. It becomes evident that knowledge of these end-members and their corresponding measurement reveals several interesting aspects of the model suite. First, the geometrical metrics display considerable variation in the range between end-member values. The range of possible volumes for the Late Birimian formation ('*Bs8*') is 1553.1 km^3, whereas the depth metrics only show a range of 200 m between end-member representatives. Considering that the smallest interval between the measured depth of a given formation is 200 m (voxel size is 200 m on the Z-axis), the results are essentially binary in that the stratigraphic unit is either one of two depths (1400 or 1200 m for *Bs8* and 400 or 200 m for the base of the Tarkwaian Series '*tks*'). The depth for the shallowest extent for *Bs8* does allow for three possibilities (1000, 1200 and 1400 m), but none of these results allow for much fidelity in terms of detecting different depths. In contrast, the volume is calculated by counting the voxels within a formation, and the range between the smallest and largest volume allows a far greater spectrum of results (between 321 915 to 352 977 voxels, a range of 31 062 voxels) than the depth metric. Much smaller variation in geometry can be detected in the volume of a formation throughout the model suite, suggesting that the volume metric will be more useful as a variable in PCA. Decreasing the voxel Z-axis interval may improve the effectiveness of the depth metric, but the impact on computing requirements (both storage and computation) would need to be considered.

The complexity metric shows that the *tkc* stratigraphic unit has an average of between 3.1141 and 3.3186 different lithologies at any given point. This result is high in comparison to the 1.3825 to 1.9825 calculated for the Early Birimian ('*Bs*'). A relatively high degree of complexity that also describes model suite geometrical variability (determined during first stage of PCA) suggests that *tkc* is an important unit within the model suite. The location of *tkc* in the stratigraphy makes the complexity result even more interesting as *tkc* forms the base to the Tarkwaian Basin and is gold prospective. Intuitively, the average of over three different units at any given point seems high and indicates that this unit may define the geometry of units overlying it.

The relatively large range of values seen within the contact relationship metrics *Bs* and *Bs8* (5116 voxels) and between *Bs* and *tkc* (1869 voxels) reveal how geometrically variable the model suite is. The large range indicates that the surface area between these contacts can be easily related to model uncertainty. The *Bs/Bs8* contact variability shows that the contact between the Early and Late Birimian units is inadequately constrained. This is confirmed as only a single orientation point defines the orientation of this contact. The *Bs8/tkc* contact is gold prospective, so a large range of values may suggest that this model suite may not offer enough certainty for use with further prospectivity modelling. These results do not necessarily mean that the model suite cannot be used, but that the variability of this geometrical aspect of the model suite should be kept in mind in future applications.

The curvature results reveal some aspects of the *Bs/Bs8* and *Bs8/tkc* contacts that are difficult (if not impossible) to determine from visual analysis of the surfaces. The average mean curvature (k_m) of the gold prospective *Bs8/tks* contact shows that all the

Table 3. End-member representatives for each geodiversity metric and the corresponding observation. Note how each end-member is represented by a single scalar value. This is a requirement for input into PCA, but also allows commonality constraint values to be determined.

Metric	Measures	Minimum end-member	Maximum end-member
Geometric geodiversity end-members			
Volume	Volume of $Bs8$	Model 37: 16 095.75 km^3	Model 21: 17 648.85 km^3
Volume	Volume of tks	Model 78: 197.95 km^3	Model 21: 229.7 km^3
Complexity	Complexity of tkc	Model 70: 3.1141	Model 26: 3.3186
Complexity	Complexity of Bs	Model 61: 1.3825	Model 1: 1.9825
Depth—deepest	Deepest occurrence of $Bs8$	Model 11: 1400 m	Model 100: 1200 m
Depth—deepest	Deepest occurrence of tks	Model 31: 400 m	Model 12: 200 m
Depth—shallowest	Shallowest occurrence of Bs	Model 101: 1400 m	Model 32: 1200 m
Depth—shallowest	Shallowest occurrence of $Bs8$	Model 92: 1400 m	Model 61: 1000 m
Contact relationship	Contact between Bs and $Bs8$	Model 32: 21 348 voxels	Model 72: 26 464 voxels
Contact relationship	Contact between Bs and tkc	Model 90: 1106 voxels	Model 32: 2975 voxels
Curvature	km of contact between $Bs8$ and tks	Model 57: 2.9995e−05	Model 28: 1.8792e−05
Curvature	km of contact between Bs and $Bs8$	Model 54: −4.6455e−05	Model 84: 7.4882e−04
Curvature	kg of contact between Bs and $Bs8$	Model 89: −1.0559e+21	Model 31: −1.4424e−08
Curvature	kg of contact between $Bs8$ and tks	Model 8: −4.1816e+14	Model 12: −6.5508e−08
Geophysical geodiversity end-members			
rms	rms misfit between observed and calculated response	Model 37: 8.2963	Model 26: 8.4930
Standard deviation	Global measure of geophysical variability within the calculated response grid	Model 70: 17.4038	Model 39: 17.5306
Entropy	Global measure of geophysical randomness within the residual grid	Model 70: 4.9151	Model 39: 4.9509
2-D correlation coef.	Comparison between observed and calculated grids—accounts for similar patterns	Model 32: 0.9465	Model 37: 0.9503
Hausdorff distance	Global measure of geophysical misfit—accounts for pattern translation, rotation and dilation	Model 26: 482.8557	Model 76: 487.8138

surfaces throughout the model suite display an overall antiformal curvature. This result seems to be counter-intuitive to geological reason. The $Bs8$/tkc contact forms the basement to the Tarkwaian Basin which one would expect to be synformal, even with the four deformation events that have shaped it to the present-day geometry. Visual inspection also suggests the same conclusion (Fig. 11). The reason for this seemingly anomalous result is the fact the curvature is calculated between adjacent voxels. Overall curvature appears to be synformal, influenced heavily by curvature at the edges of the basin. However, throughout the centre of the basin the curvature exhibits more antiformal geometry, more than the synformal, resulting in the average mean curvature detected by the geodiversity metrics.

An interesting result from the Gaussian curvature (k_g) metric is that all the results return a negative number. This results from both principal curvatures (k_1 and k_2) having opposite signs, meaning that the surfaces all exhibit a saddle or inverted saddle geometry (Lisle & Toimil 2007). The magnitude of k_g values range from close to zero (model 31, Bs/$Bs8$–model 12, $Bs8$/tks) to extreme (model 89, Bs/$Bs8$ and model 8, $Bs8$/tks), meaning that models 31 and 12 exhibit almost cylindrical fold geometry, whereas models 89 and 8 would exhibit distinctive saddle geometries. The saddle geometries are intuitively expected given the polyphase deformation history of the Tarkwaian and differing axes of shortening (Perrouty *et al.* 2012). The cylindrical fold geometries are not expected for the same reasons. Therefore, any use of models 31 and 12 (or others showing near-cylindricity) in further analysis should be performed with caution.

4.2 Gravity misfit comparison

Sensitivity analysis of all the gravity geodiversity metrics was performed to confirm that each metric has a similar variation to range ratio and that no correction was required to normalize any extremely high or low measurements. This was initially difficult to determine as each geophysical geodiversity measure has different units of measurement. Table 4 shows the comparison of metrics. The value to note is the range–standard deviation ratio, which is similar for all metrics (highlighted in grey within Table 4). This means that the degree of variability within each measure is proportional to its range and that comparison can be made between these metrics as end-members and no data levelling is required for PCA.

Geophysical response modelling is by no means the only method by which a model can be verified. For example, drill logs and/or field mapping can be integrated with geophysical observations to provide a less ambiguous solution. In practice though, geophysical data are often the only data set that has full coverage over the study area and subsequently suffices for first-pass model validation.

4.2.1 rms misfit

Overall, the rms misfit values are reasonably high for this kind of study. Obviously, one would hope for slightly lower values, but the magnitude is also not high enough to reject the geology model outright. Fig. 12 shows the difference between the end-member residual grids is not immediately obvious, though regions through the centre and the edges of the Tarkwaian Basin appear to show the greatest variation. The highest magnitude anomalies appear to
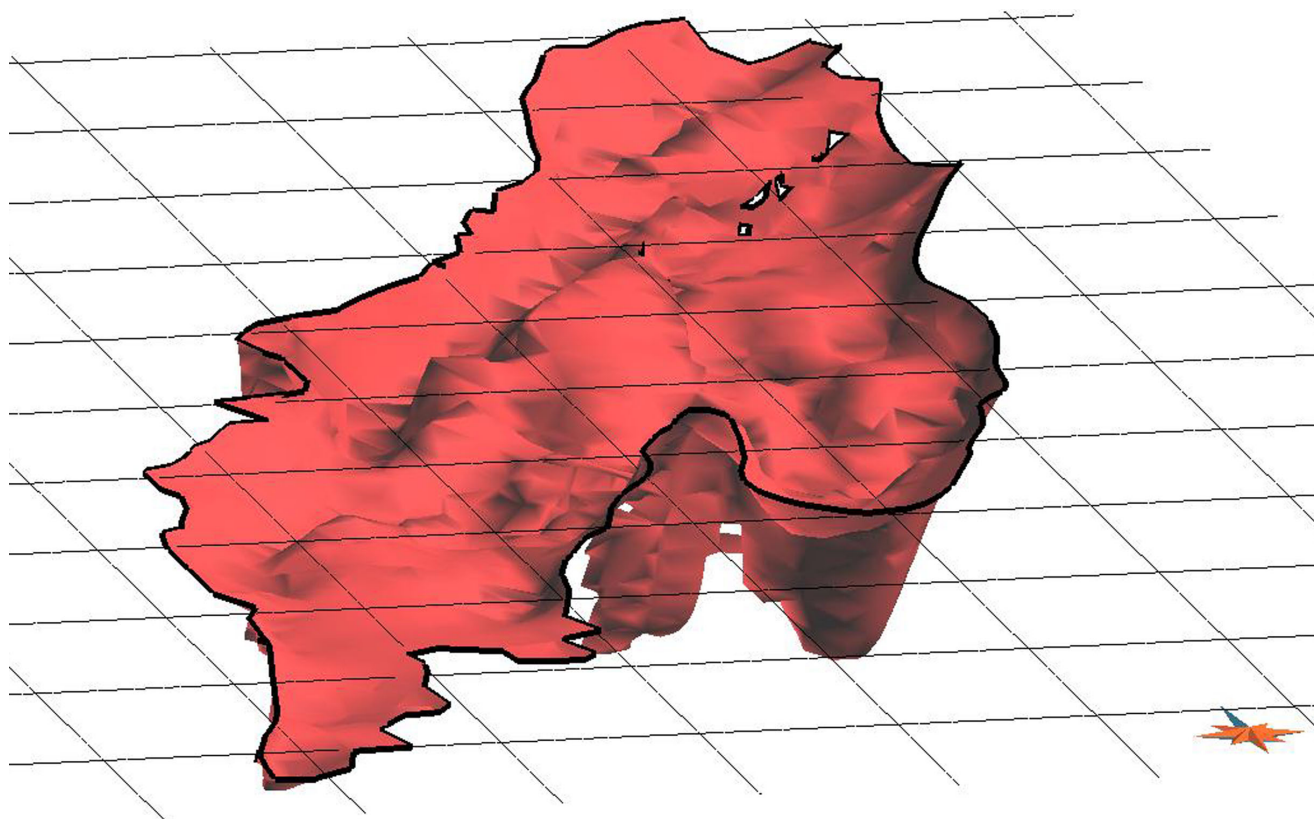
Figure 11. View of the ***tkc*** surface from the Ashanti Greenstone Belt model (from the south), showing an overall synformal curvature (vertical exaggeration x4).

Table 4. Statistical analysis of geophysical geodiversity metrics. A ratio of range to standard deviation has been calculated to evaluate whether the proportion of variation to range in the results across the model suite were consistent, which appears to be the case with all metrics giving similar ratio values. This confirms that there is no metric that determines outliers (extremely high or low values) the presence of which may distort further analysis in PCA.

	Hausdorff distance	rms	2-D correlation coefficient	Entropy	Standard deviation
Range	4.9580	0.1967	0.0039	0.0358	0.1268
Standard deviation	1.1325	0.0424	0.00085	0.0081	0.0283
Range/standard deviation	4.3779	4.6392	4.5240	4.4164	4.4814
Mean	485.4841	8.4034	0.9482	1.8077	17.4716
Max	482.8557	8.4930	0.9503	1.8729	17.5306
Min	487.8138	8.2963	0.9465	1.7290	17.4038

be concentrated in the northwest and southeast. The reason for this is that measured gravity in these areas is responding to deeper Birimian-age structures (Perrouty *et al.* 2012) that were not included in the model, and this is subsequently reflected in the relatively high misfit values. The residual observed over the Tarkwaian is not as severe as in the northwestern and southeastern regions of the map, and the modelled geology appears to match the gravity data relatively well.

Low entropy ('*E*') reflects less randomness in the residual image. If the entire image has predictably (i.e. less random) high misfit values, then the misfit is predictably high. Therefore, low *E* values need to be considered in combination with the mean misfit value for the residual. A low rms value reported with a high-entropy value will represent an image with an overall low misfit but riddled with high-magnitude anomalies. Choosing a model for further processing simply because it has a low rms, without considering the corresponding entropy may prove problematic if high-magnitude anomalies are present.

The remaining geophysical metrics produce a single scalar value that describes the phenomenon measured by each particular technique. The degree to which scores vary from the mean for each metric is relatively similar. Therefore, we can expect that both techniques may be measuring relatively similar degrees of misfit for each calculated model response against the observed response. However, identical results have not been produced for a given model, meaning that each technique may be measuring different phenomena, as was hoped. The PCA should reveal what phenomena each geophysical metric is measuring, with the proviso that a companion geodiversity metric exists that also quantifies the same phenomena.

4.3 Combined PCA and geodiversity metric relationships

PCA results are best analysed using a biplot diagram shown in Fig. 13(a). A short explanation is required to explain how to read a PCA biplot. The first two principal components are plotted along
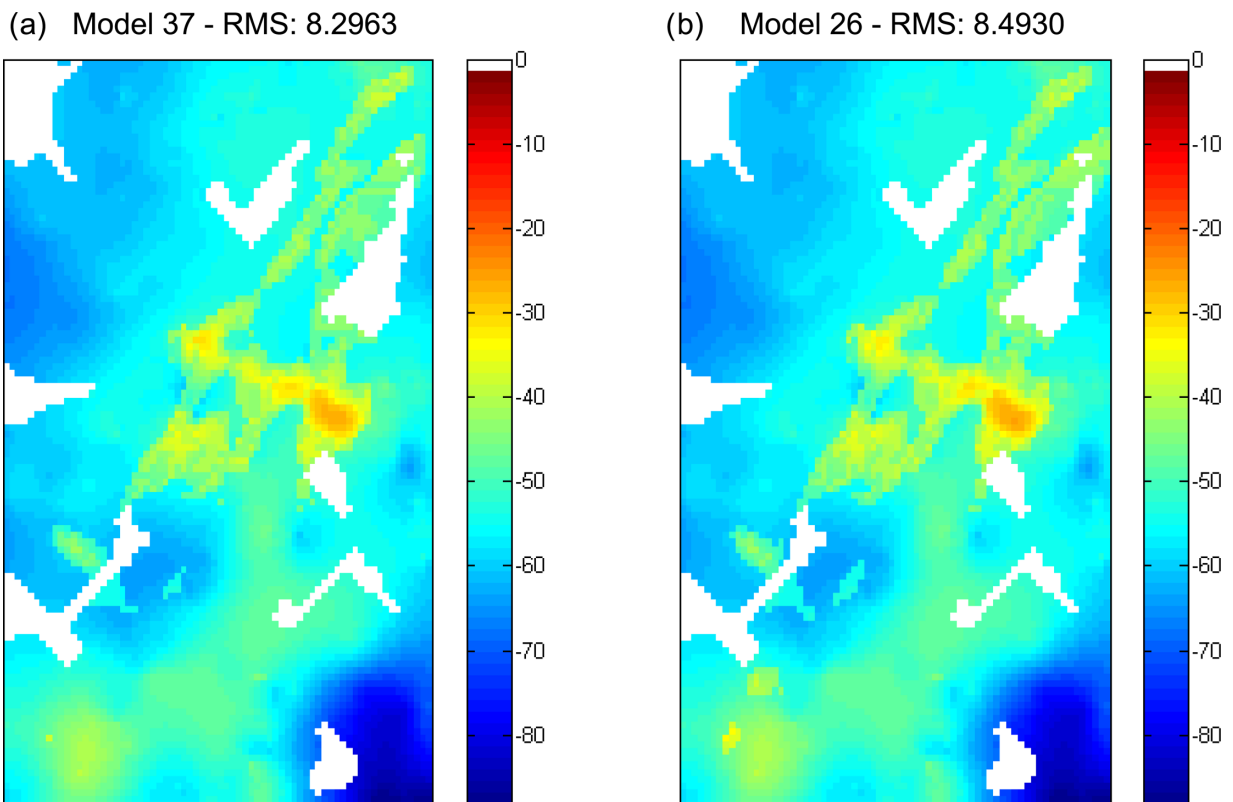
(a)   Model 37 - RMS: 8.2963

(b)   Model 26 - RMS: 8.4930

Figure 12. Residual grids from rms misfit end-member models 37 (rms: 8.2963) and 26 (rms: 8.4930).

the *X*- and *Y*-axes, respectively. The first two principal components together explain just under 50 per cent of model suite variability (Fig. 13b). The PCA plot shows each metrics contribution to model suite variability and how each model is represented in terms of model suite variability. The points represent models within the model suite and the distance of the point from the 0,0 intersection represents the distance of the model from the 'barycentre' of the model suite. The barycentre is a region containing models that share common characteristics, therefore the further away from the barycentre, the more diverse the models become. The models that plot around the edges of the diagram, the 'outliers', define the boundaries of model space. These outlier points represent the models that are dissimilar in terms of the geodiversity metrics that we have used to define the models within model space. The length and direction of the vectors extending from the 0,0 intersection represent how much they contribute to the principal components in the plot. Vector direction represents association with the component (or axis) the vector plots closest to. A long vector that plots close to the *x*-axis would shows a close association to the first principal component, representing a variable that contains a high proportion of model suite variability.

4.4 Ashanti Greenstone Belt PCA

The most distinctive feature of this PCA diagram is the clustering of the rms, standard deviation, entropy and complexity (*tkc*) vectors around the positive *x*-axis. The clustering shows that this group of metrics more effectively measure intrinsic model suite variability than the others in the diagram. Perhaps more interesting is that the rms of the gravity residual of the model suite is linked with the complexity of *tkc* (Fig. 14), and is a surprising result. Given that *Bs*

and *Bs8* make up a high proportion of the model volume, it would be reasonable to expect that these units would have correspondingly high influence over the overall geophysical response, but is not the case in this example. The volume of *Bs8* is still influential in terms of model suite variability, plotting close to the *y*-axis (along with the complexity of the *Bs*/*Bs8* contact), but is not associated with the geophysical metrics. The positive correlation between the complexity of the *tkc* contact and gravity residual rms can be attributed to the sampling. The sampling parameters for the complexity metric are of a higher resolution than that of gravity survey (500 m cell size complexity versus 4.6 km for the gravity survey). The gravity signal is unable to adequately resolve the geological complexity of the *Bs8*/*tkc* contact at this lower resolution, suggesting that higher resolution gravity sampling needs to be performed if gravity is to be used effectively in resolving the architecture of the contact between the Late Birimian and the Tarkwaian Group. The remaining 2-D correlation and Hausdorff distance geophysical metrics show no close association with other geometrical metrics, though they are almost covariant. The association the two metrics have with each other is due to their ability to recognize patterns within the data.

The opposing direction of the vectors for 2-D correlation and the Hausdorff distance to the other geophysical metrics is due the inverse relationship between these measures. For example, the Spearman's Rank between the rms and Hausdorff distance is -0.81, showing a medium-to-high correlation (Fig. 15). Table 3 shows that the rms end-member representing the highest geophysical misfit is model 26, whereas the Hausdorff distance end-member representing the least distance from the observed grid is model 26. We suspect that this reversal of end-members is due to the Hausdorff distance metric recognizing a pattern in the model 26 calculated grid that has a close resemblance to the pattern in the observed grid. This pattern may have been dilated or translated, as the rms misfit is
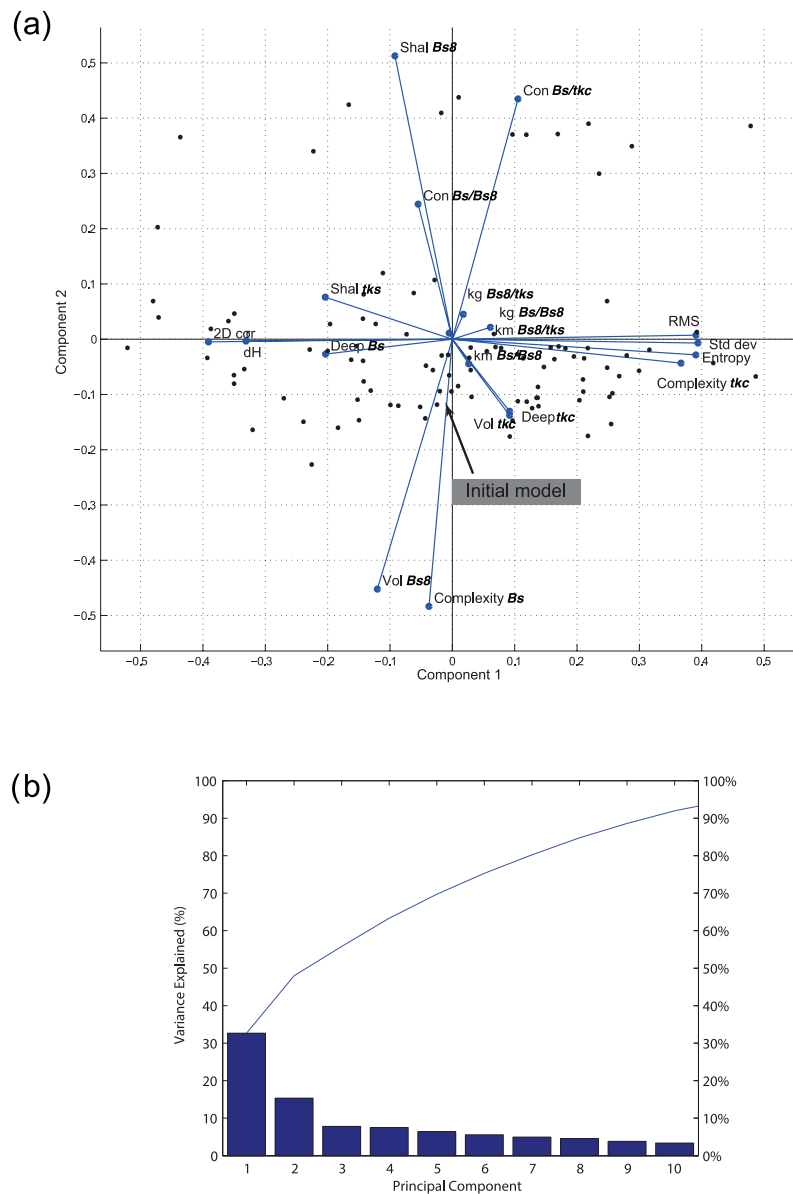
Figure 13. PCA for the Ashanti Greenstone Belt model suite and geodiversity metrics. (a) Biplot diagram of the combined PCA showing that rms, standard deviation, entropy geophysical metrics all measure similar variations between models as does the complexity metric of ***tkc***. Note the location of the initial model. (b) Pareto diagram of principal components and the variance explained. Almost half the variance observed by the geodiversity metrics within the model suite is explained by the first two components, negating the need to examine components three onwards.

low, but may exhibit the same geometrical features. Identification of the exact transformation the Hausdorff distance has identified was not performed within this study, but would obviously be a useful outcome for future work.

Most of the metrics employed in the Ashanti Belt geodiversity analysis have proven useful in describing model suite variability. Some metrics, such as depth and volume tend to be more effective when applied to certain stratigraphic units. For example, the shallowness of ***Bs8*** appears to explain a considerable proportion of the variability in the second principal component, whereas the shallowness of ***tks*** appears ineffective in this respect (Fig. 13a). The shallowness of ***tks*** was found to be influential on the variability of the model suite when compared to other units, but when compared with other metrics, such as the geophysical, complexity or contact surface area, appear less so. The mean and Gaussian curvature met-

rics seem less useful than the other metrics in identifying variability in the Ashanti Greenstone Belt model suite. Two implications of this result are possible: (1) the curvature metrics may be more effective in models displaying different, more curvaceous surfaces, or (2) need to be revised to more effectively represent differences in curvature between models. The inability of the curvature metrics to determine variability in the model suite is consistent with the results from the Lindsay *et al.* (2013) study suggesting that the curvature metrics need to be revised.

The reason to construct this model was to better understand the geometrical nature of the Tarkwaian Basin, therefore little attention was paid to the underlying Birimian structures to accommodate petrophysical property heterogeneities in the Birimian basement. The complication is that if this model was subjected to geophysical inversion, Tarkwaian geometries may be adjusted
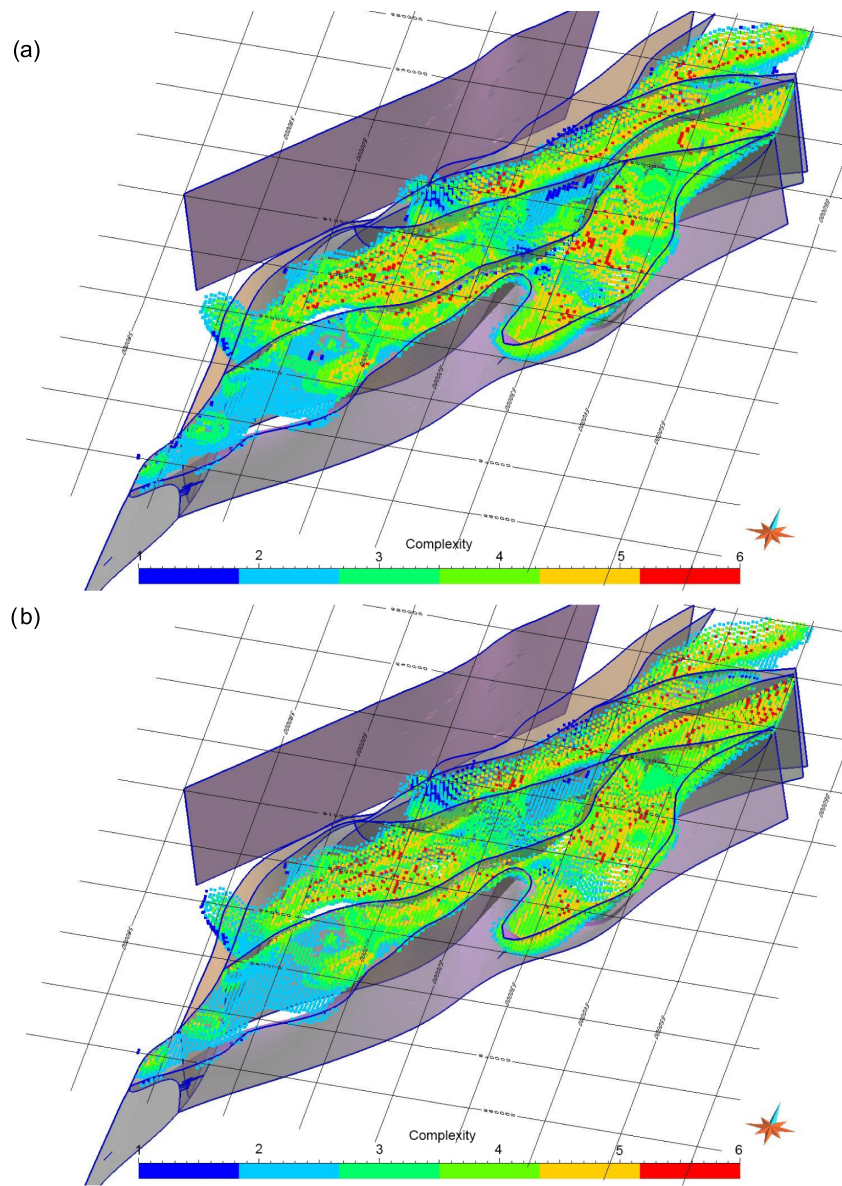
Figure 14. Complexity maps of *tkc* formation from models 70 (a) and 26 (b). Modelled fault surfaces (grey with blue borders) are shown.

beyond the boundaries of geological feasibility. As the Birimian geophysical response has not been adequately represented geometrically, the inversion process would likely adjust Tarkwaian structures or petrophysical properties. Tarkwaian structures are the only features that will likely produce an improvement in misfit when adjusted in the model, and these changes may produce geologically unreasonable results. The obvious course of action is to better represent Birimian age structures in future model versions to avoid unnecessary adjustment of model attributes during inversion.

Hotelling's T^2 scores have been calculated to rank each model according to distance from the barycentre (Table 5). All outlier models (8, 84, 32, 61 and 37) feature in the end-member analysis (Table 3), though determining their influence on overall model suite variability is difficult to determine by simple ranking procedure. Model rank needs to be weighted according to the influence of the metric to overall model suite variability. A combined approach is needed to completely acknowledge all geodiversity within the model suite. Similar to results obtained in Lindsay *et al.* (2013),

the initial model does not feature in the bottom five ranked models. The initial model is reasonably close to the barycentre, being ranked in 88th position, but there are 13 other models that are considered more similar according to the geodiversity metrics used in this study. The Hotelling's T^2 ranking results have implications on typical modelling procedures that produce a single realization of the input data, realizations which potentially misrepresent input data. Uncertainty inherent in 3-D geological models necessitates employing modelling techniques that produce multiple geological realizations to aid understanding of the full range and variation of possibilities. Modelling procedures may therefore benefit from probabilistic models, rather than models that have been produced from processes that optimize data.

The method presented here focuses on precisional uncertainties associated with input orientation measurements. Conceptual uncertainty associated with model topology (i.e. stratigraphy, fault–fault and fault–stratigraphy relationships; Cherpeau *et al.* 2010; Lindsay *et al.* 2012) can also be assessed with the geodiversity method, however metrics would need to be developed that recognize
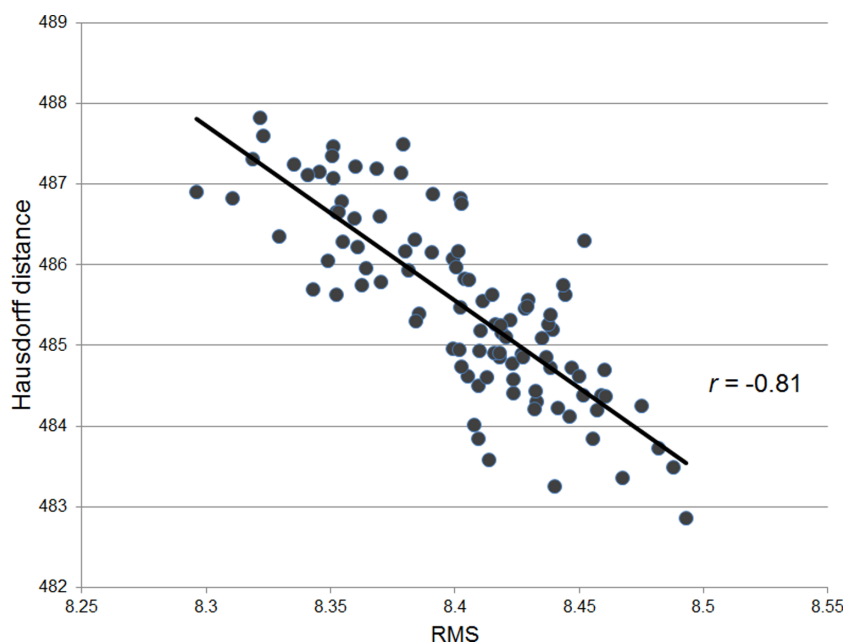
Figure 15. Comparison of rms misfit scores and Hausdorff distances (*y*-axis) between calculated and observed geophysical grids for each model (*x*-axis). A reasonably strong negative correlation between both data sets ($r = -0.81$).

Table 5. Hotelling's T^2 score rankings for the model suite. The top five models are those that exhibit the greatest distance from the model space barycentre. The bottom five models represent are closest to the barycentre. The ranking of the initial model is also shown.

	Model	Model
Top 5	1	8
	2	84
	3	32
	4	61
	5	37
	88	**Initial model**
Bottom 5	97	50
	98	9
	99	67
	100	34
	101	11

changes in model topology. Fault intersection and fault complexity maps or stratigraphic orientation vectors can be added to the stable of geodiversity metrics to quantify conceptual uncertainty.

5 CONCLUSION

Adding geophysical geodiversity metrics adds further information to model space exploration. The process of comparing different geophysical techniques has revealed that some are associated with the geological complexity of a gold prospective layer, *tkc*, which forms the base of the Tarkwaian Basin. Complexity maps of the two end-member models (Fig. 14) could provide useful input into prospectivity modelling, but also guide further modifications and changes to the initial model in combination with uncertainty analysis by identifying regions requiring specific focus.

The geophysical metrics included in this study have provided additional methods to calculate the misfit between calculated and

observed grids. The pattern recognition feature of the Hausdorff distance provides a useful companion technique to typical rms misfit calculations. The Hausdorff distance could be an inclusion for further studies due to its ability to detect patterns in the observed response that may exist, translated, rotated or dilated, in the calculated response.

PCA has also revealed that the geophysical response is associated with a geometrical phenomenon. Further investigation of the model space through geophysical inversion can be streamlined by identifying models and elements that should be analysed. Based on the results obtained in the PCA, we suggest that geophysical inversion on *tks*/*tkc* or the ***Bs8***/*tkc* contact may aid in producing a model honouring both geological and geophysical data. We can be more confident with the final result if models exhibiting both common (barycentre) and unusual (outlier) geometries are included in analysis. The inversion process can also be guided to focus on elements that have been determined to be important to reconcile when trying to obtain a model honouring both geological and geophysical data. This will be achieved by (1) eliminating model realizations that do not represent either the barycentre or outlier regions of model space and (2) identifying geometrical elements, while potentially important in other non-geophysical contexts, that do not contribute significantly to model suite variability. The production of models that best represent all input data and geological possibility will be aided by following the procedure outlined here.

Geodiversity PCA has implications for the future of 3-D geological modelling. The initial model was found not to be the most representative model in the model suite. This finding exposes flaws in workflows that consider a single model to adequately represent geological possibility. We propose two processes that address these flaws: (1) multiple models, identified through PCA and geodiversity analysis, to be produced as the product of simulation workflows and (2) combining all models in the model suite to produce a probabilistic model, where surfaces and structures are represented by probability measures, rather than as discrete surfaces. The result from both proposals is that uncertainty inherent in geological studies is acknowledged and is communicated effectively.

ACKNOWLEDGEMENTS

REFERENCES

Adadey, K., Clarke, B., Théveniaut, H., Urien, P., Delor, C., Roig, J.Y. & Feybesse, J.-L., 2009. *Geological map explanation - Map sheet 0503 B (1:100 000), CGS/BRGM/Geoman,* Geological Survey Department of Ghana (GSD). No MSSP/2005/GSD/5a.

Allibone, A.H., McCuaig, T.C., Harris, D., Etheridge, M.A., Munroe, S., Bryne, D., Amanor, J. & Gyapong, W., 2002. Structural controls on gold mineralization at the Ashanti gold deposit, Obuasi, Ghana, in *Integrated Methods for Discovery: Global Exploration in the 21st Century,* pp. 65–93, eds Goldfarb, R.J. & Neilson, R.L., Society of Economic Geologists.

Barritt, S.D. & Kuma, J.S., 1998. Constrained gravity models and structural evolution of the Ashanti Belt, southwest Ghana, *J. Afr. Earth Sci.,* **26**(4), 539–550.

Betts, P.G., Valenta, R.K. & Finlay, J., 2003. Evolution of the Mount Woods Inlier, northern Gawler Craton, Southern Australia: an integrated structural and aeromagnetic analysis, *Tectonophysics,* **366**(1–2), 83–111.

Blenkinsop, T., Schmidt, M., Kumi, R. & Sangmoor, S., 1994. Structural geology of the Ashanti Gold Mine, in *Geologisches Jahrbuch D100,* pp. 131–153, ed. Oberthür, T., E Schweizerbart Science Publishers.

Bond, C.E., Gibbs, A.D., Shipton, Z.K. & Jones, S., 2007. What do you think this is? "Conceptual uncertainty" in geoscience interpretation, *GSA Today,* **17**(11), 4–10.

Bond, C.E., Philo, C. & Shipton, Z.K., 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.

Calcagno, P., Chilès, J.P., Courrioux, G. & Guillen, A., 2008. Geological modelling from field data and geological knowledge: part I. Modelling method coupling 3D potential-field interpolation and geological rules, *Phys. Earth planet. Inter.,* **171**(1–4), 147–157.

Cherpeau, N., Caumon, G. & Lévy, B., 2010. Stochastic simulations of fault networks in 3D structural modeling (Simulations stochastiques de réseaux de failles en modélisation structurale 3D), *C. R. Geosci.,* **342**(9), 687–694.

Chugunov, N., Shepelyov, G. & Sternin, M., 2008. *Probabilistic methods for uncertainty quantification,* in *Encyclopedia of Decision Making and Decision Support Technologies,* pp. 732–742, IGI Global.

Feybesse, J.-L., Billa, M., Guerrot, C., Duguey, E., Lescuyer, J.-L., Milesi, J.-P. & Bouchot, V., 2006. The paleoproterozoic Ghanaian province: geodynamic model and ore controls, including regional stress modeling, *Precambrian Res.,* **149**(3–4), 149–196.

Foudil-Bey, N., 2012. *Développement d'outils d'interprétation de données géophysiques, PhD thesis,* Université du Québec en Abitibi-Témiscamingue et Université de Lorraine.

Gao, Y. & Leung, M.K.H., 2002. Face recognition using line edge map, *IEEE Trans. Pattern Anal. Mach. Intell.,* **24**(6), 764–779.

Gershon, N., 1998. Visualization of an imperfect world, *IEEE Comput. Graph. Appl.,* **18**(4), 43–45.

Gonzalez, R.C., Woods, R.E. & Eddins, S.L., 2003. *Digital Image Processing using MATLAB,* Vol. 11, Prentice Hall.

Gray, A., Abbena, E. & Salamon, S., 2006. *Modern Differential Geometry of Curves and Surfaces with Mathematica,* Chapman & Hall/CRC.

Hastings, D.A., 1982. On the tectonics and metallogenesis of West Africa: a model incorporating new geophysical data, *Geoexploration,* **20**(3–4), 295–313, 317–327.

Hirdes, W. & Nunoo, B., 1994. The Proterozoic paleoplacers at Tarkwa Gold Mine, SW Ghana: sedimentology, mineralogy, and precise age dating of the main reef and west reef, and bearing of the investigations on source area aspects, *Geologisches Jahrbuch,* **D100,** 247–311.

Holstein, H., 2003. Gravimagnetic anomaly formulas for polyhedra of spatially linear media, *Geophysics,* **68**(1), 157–167.

Holstein, H., Schürholz, P., Starr, A. & Chakraborty, M., 1999. Comparison of gravimetric formulas for uniform polyhedra, *Geophysics,* **64**(5), 1438–1446.

Hotelling, H., 1931. The generalization of student's ratio, *Ann. Math. Stat.* **2**(3), 360–378.

Huttenlocher, D.P., Klanderman, G.A. & Rucklidge, W.J., 1993. Comparing images using the Hausdorff distance, *IEEE Trans. Pattern Anal. Mach. Intell.,* **15**(9), 850–863.

Jessell, M.W., Ailleres, L. & de Kemp, E.A., 2010. Towards an integrated inversion of geoscientific data: what price of geology? *Tectonophysics,* **490**(3–4), 294–306.

Jolliffe, I.T., 2002. *Principal Component Analysis,* 2nd edn, Springer.

Joly, A., Martelet, G., Chen, Y. & Faure, M., 2008. A multidisciplinary study of a syntectonic pluton close to a major lithospheric-scale fault—relationships between the Montmarault granitic massif and the Sillon Houiller Fault in the Variscan French Massif Central: 2. Gravity, aeromagnetic investigations, and 3-D geologic modeling, *J. geophys. Res.,* **113**, B01404, doi:10.1029/2006JB004744.

Krzanowski, W.J., 1995. Selection of variables, and assessment of their performance, in mixed-variable discriminant analysis, *Comput. Stat. Data Anal.,* **19**(4), 419–431.

Lajaunie, C., Courrioux, G. & Manuel, L., 1997. Foliation fields and 3D cartography in geology: principles of a method based on potential interpolation, *Math. Geol.,* **29**(4), 571–584.

Lindsay, M.D., Aillères, L., Jessell, M.W., de Kemp, E.A. & Betts, P.G., 2012. Locating and quantifying geological uncertainty in three-dimensional models: analysis of the Gippsland Basin, southeastern Australia, *Tectonophysics,* **546–547,** 10–27.

Lindsay, M.D., Jessell, M.W., Ailleres, L., Perrouty, S., de Kemp, E.A. & Betts, P.G., 2013. Geodiversity: exploration of 3D geological model space, *Tectonophysics,* **594,** 27–37.

Lisle, R.J. & Robinson, J.M., 1995. The Mohr circle for curvature and its application to fold description, *J. Struct. Geol.,* **17**(5), 739–750.

Lisle, R.J. & Toimil, N.C., 2007. Defining folds on three-dimensional surfaces, *Geology,* **35**(6), 519–522.

Loh, G., Hirdes, W., Anani, C., Davis, D.W. & Vetter, U.K., 1999. Explanatory notes for the geological map of Southwest Ghana 1: 100,000, Geologisches Jahrbuch, Reihe B, Heft 93, 150p, Bundesanstalt für Geowissenschaften und Rohstoffe.

Mallet, J.-L., 2002. *Geomodeling,* Oxford University Press, 599pp.

Metelka, V., Baratoux, L., Naba, S. & Jessell, M.W., 2011. A geophysically constrained litho-structural analysis of the Eburnean greenstone belts and associated granitoid domains, Burkina Faso, West Africa, *Precambrian Res.,* **190,** 48–69.

Mynatt, I., Bergbauer, S. & Pollard, D.D., 2007. Using differential geometry to describe 3-D folds, *J. Struct. Geol.,* **29**(7), 1256–1266.

O'Gorman, L., Sammon, M.J. & Seul, M., 2008. *Practical Algorithms for Image Analysis: Description, Examples, Programs, and Projects,* Cambridge University Press.

Okabe, M., 1979. Analytical expressions for gravity anomalies due to homogeneous polyhedral bodies and translations into magnetic anomalies, *Geophysics,* **44**(4), 730–741.

Olson, C.F. & Huttenlocher, D.P., 1997. Automatic target recognition by matching oriented edge pixels, *IEEE Trans. Image Process.,* **6**(1), 103–113.

Perrouty, S., 2012. *Evolution structurale de la Ceinture Minéralisée d'Ashanti, Sud-Ouest Ghana, PhD thesis,* Université Toulouse III – Paul Sabatier.

Perrouty, S., Aillères, L., Jessell, M.W., Baratoux, L., Bourassa, Y. & Crawford, B., 2012. Revised Eburnean geodynamic evolution of the gold-rich southern Ashanti Belt, Ghana, with new field and geophysical evidence of pre-Tarkwaian deformations, *Precambrian Res.,* **204–205,** 12–39.

Pigios, J.-P., Groves, D.I., Fletcher, I.R., McNaughton, N.J. & Snee, L.W., 2003. Age constraints on Tarkwaian palaeoplacer and lode-gold formation in the Tarkwa-Damang district, SW Ghana, *Miner. Deposita,* **38,** 695–714.

Plouff, D., 1976. Gravity and magnetic fields of polygonal prisms and application to magnetic terrain corrections, *Geophysics,* **41**(4), 727–741.

Polson, D. & Curtis, A., 2010. Dynamics of uncertainty in geological interpretation, *J. geol. Soc.,* **167**(1), 5–10.

Putz, M., Stüwe, K., Jessell, M. & Calcagno, P., 2006. Three-dimensional model and late stage warping of the Plattengneis Shear Zone in the Eastern Alps, *Tectonophysics,* **412**(1–2), 87–103.

Rankey, E.C. & Mitchell, J.C., 2003. That's why it's called interpretation: impact of horizon uncertainty of seismic attribute analysis, *Leading Edge,* **22**(9), 820–824, 826, 828.

Rucklidge, W.J., 1997. Efficiently locating objects using the Hausdorff distance, *Int. J. Comput. Vis.,* **24**(3), 251–270.

Sestini, G., 1973. Sedimentology of a Paleoplacer: The Gold-Bearing Tarkwaian of Ghana, in *Ores in Sediments,* pp. 275–305, Springer.

Shannon, C.E., 1948. A mathematical theory of communication, *Bell Syst. Tech. J.,* **27,** 379–423.

Sim, D.-G., Kwon, O.-K. & Park, R.-H., 1999. Object matching algorithms using robust Hausdorff distance measures, *IEEE Trans. Image Process.,* **8**(3), 425–429.

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.

Thomson, J., Hetzler, E., MacEachren, A., Gahegan, M. & Pavel, M., 2005. A typology for visualizing uncertainty, in *Proceedings of SPIE,* The International Society for Optical Engineering, Vol. 5669, pp. 146–157.

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

Tunks, A.J., Selley, D., Rogers, J.R. & Brabham, G., 2004. Vein mineralization at the Damang Gold Mine, Ghana: controls on mineralization, *J. Struct. Geol.,* **26,** 1257–1273.

Viard, T., Caumon, G. & Lévy, B., 2010. Adjacent versus coincident representations of geospatial uncertainty: which promote better decisions? *Comput. Geosci.,* **37**(4), 511–520.

Wang, L. & Suter, D., 2007. Learning and matching of dynamic shape manifolds for human action recognition, *IEEE Trans. Image Process.,* **16**(6), 1646–1661.

Wellmann, J.F. & Regenauer-Lieb, K., 2012. Uncertainties have a meaning: information entropy as a quality measure for 3-D geological models, *Tectonophysics,* **526–529,** 207–216.

Wellmann, J.F., Horowitz, F.G., Schill, E. & Regenauer-Lieb, K., 2010. Towards incorporating uncertainty of structural data in 3D geological inversion, *Tectonophysics,* **490**(3–4), 141–151.

Williams, H.A., Betts, P.G. & Ailleres, L., 2009. Constrained 3D modeling of the Mesoproterozoic Benagerie Volcanics, Australia, *Phys. Earth planet. Inter.,* **173**(3–4), 233–253.

Yeten, B., Brouwer, D.R., Durlofsky, L.J. & Aziz, K., 2004. Decision analysis under uncertainty for smart well deployment, *J. Petrol. Sci. Eng.,* **43**(3–4), 183–199.