

# Categorising features of geological terranes with geodiversity metrics: Enhancing exploration of multiple geological models

**M.D. Lindsay**<sup>ab</sup>, **S. Perrouty**<sup>ac</sup>, **M. J. Jessell**<sup>c</sup>, **L. Ailleres**<sup>a</sup>, **E. de Kemp**<sup>d</sup> and **P.G. Betts**<sup>a</sup>

<sup>a</sup> School of Geosciences, Monash University, Victoria, Australia

<sup>b</sup> Geosciences Environnement Toulouse (GET), Université de Toulouse III, France

<sup>c</sup> IRD, Geosciences Environnement Toulouse (GET), Université de Toulouse III, France

<sup>d</sup> Geological Survey of Canada, Ottawa, Canada

Email: mark.lindsay@monash.edu

**Abstract:** Most geoscientific research programmes benefit from three-dimensional (3D) representations of geology. Various elements of a geological target can be visualised, analysed and quantified to better understand the spatial properties of prospective terranes. For example, current technologies are able to produce useful measures that describe proven or prospective ore deposits. Essential information such as host and source rock proximity relationships can be estimated and analysed simultaneously with the location and prevalence of particular geological features (such as faults, lithologies, folds, resource estimates and mineral distribution) to generate 3D prospectivity maps that help to guide exploration activity.

The geological elements of a 3D model are defined by a suite of data including field observations, geophysical interpretation and the prevailing tectonic evolution hypothesis. Field data (consisting of orientation measurements and lithological observations) are often supported by interpreted geophysics in covered terranes. In addition, the tectonic evolution hypothesis describing the timing of important geological events also has a large influence on the stratigraphic column, fault networks and interactions between the modelled elements. All of these input data are prone to error and uncertainty and may produce a model that does not adequately represent actual geology. In particular, a heavy reliance on geophysical interpretation introduces a high risk of ambiguity as it is difficult to explicitly identify lithological and structural fabric orientations. Subsequently much effort is made to remove error and uncertainty from the inputs to produce a single, optimised model that represents the geology in a useful and reliable manner.

Removing error from the input data is difficult and, in some cases, almost impossible to perform. There is a risk that a reduced set of measurements that produces a model best representing the geology can be removed in the process. Our philosophy is to examine model reliability by simulating the error in input data. The data is subjected to uncertainty simulation prior to model input and involves varying strike and dip observations that determine modelled geological geometries. The subsequent sets of varied strike and dip observations are used to calculate multiple geological 3D models. The result is a suite of models that represent the range of possibilities offered by the input data set.

We present this technique using a part of the palaeoproterozoic Ashanti Greenstone Belt, southwestern Ghana and the Gippsland Basin, southeastern Australia in a comparative case study. Geological knowledge in these regions can benefit from this technique as it produces an interesting set of ‘what-if’ scenarios, expanding our understanding of the interaction between geological elements considered important for gold mineralisation (Ashanti Greenstone Belt) or oil and gas prospectivity (Gippsland Basin). We perform analysis on the model suites using Principal Component Analysis (PCA) to determine the important features and characteristics. The most-different models or ‘end-members’ of a model suite can be identified given a particular geological attribute, be it depth (deep or shallow) or volumes (large and small) of a particular stratigraphic unit, fault relationships or magnitude of deformation. These attributes, or ‘geodiversity’ metrics, can provide invaluable information to the geoscientist. The geodiversity metrics are then integrated into a combined study to answer questions regarding geological possibilities in the region providing a comprehensive understanding of geology in the respective geological terranes.

**Keywords:** 3D modelling, uncertainty simulation, gold, multiple models, model suite exploration

## 1. INTRODUCTION

Geological three-dimensional (3D) models are employed to aid various areas of geoscientific research. The quality of the model relies upon the correct location and relationships between various geological elements (Jessell *et al.*, 2010). Therefore the usefulness of the 3D model relies heavily on the ability of the input data (including geophysical information, field geology observations; well/drill logs and elevation or depth data) to provide sufficient information to locate and describe these geological elements. Unfortunately input data is subjected to varying degrees of error and uncertainty in collection or processing that can adversely affect the accuracy of the model (Thore *et al.*, 2002). This approach follows advances made by Jessell *et al.* (2010), Lindsay *et al.* (2010), Wellmann *et al.* (2010) and Wellmann and Regenauer-Lieb (2011) toward perturbation of input orientation data, misfit functions, multiple model generation, calculation and analysis of multiple models, visualisation and quantification of model uncertainty, model quality assessment and visualisation of information entropy. Similar to these previous studies, removal of error or uncertainty from data prior to input is not performed. Instead multiple models are calculated in an attempt to understand the range of geological possibilities that exist when uncertainty has been taken into consideration. The result is a model suite that contains models displaying typical geological geometries and models that provide extreme geological geometries, defining a boundary to the geological possibilities that can be generated from the input data set.

In much the same way many animal species at a particular location can be described using biodiversity metrics, multiple geological models within a model suite can be described using a set of geological diversity, or ‘geodiversity’, metrics. The use of the term geodiversity in the modelling context of this study should not be confused with its use in conservation that, while related, describes the diversity of processes, morphology and mineralogy of the earth. Fundamental questions that should be answered to describe model suite geodiversity are: (1) are these models different; (2) how are these models different and (3) what modelling implications can be drawn from the information that geodiversity metrics provide? An additional aim of this exploratory study is to determine whether there is a universal metric that will always appropriately describe the geodiversity, regardless of the input data set parameters or geological terrane being modelled.

## 2. MULTIPLE MODEL GENERATION

A model suite is created by calculating many models from a data set that has been subjected to uncertainty simulation. Each model is created using an implicit geological modelling application, 3D Geomodeller, that uses the potential field method to calculate geological interfaces as implicit surfaces (Lajaunie *et al.*, 1997). The geological interfaces represent the contact between different geological formations from which geological formations can be interpolated (see Calcagno *et al.* 2008 for additional information). It is required that three different sources of geological information are specified before model calculation can take place:

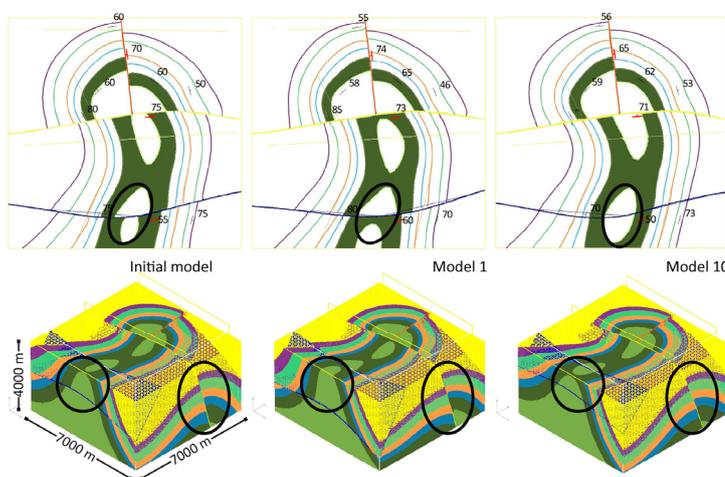
1. A stratigraphic column. This describes the temporal relationships between the geological formations being modelled.
2. The location of geological contacts for each stratigraphic group must be known at some locations.
3. Orientation measurements in the form of strike and dip measurements that describing the orientation of the geological interface.

The uncertainty simulation varies the orientation measurements from the data set  $\pm 5$  degrees (Table 1), a reasonable amount of variation that may be observed between measurements taken by different geologists, especially in weathered or highly deformed terranes. Each data set perturbation is then used to recalculate the model potential field to accommodate the new, varied input data. The results of uncertainty simulation can be seen in the synthetic model (Figure 1). There is potential to generate millions of models in this manner. For the purposes of this study we have generated 100 new models for a total of 101 (including the original model), providing enough opportunity to generate a suite of different and varied model geometries to make a reasonable comparison of the possibilities within the model suite.

Each model is then converted to a voxel – a set of volumetric pixels (voxels) that represent the model as a grid in 3D space. The voxel parameters are stored in Universal Transverse Mercator

	Azimuth	Dip	New Azimuth	New Dip
Formation 5	315	60	320	58
Formation 4	35	60	38	65
Formation 2	240	80	245	85
Formation 2	105	75	107	70
...				

**Table 1.** Sample of original geological orientation measurements (columns to the left) that have been subjected to uncertainty simulation (columns to the right).



**Figure 1.** Comparison of models that have been calculated using data subjected to uncertainty simulation. Note areas of significant difference circled in black. The first row of models are surface maps with input strike and dip measurements, the second are 3D block model representations of the above surface maps.

tor (UTM) co-ordinates so that distances (including depth) and location can be measured in metres and related back to the real-world area of interest. The property of each cell is an integer representing a stratigraphic unit ('stratigraphic ID') in relative position according to the stratigraphic column. A stratigraphic ID value of '1' represents the basement unit of the model, whereas a stratigraphic ID value of '2' represents the unit overlying the basement and so on.

### 3. GEODIVERSITY METRICS

The geodiversity seen within the model suite is clear when comparing the models in Figure 1. The different degrees of displacement along faults (3D block models) and the geometry of rock outcrop (surface maps) can be easily identified. But often the differences are more subtle than can be seen when using purely visual comparative techniques. This is the case when studying the difference in curvature of a particular geological contact. In addition some model properties, such as the volume or depth ranges of a particular geological unit of interest, require quantitative analytical methods to provide useful information to the operator. Visual comparison is also inefficient for model comparison in suites with more than a few members. A set of mathematical analysis techniques have been developed to automatically assess model suite geodiversity that can be used independent of geological terrane or model suite size. The following metrics have been developed to appropriately assess the geodiversity contained within a model suite. These metrics are used to characterize the geology of model allowing relevant comparisons to be made.

#### *Formation depth and volume*

The shallowest and deepest extents of each stratigraphic unit can be determined from each model. This type of information can be of interest to both traditional geoscientific industries, such as oil, gas and minerals exploration (depth of reservoir or deposit) and groundwater and environmental management (depth of aquifer), but also to emerging energy industries such as geothermal and coalbed methane exploration. Shallowest extent of a formation is calculated by determining the shallowest voxel in the formation under study and vice versa for the deepest extent. It is also possible to calculate the volume of each stratigraphic unit within each model of the suite. Particular units can be delineated and examined for economic or scientific research interest to answer questions regarding geological possibilities. Volumes are determined by a count of formation voxels

#### *Average mean curvature*

Calculating the average mean curvature of a geological contact can provide information that describes the degree of geological interface deformation. Curvature is determined by calculating the magnitude of principal curvatures  $k_1$ , a section plane through the fold displaying the maximum curvature, and  $k_2$ , a plane perpendicular to  $k_1$ , displaying the minimum curvature. Mean curvature ( $M$ ) is the arithmetic average of  $k_1$  and  $k_2$ . Gaussian curvature ( $G$ ) is the product of  $k_1$  and  $k_2$ . A positive  $G$  value indicates that both principal curvatures have the same sign (the surface resembles a dome, or if inverted, a bowl or basin) and a negative value indicates the principal curvatures have different signs (the surface resembles a saddle or inverted saddle). Curvature calculations can locate and determine the magnitude of curvature (or folding) observed within a contact. Further analysis of mean curvature and Gaussian curvature in combination can reveal folding directions along two perpendicular axes, allowing the identification of fold geometry (i.e. antiformal synforms or synformal anti-forms) (Lisle and Toimil, 2007).

#### *Neighbourhood relationships*

Neighbour relationships can describe the manner in which different stratigraphic units are juxtaposed. Two types of relationship can be described, 'short-distance' and 'long-distance'. Only short-distance metrics are employed at this stage, though long-distance metrics can be added as their method becomes more refined. Short-distance relationships are those that concern cells directly adjacent to the cell under study. In this study two short-distance relationships are examined. Firstly, the union between two formations, or in geological terms, the contact between two units can be determined. The proportion of the model consisting of differing contact relationships can be calculated. The surface area of the contact between various stratigraphic units can be identified individually and their respective proportions calculated. This information can be beneficial to mineral exploration studies, where the contact between particular units can lead to the identification of a potential mineral resource. For example, an explorer may be interested in under what conditions and which model displays the largest surface-area between a psammitic and psammopelitic unit within a model being used to target a Broken Hill-type deposit.

Secondly, if the stratigraphic units directly adjacent to a given cell can be identified, then the number of different adjacent stratigraphic units can be determined. This number can represent the degree of geological complexity at that point (Figure 2). For example, if a cell had one type of stratigraphic unit adjacent, then it could be assumed that the cell is surrounded by the same stratigraphic unit. This would indicate that the cell is located in an area of low complexity, possibly in the middle of a particular formation. Complexity increases as the number of adjacent stratigraphic units increases, so that a value of two may indicate that the cell is near or on a geological contact between two different units and three indicates that the cell is near or on a triple-junction between units. Individual cells can give information about the complexity of a given point, but a mean value for a given stratigraphic unit can also be calculated representing the geological complexity of that unit as a whole. This complexity information can be used to determine which stratigraphic units within the model may be more difficult to target in drilling programs.

Neighbourhood relationships are calculated with a  $k$ -nearest neighbour algorithm ( $k$ -NN). This method classifies objects based on training areas within the model and finds the closest points in terms of Euclidean distance (Bremner *et al.*, 2005; Friedman *et al.*, 1977). The distances determined by this technique are used to constrain which voxels are counted as neighbours. In the case of the short-distance metrics it follows that only the shortest distances (a six neighbour relationship) are included in determination of the geological relationships.

#### 4. PRINCIPAL COMPONENT ANALYSIS

Can the variability of a particular geodiversity metric describe the geometrical variation between models within the model suite? The measured attributes of the chosen metric can be used to identify which models are ‘outliers’ or display a large difference from the more common and similar examples within a model suite. This similarity defines a ‘barycentre’ of the model suite in terms of the metric under examination. Further, and perhaps more interestingly, the models that display a large degree of difference from the barycentre can also be identified identifying model suite ‘end-members’.

This knowledge helps to define the limits of geological possibility given the input dataset, method of model calculation and geodiversity metrics employed.

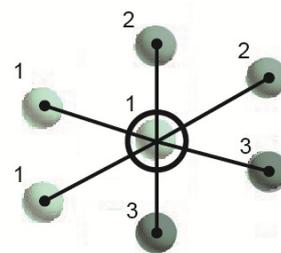
This study employs multiple geodiversity metrics. While expert opinion may hypothesise that co-variance exists between each geodiversity metric, it is required that rigorous analysis be performed to confirm this hypothesis. For example, the differences seen between models in terms of formation volume and formation depth are likely not to be related and their relationship should come under examination to better understand the characteristics of the model suite. Subsequently the problem has become multidimensional and to adequately analyse the complexities of the model suite multivariate analysis must be used. Principal Component Analysis (PCA) has been chosen to perform this task. PCA is a technique that allows complex data interactions to be displayed by orthogonal transformation of the data and re-organisation in terms of relevance to the attribute being analysed. The original, potentially correlated variables (in this case the geodiversity metrics) are converted into uncorrelated variables or principal components. The conversion of data is performed so that the first principal component displays the greatest variance, with each component thereafter displaying progressively lower degrees of variance. The ultimate aim is to contain as the highest degree of variance within the first principal component. This means that each component contains a combination of variability across all the metrics, rather than just measuring the variability of just one. This way the combined effect on variability of all the metrics can be measured, and also allows metrics that use different unit of measurement to be included. Each further component contains the next highest degree of remaining variance, so long as it is uncorrelated to preceding components (Jolliffe, 2002). By combining metric variability into principal components a common cause (or causes) behind that variability can be identified.

The principal components are calculated in the following order: i) statistics (mean, subtract deviations from mean, covariance matrix); ii) sorted eigenvectors and eigenvalues of the covariance matrix in descending order; iii) contributions of eigenvectors to eigenvalues; iv) determination of basis vectors; v) projection of z-score-converted original dataset onto basis vectors. PCA has been chosen instead of other multivariate techniques such as Factor Analysis or Nonnegative matrix Factorization due the larger range visualisation methods that are possible. PCA was performed in MATLAB, primarily with the ‘princomp’ function. The coefficients, or ‘loadings’, of the linear combinations of the metrics that were using to calculate the principal component data are obtained from this function (Jolliffe, 2002). Plotting loadings as vectors can show the contribution in variability a particular metric has toward the principal components (see Figures 5 and 6). Hotelling’s  $T^2$  statistic is also determined from the MATLAB ‘princomp’ function. It allows the determination of the multivariate distance of each model from the centre of the dataset, aiding identification of model suite outliers and barycentre examples (Hotelling, 1931; Krzanowski, 1995).

A two-stage PCA method is employed. The first stage determines which individual formations describe variability within each metric. Each stratigraphic unit is analysed for the formation depth and volume metrics and the short-distance neighbourhood relationships. The stratigraphic unit from each metric best describing observed variability within the model suite is identified by plotting the loadings. Those contributing the most to the first two principal components are retained for the second stage. Each metric representative identified in the first stage is then collated into a combined matrix with other geodiversity metrics compared with the others in a second stage. For example, if lithology  $X$  and lithology  $Y$  were seen to contribute most to the first principal component and second component variability respectively, both would be combined with other representative lithologies (say lithology  $A$  and  $B$  from the formation depth geodiversity metric) into a single matrix that combines all metrics together. By performing PCA on the combined matrix it is possible to 1) determine which metric best describes model suite variability overall and 2) which models represent the outliers and barycentre of the model suite.

#### 5. GIPPSLAND BASIN AND ASHANTI GREENSTONE BELT, SOUTHWESTERN GHANA

Two geological data sets have been used to build case study models representing two different geological terranes (off-shore basin versus craton) to discover whether each data set produces different results from the PCA. The first data set represents the Gippsland Basin, southeastern Australia. A variety of data types have been used when building this model



**Figure 2.** A complex short-distance neighbour relationships. Note the voxel under study, represented by the circled sphere (unit ‘1’) is surrounded by three different stratigraphic units (units ‘1’, ‘2’ and ‘3’).

including hard rock observations, geophysical interpretation (2D seismic, aeromagnetic and gravity data), bathymetry information and well logs. The second data set represents the southern part of the Ashanti Greenstone Belt located in southwestern Ghana. The input data types are similar, except seismic and bathymetry information have been replaced with radiometric and digital elevation model (DEM) data respectively.

The Mesozoic to Cenozoic Gippsland Basin is a mature oil and gas field located in southeastern Australia that also hosts brown coal deposits and is prospective for CO<sub>2</sub> sequestration (Cook, 2006; Rahmanian *et al.*, 1990) (Figure 3). Ordovician formations comprise the basement of the model and the Oligocene to Pliocene Seaspray and Angler comprise the cover sequences. The Paleocene to Late Miocene Latrobe Group is primary target for oil and gas and includes the Cobia, Golden Beach and Emperor Subgroups (Bernecker *et al.*, 2001). The basin is cross-cut by a number of transfer and normal faults, with the model bounded by the Cape Howe Fault in the west, the Lake Wellington and Combienbar faults in the north and the Cape Everard Fault in the south. The initial assumption was made that either the volume or depth of formation metrics would have the greatest influence on model suite variability as the model has layer-cake stratigraphy typical of a basin.

Granitoid intrusions, the Birimian Supergroup mafic volcanics and sediments and the Tarkwaian Group sediments comprise the basement of western Ghana (Figure 4). The Birimian Supergroup comprises alternating ‘belts’ of northeast striking mafic volcanics and ‘basins’ of shales, cherts and turbiditic sediments separated by major faults (Leube *et al.*, 1990). Fault-bounded Tarkwaian rocks appear as slices adjacent to or unconformably overlying Birimian volcanics (Davis *et al.*, 1994). Gold prospective regions include the contact between the Palaeoproterozoic Birimian Supergroup and Tarkwaian Group (Allibone *et al.*, 2002; Klemd *et al.*, 1993). In contrast to the Gippsland basin model, the assumption was made that the average mean curvature metric would have the greatest influence on model suite variability given that the model represents a more heavily deformed cratonic geological terrane.

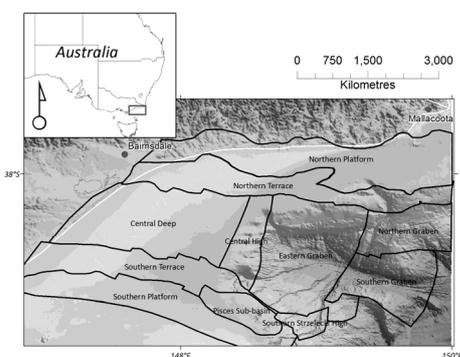


Figure 3. Location of the Gippsland Basin study area.

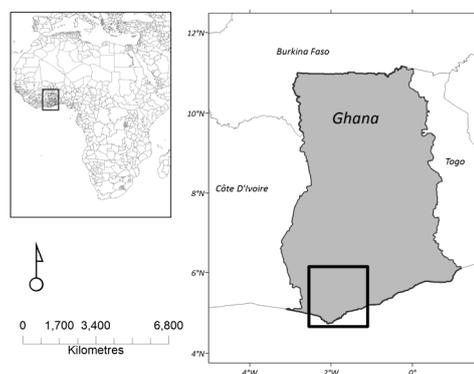


Figure 4. Location of the Ashanti Greenstone Belt study area.

## 6. RESULTS

Figure 5a and 5b show loading and score plots for short-distance neighbourhood relationships representing adjacent stratigraphic complexity. Each vector represents a stratigraphic unit. The direction and length of the vector indicates how each stratigraphic unit contributes to the two principal components in the plot. For example, Figure 5a shows that stratigraphic units fourteen, ten and fifteen contribute the most to the first principal component (x axis), whereas thirteen contributes the most toward the second principal component (y axis). Figure 5b shows that variability between models is best

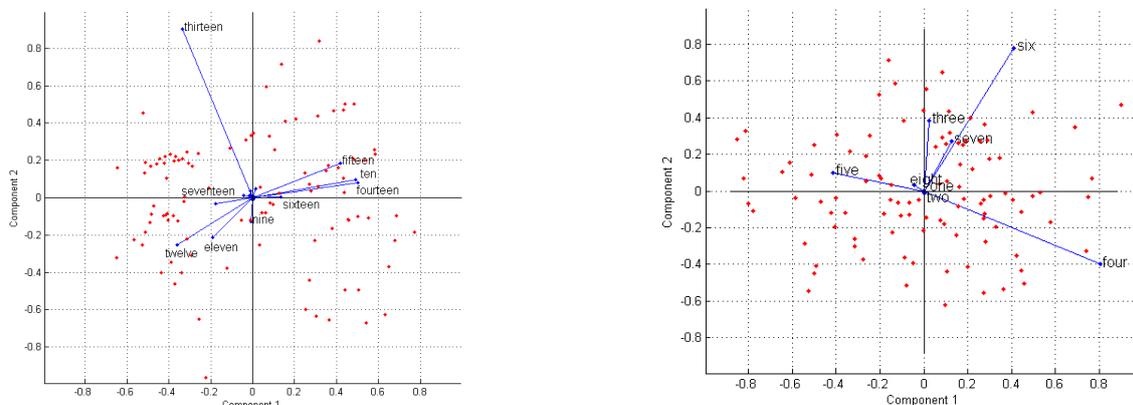
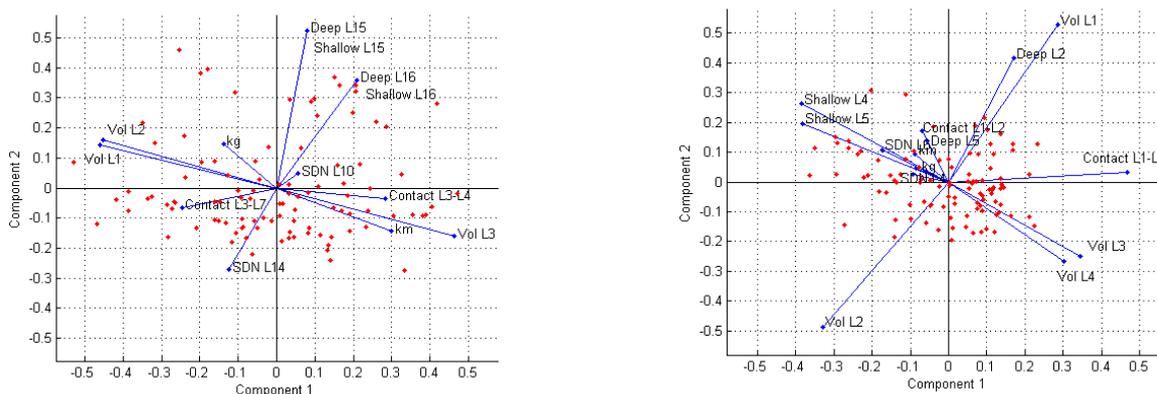


Figure 5. Loading (vectors) and score plots showing contributions to model suite variability: SD neighbour relationships representing adjacent stratigraphic complexity. (a) Gippsland Basin model suite study shows that for this geodiversity metric stratigraphic units fourteen, ten and fifteen plot along the x axis, indicating they are influential in terms of describing model suite variability. Unit thirteen is less influential as it plots along the y axis. (b) Ashanti Greenstone Belt model suite study shows that unit four is influential and six also, but to a lesser degree. Points represent a model and their distance from 0,0 represents distance from the barycentre.



**Figure 6.** Loading and score plots showing the contribution to model suite geodiversity for each metric defined in Section 3. The vector labels indicate the metric and the selected unit i.e. (a) shows that ‘Vol L3’ (unit three volume) is the most useful in terms of describing model variability for the Gippsland Basin model suite as it plots along the x axis, whereas (b) shows that ‘Contact L1-L3’ (the surface area of the contact between units one and three) best explains the variability of the Ashanti Greenstone Belt model suite. Points represent a model and their distance from 0,0 represents distance from the barycentre. Additional metrics are - Deep L’X’: deepest occurrence of unit X; Shallow: shallowest occurrence of unit; SDN: Short-distance neighbour adjacent stratigraphic complexity; km: mean average curvature; kg: mean Gaussian curvature (see Lisle and Toimil, 2007 for detailed description of curvature metrics). Points represent a model and their distance from 0,0 represents distance from the barycentre.

explained by stratigraphic unit four (first principal component) and, to a lesser degree, stratigraphic unit six (second principal component). These units are then selected and combined with units from other geodiversity metrics for the second stage combined PCA to determine which models represent the barycentre of the model suite (models represented by points close to 0,0) and those that represent the outliers (points plotted towards the extremities of the chart). Figure 6 shows the results from a combined geodiversity metric PCA study. All geodiversity metrics defined in Section 3 have been input into calculations in order to determine which models represent ‘end-members’ or outlier representations of the model suite and those that represent the barycentre. A comparison of the end-member models and the barycentre models is shown in Table 2.

### 7. DISCUSSION AND CONCLUSIONS

Most of the metrics used in this study could be of some use to various geoscientific studies. The method described can illuminate otherwise difficult to obtain geometrical information across multiple models and allow comparison with other model suites. End-member and barycentre models have been identified for both Gippsland Basin and Ashanti Greenstone Belt model suites. Table 2 reveals that the SD neighbour adjacent stratigraphic complexity metric PCA shows different

end-member and barycentre models to the combined PCA. In addition Figure 6a shows that the most influential metric for the Gippsland Basin is the volume of stratigraphic unit three, while the contact surface area for the contact between units one and three is the most influential for the Ashanti model suite. That the volume of unit three (top of the Ordovician basement) in the Gippsland Basin model is the most influential has implications for model uncertainty. As unit three underlies most of the stratigraphy in a basin environment, any geometrical perturbations of this unit will have subsequent effect on the volume of

the formation, which in turn could affect the spatial location of the the overlying units, including the oil and gas prospective Latrobe Group units. The contact between the units one and three for the Ashanti model suite are also important. This is the contact between the Early Birimian and the base of the Tarkwaian basin. This relationship contains an unconformity and defines the geometry of the Tarkwaian basin as whole, but also is a gold prospective geological contact. Any geometrical perturbations to this contact will have a large impact on variability between models in the suite, as identified by PCA, but also would be of interest to mineral explorers in the region.

These results can aid further revisions of the model. Measurements, observations and data types that are effective at analysing the volume of Gippsland Basin stratigraphic unit three (for example seismic tomography or sections) or the contact between units one and three in the Ashanti model (drill core or aeromagnetic interpretation) can be pursued. If these aspects of the model can be better constrained, the possibility of high variability between models may be less, which sub-

**Table 2.** Comparison of top three end-member models and top three barycentre models calculated from the Gippsland Basin and Ashanti Greenstone Belt model suites ( $n = 101$ ) in terms of SD neighbour adjacent stratigraphic complexity and the combined geodiversity metric results are compared. Hotelling’s  $T^2$  was used to determine these rankings.

	SD neighbour adjacent stratigraphic complexity		Combined geodiversity metrics	
	Gippsland Basin	Ashanti GS Belt	Gippsland Basin	Ashanti GS Belt
End-member models	Model 2	Model 61	Model 10	Model 13
	Model 16	Model 52	Model 30	Model 59
	Model 66	Model 78	Model 38	Model 82
Barycentre models	Model 7	Model 96	Model 20	Model 84
	Model 37	Model 50	Model 61	Model 50
	Model 75	Model 40	Model 95	Model 1

sequently reduces overall model uncertainty. The different geodiversity metrics identified between model suites also suggests that using a number of metrics is appropriate to encompass the variety of model geometries presented when a data set has subjected to uncertainty simulation. The hypothesis that a universal metric exists can be rejected in terms of the data sets and metrics used in this study, but the concept of a universal metric has not been completely examined, as many more exist than those analysed here.

Model suites can be analysed by PCA to quickly determine end-member and barycentre models for both individual geodiversity metrics and in combination. The relative contribution of different stratigraphic units within individual metrics can be assessed to find their contribution to model suite variability which can potentially answer questions relating to the specific aims of a geoscientific study. The location of areas of interest can be easily located using these metrics as the data is georeferenced within the voxel. Different metrics can also be identified as contributing more or less to model suite variability. Additional metrics can be easily added into combined PCA analyses as they are developed. This information can allow the operator to choose which metrics to utilise more effectively inspect the breadth of geological possibility.

## ACKNOWLEDGMENTS

Thanks go to Geoscience Victoria, Department of Primary Industries, Victoria, Australia; the Society of Economic Geologists Hugo Dummett Memorial Fund; Intrepid Geophysics for technical assistance and access to the 3D Geomodeller API and Golden Star Resources Ltd for Ashanti Greenstone Belt data sets.

## REFERENCES

- Allibone, A. H., McCuaig, T. C., Harris, D., Etheridge, M. A., Munroe, S., Bryne, D., Amanor, J., and Gyapong, W., (2002). Structural controls on gold mineralization at the Ashanti gold deposit, Obuasi, Ghana. In Goldfarb, R. J., and Neilson, R. L., eds., *Integrated Methods for Discovery: Global Exploration in the 21st Century*, Society of Economic Geologists, 65-93.
- Bernecker, T., Woollands, M., Wong, D., Moore, D., and Smith, M., (2001). Hydrocarbon prospectivity of the deep water Gippsland Basin, Victoria, Australia, *APPEA Journal*, 41, 91-113.
- Bremner, D., Demaine, E., Erickson, J., Iacono, J., Langerman, S., Morin, P., and Toussaint, G., (2005). Output-Sensitive Algorithms for Computing Nearest-Neighbour Decision Boundaries, *Discrete & Computational Geometry*, 33(4), 593-604.
- Calcagno, P., Chilès, J. P., Courrioux, G., and Guillen, A., (2008). Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules, *Physics of the Earth and Planetary Interiors*, 171(1-4), 147-157.
- Cook, P. J., (2006). Carbon dioxide capture and geological storage: research, development and application in Australia, *International Journal of Environmental Studies*, 63(6), 731 - 749.
- Davis, D. W., Hirdes, W., Schaltegger, U., and Nunoo, E. A., (1994). U-Pb age constraints on deposition and provenance of Birimian and gold-bearing Tarkwaian sediments in Ghana, West Africa, *Precambrian Research*, 67, 89-107.
- Friedman, J. H., Bentley, J., and Finkel, R. A., (1977). An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software*, 3(209).
- Hotelling, H., (1931). The Generalization of Student's Ratio, *The Annals of Mathematical Statistics*, 2(3), 360-378.
- Jessell, M. W., Ailleres, L., and 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*, second edition, Springer.
- Klemm, R., Hirdes, W., Olesch, M., and Oberthür, T., (1993). Fluid inclusions in quartz-pebbles of the gold-bearing Tarkwaian conglomerates of Ghana as guides to their provenance area, *Mineralium Deposita*, 28(5), 334-343.
- Krzanowski, W. J., (1995). Selection of variables, and assessment of their performance, in mixed-variable discriminant analysis, *Computational Statistics and Data Analysis*, 19(4), 419-431.
- Lajaunie, C., Courrioux, G., and Manuel, L., (1997). Foliation fields and 3D cartography in geology: Principles of a method based on potential interpolation, *Mathematical Geology*, 29(4), 571-584.
- Leube, A., Hirdes, W., Mauer, R., and Kesse, G. O., (1990). The Early Proterozoic Birimian Supergroup of Ghana and some aspects of its associated gold mineralization, *Precambrian Research*, 46, 139-165.
- Lindsay, M., Ailleres, L., Jessell, M. W., De Kemp, E. A., and Betts, P., Integrating geological uncertainty into combined geological and potential field inversions., in *Proceedings GeoMod 2010*, University of Lisbon, Lisbon, Portugal, 27-29 September 2010.
- Lisle, R. J., and Toimil, N. C., (2007). Defining folds on three-dimensional surfaces. *Geology*, 35(6), 519.
- Rahmanian, V. D., Moore, P. S., Mudge, W. J., and Spring, D. E., (1990). Sequence stratigraphy and the habitat of hydrocarbons, Gippsland Basin, Australia. *Geological Society, London, Special Publications*, 50(1), 525-544.
- Thore, P., Shtuka, A., Lecour, M., Ait-Ettajer, T., and Cognot, R., (2002). Structural uncertainties: Determination, management, and applications. *Geophysics*, 67(3), 840-852.
- Wellmann, F. J., Horowitz, F. G., Schill, E., and Regenauer-Lieb, K., (2010). Towards incorporating uncertainty of structural data in 3D geological inversion, *Tectonophysics*, 490(3-4), 141-151.
- Wellmann, J. F., and Regenauer-Lieb, K., (2011). Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models, *Tectonophysics*.