ASEG-PESA 2013
'THE EUREKA MOMENT'
11-14 AUGUST 2013 • MELBOURNE, AUSTRALIA

Unsupervised soft clustering of high resolution airborne geophysical and satellite data suites from the Sperrgebiet, Karas region, Southern Namibia, to enhance lithology mapping

D.G. Eberle
Council for Geoscience
Private Bag X112
Pretoria, South Africa
deberle@geoscience.org.za

D. Hutchins
Geological Survey of Namibia
1 Aviation Rd.
Windhoek, Namibia
dhutchins@mme.gov.na

N. Titus
Geological Survey of Namibia
1 Aviation Rd.
Windhoek, Namibia
ntitus@mme.gov.na

SUMMARY

Conventional geological mapping using aerial photo interpretation coupled with field visits is often handicapped in areas with tropical weathering or sand cover in arid regions. High resolution airborne geophysical data acquisition over large areas covered by overburden can augment geological mapping of these areas provided information extraction is done in an adequate, automated and objective way. We suggest that the integration of airborne geophysical data with space-borne remote sensing data suites using fuzzy partitioning clustering meets these criteria.

The outcome of unsupervised clustering is a classified zonal map which, in combination with some field inspections, produces a high resolution lithology map, in this case 200 m by 200 m pixel size, which is by power of spatial resolution far superior to any conventional mapping. It can therefore provide new incentive for geological understanding of the area, modelling or mineral exploration.

Our new approach to map hard rock geology hidden beneath overburden has been applied to a portion of largely sand and sediment covered southern Namibia, the so-called Sperrgebiet, by integrating airborne magnetic, radiometric and Landsat 7 data suites. The classified zonal maps obtained from these data were converted to a (pseudo) lithology map by attributing what is known from previous geological investigations of this area.

Key words: Data integration, clustering, mapping, airborne geophysics, satellite imagery.

INTRODUCTION

Sand cover in arid regions and thick weathering in the tropics have always posed an obstacle to conventional mapping where the use of aerial photographs in combination with field visits is standard practise. In recent years, however, with the availability of large multi-parameter high resolution earth science related data sets acquired from space or aircraft, the upgrading of conventionally mapped geology is now feasible. Contrary to space-borne and hyper-spectral imagery, airborne geophysics has the ability to penetrate tropical weathering or extensive sand layers lying on hidden bedrock. Furthermore, airborne geophysical data, such as magnetic, gravity or conductivity data, provide information not only from the surface but also from the sub-surface thus enabling geological mapping at selected depth levels.

We therefore suggest a new method based on clustering and integration of high density airborne geophysical data with remote sensing data to facilitate high resolution mapping over areas with extensive overburden. We discuss the method using the results obtained from an airborne geophysical interpretation project recently conducted with data from sand and sediment covered southern Namibia.

The use of multivariate statistical methods in geophysical data interpretation is not new. First suggestions to use decision theoretical and Bayesian techniques for the integrated interpretation of magnetic and gravity data were published as early as 1972 by J. Harff. With the advent of powerful computers, case studies of crisp and fuzzy clustering have increasingly been carried out over smaller selected areas for mineral exploration, hydro- and engineering geology (e.g. Eberle, 1993, Paasche and Tronicke, 2007).

Here, we discuss results achieved from clustering high resolution data sets from the former Diamond Sperrgebiet of southwestern Namibia extending along the South Atlantic Coast over a distance of more than 200 km and 100 km inland. The area is largely covered by sands of the Namib desert camouflaging hard bedrock. Geologically the area is part of the late Proterozoic Gariep Belt (Frimmel, 2008). Our study results in the first ever complete high resolution lithology map of the Sperrgebiet with a pixel size of 200 m by 200 m revealing incentives to re-consider the geology and mineral potential of the Gariep Belt.

METHOD

Airborne geophysical data are integrated with hyper-spectral data and satellite imagery using fuzzy partitioning clustering algorithms (Paasche and Eberle, 2009, 2011). The integration process is based on fuzzy clustering, which is widely applied in natural sciences, particularly, in medical and biological sciences, but rarely proactively used in the Earth sciences. Clustering requires grid data with identical grid spacing and grid points for all the data sets to be analysed. Each grid point is characterized by a multi-element sample vector containing, for example, electrical conductivity, gravity or reflectance variables. Some data pre-processing is required (Paasche and Eberle, 2009). Ideally, all data sets should have a bell-shaped

Gaussian distribution curve. The data sets are normalized to mean zero and unit standard deviation. The correlation matrix is computed to study the inter-relationship of the data sets. For example, it is known from radiometric data sets that they contain abundant information and are therefore highly correlating with one another. To avoid that data sets which are abundant may skew the clustering result we check the correlation amongs the data sets removing one of two if they are highly correlating. We have noted that optimum clustering results are obtained when data sets with minimum correlation are used (Eberle and Paasche, 2012). Data are then submitted to clustering as shown in the flow diagram (Figure 1).

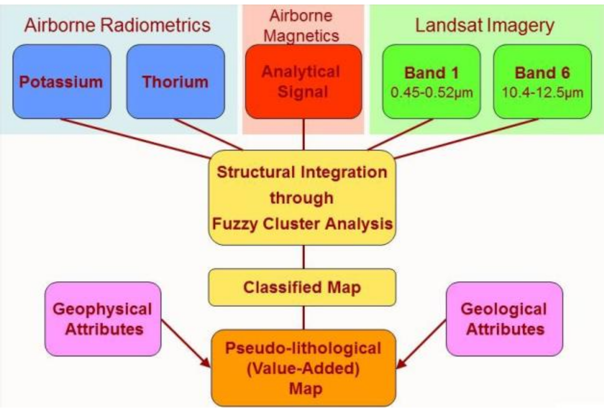


Figure 1. Flow diagram.

Through clustering, all sample vectors are grouped into a pre-determined number of classes. The principle of clustering is to have class centres as far as possible from one another in the multi-dimensional space, with individual samples as close as possible to one of the class centres. Fuzzy clustering assigns each sample to all classes and the soundness of the assignment is provided by the value of statistical probability, or degree of membership.

To obtain the classified zonal map (Paasche and Eberle, 2009, 2011) from the clustering result, each class is assigned to a specific colour and each pixel centred around its grid point is printed using only the colour of the class that has the highest degree of membership at this specific point. The higher the degree of membership, the more saturated is the colour of the pixel.

RESULTS

High resolution airborne geophysical data (magnetics and radiometrics) and Landsat 7 satellite imagery were made available for this study (Figure 2). All the data sets were projected onto a grid with a spacing of 200 m. After data inspection and pre-processing we submitted the analytical signal of the magnetic field (Figure 2a), the calculated surface concentrations of Potassium and Thorium and Landsat 7, bands 1 and 6 (Figures 2c and d), to fuzzy clustering. The selection of these variables was done based on the inspection of the 2D correlation matrix which confirms low correlation within the data sets (Figure 3).

Classified zonal maps for four to nine classes were successively computed. Evaluating statistical quality parameters, such as the Xie-Beni index or Normalized Classification Entropy (Paasche and Eberle, 2011), indicated that the 9-class case was the optimum classification output as displayed by Figure 4.

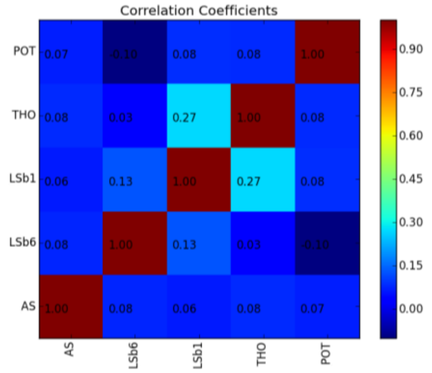


Figure 3. 2D-correlation matrix. AS: Analytical signal, POT: potassium, THO: thorium, LSb1: Landsat band 1, LSb6: Landsat band 6.

At this point, all nine classes are derived from unsupervised data-driven classification and have no geological counterparts. The next step is to assign the classes in relation to what is known from previous investigations including any existing mapped geology. Human skills are required to upgrade the classified map to a more user-friendly, value-added lithology map. Each class is correlated with as many geological and geophysical attributes as are known. In some cases ground truthing will be necessary to identify the petrographic expression of a specific class. Table 1 summarizes the subjective effort to attribute the eight classes to airborne geophysical features, satellite imagery patterns and the mapped geology.

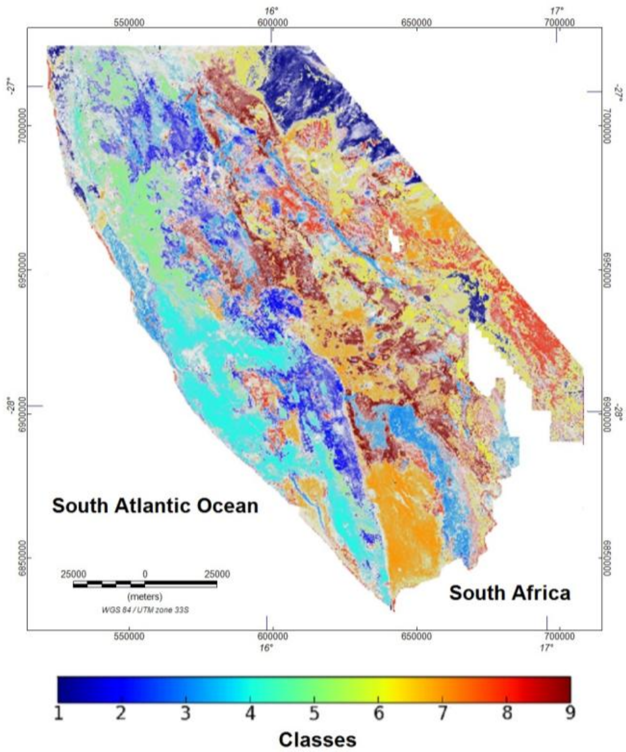


Figure 4. Classified zonal map as result of fuzzy clustering a data suite consisting of three airborne geophysical and two satellite imagery data sets.

The fact that airborne magnetic data which 'look through' ground surface have been included for clustering partially explains why the classified zonal map cannot be expected to be congruent with the mapped geology of this area. However, the high resolution (pixel size is 200 m by 200 m) of the map provides high resolution geologic mapping and identifies anomalous areas where a re-visiting mapping geologist can concentrate on very local specific sites.

Class No.	Attributes as per mapped geology (GSN, 1999) and remote sensing
1	Along NW-coast: Magnetic member of Port Nolloth Zone, in the extreme NE corner of area: High LSb1 reflectance; in the SE: Possibly magnetic member of Numees Fmt.
2	Occurring between classes 4 and 5 in the west and class 9 in the east, in the NW supported by magnetic response, in the SE by increased surface thorium concentarations. No obvious geological equivalent.
3	Schakalsberge Subterrain: Chameis Group, Metabasalts of the Grootderm Formation with distinct compact magnetic signature,
4	Namib sands with elevated thorium surface concentrations of > 20 ppm
5	Bloeddrif member of the Port Nolloth Group
6	Undifferentiated basement cover in the eastern part of the study area
7	Namib sands with elevated potassium surface concentrations > 1.5 %
8	Nama Group: Shale, limestone, quartzite
9	Surficial cover enriched in thorium, mostly over the Port Nolloth Zone

Table 1. Linking geological and geophysical attributes to objectively identified classes displayed in Figure 4.

To study the effect of removing the magnetic analytical signal from the data suite we integrated the airborne radiometric surface concentrations of potassium and thorium with the Landsat 7 bands 1 and 6 for clustering. These data contain only near surface information and the resulting classified zonal map allows immediate comparison with the mapped geology of the Sperrgebiet area. Figure 5 displays the classified zonal map obtained from clustering of two radiometric and two Landsat 7 data sets.

We recognize that the shape of the nine classes as displayed by Figure 5 appears more compact and solid if compared with the previous clustering result shown in Figure 4. The classified zonal map of Figure 5 provides a complete high resolution image with regard to the spatial distribution of nine units with different physical properties at the surface.

The classified zonal map of Figure 5 is ideal for comparison with the mapped geology (GSN, 1999; Frimmel, 2008) which is provided in Figure 6. Green tints covering most of the study area reflect undifferentiated overburden, such as dune sands, calcrete or gravel. Pinkish tints along the coast of the South Atlantic Ocean represent magnetic members of the Port Nolloth Zone of the Gariep orogeny, possibly intercalated in the Holgat Formation. Petunia tints in the east and southeast indicate outcrop of the Marmora Terrane and Port Nolloth Zone.

Using the mapped geology (Figure 6) together with geophysical characteristics identified from Figures 2a and 2b and specific patterns from satellite imagery (Figures 2c and 2d) we undertake the subjective effort to link the classes with earth science related attributes. The result is shown in Table 2.

Class No.	Attributes as per mapped geology (GSN, 1999) and remote sensing
1	Along NW-coast: Magnetic member of Port Nolloth Zone, in the extreme N: magnetic Namaqua intrusive rock
2	South: Namib sands or detritus containing < 1.75 % K Central-east: Outcrop of Namaqua Belt rock rich in potassium (> 2 %)
3	Autochtonous overburden (colluvium), mostly of the Port Nolloth Zone
4	Namib sands with elevated thorium surface concentrations of > 25 ppm
5	Nama Group: Shales, limestone, sandstone
6	South: Magnetic member of the Dernburg Fmt., Marmora Terrane. North-central: Low reflectance in both bands, shallow magnetic response
7	Very strong LSb1 reflectance, shallow magnetic field
8	Namib dune sands with elevated average Th contents of 25 ppm and less
9	Cover slightly enriched in thorium (< 20 ppm), situated in the wind blown segment of the Port Nolloth Zone

Table 2. Linking geological and geophysical attributes to objectively identified classes displayed in Figure 5.

The Sperrgebiet geology map (Figure 6) was surely compiled from hundreds of observations on the ground collected by several generations of geologists over many years. Furthermore, the majority of site observations including descriptions in writing has been done mostly in the immediate vicinity of hard rock outcrop. The larger portion of the area, however, covered by overburden has been visited to a much lesser degree and the density of data acquisition is irregular.

In contrast, the classified zonal maps shown in Figures 4 and 5 have been compiled from high resolution data projected onto a regular grid with a spacing of 200 m by 200 m. All observation points have been equally sampled over the study area. This is the first time that the Sperrgebiet has been mapped with this kind of methodology which we believe is far superior to any regional map compiled from data collected on the ground.

It is also important to note that the computing time on a standard notebook was less than five minutes to obtain the classified zonal maps shown in Figures 4 and 5, and each map consists of 500 000 pixels. To classify 500 00 multi-element sample vectors with five variables, 2.5 million individual data values had to be analysed. Consider the time and number of mapping geologists that would be required on the ground to achieve a comparable high density of observations with this map quality.

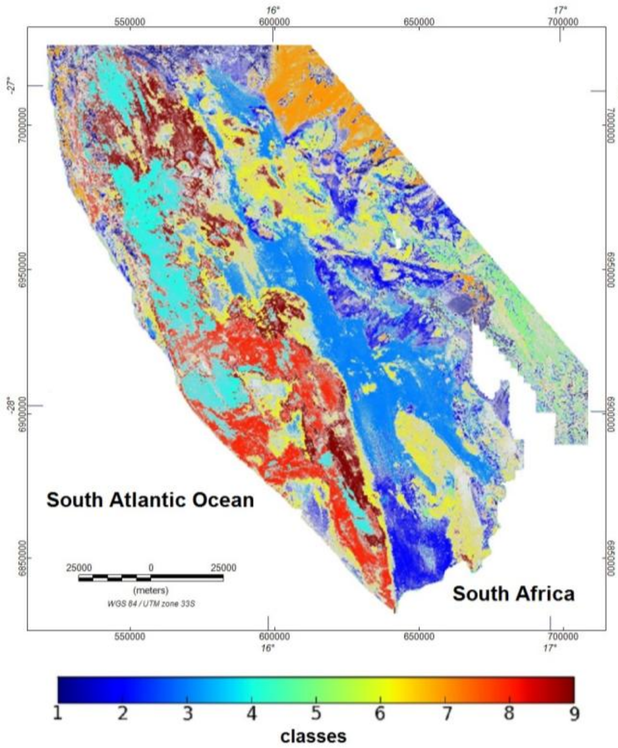


Figure 5. Classified zonal map as result of fuzzy clustering a data suite consisting of two airborne radiometric and two satellite imagery data sets.

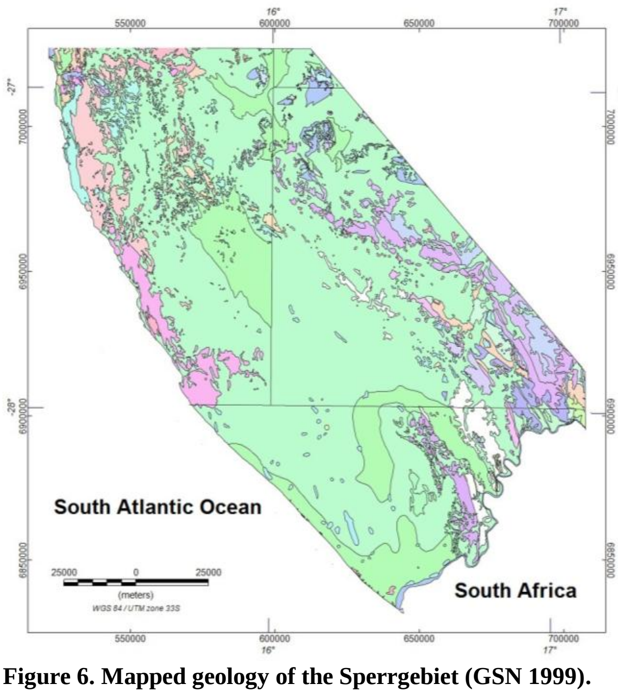


Figure 6. Mapped geology of the Sperrgebiet (GSN 1999).

CONCLUSIONS

The integration of high resolution airborne geophysical data with space-borne remote sensing data suites using unsupervised fuzzy partitioning clustering is a powerful tool to obtain classified zonal maps for lithology mapping of large areas.

Data driven clustering is a novel approach to support geological mapping in difficult terrain, such as over desert sands or areas with thick tropical weathering. The method is fully automated and objective though human skills are still required to determine the number of classes, choice of the cluster algorithm and attributing geological and geophysical properties to the classes. Explorative clustering as applied in the case study does not require any training data and objectively reveals the structure of the data suite.

The two zonal maps (Figures 4 and 5) should be assessed together to produce a more complete interpretation. Using the magnetic data provides valuable insight to the subsurface concealed geology (Figure 4), whilst the surface geology and overlying cover sediments are mapped in Figure 5.

ACKNOWLEDGMENTS

The authors want to express their gratitude to the Director of the Geological Survey of Namibia, Ministry of Mines and Energy, for making available the various airborne geophysical, remote sensing and geological data sets.

REFERENCES

Eberle, D., 1993, Geologic mapping based upon multivariate statistical analysis of airborne geophysical data: International Institute for Aerospace Survey and Earth Sciences (ITC) Journal, Special Issue 1993-2, 173-178.

Eberle, D. and Paasche, H., 2012, Integrated data analysis for mineral exploration: A case study of clustering satellite imagery, airborne gamma-ray, and regional geochemical data suites: Geophysics, 77, 4, B167-B176, doi: 10.1190/ GEO2011-0063.1

Frimmel, H.E., 2008, Neoproterozoic Gariep Orogen, *In*: Miller, R.McG., ed., *The Geology of Namibia, Vol. 2*, 14-1 – 14-39, Ministry of Mines and Energy, Windhoek.

Harff, J., 1972, Komplexinterpretation geologischer Informationen mit Klassifizierungsmethoden. *In*: Peschel, G., ed., *Beiträge zur Theorie der Komplexinterpretation geowissenschaftlicher Informationen*, ZGI Berlin.

Geological Survey Namibia (GSN), 1999, 1 : 250 000 Geological Map Series, Sheets 2716 Ai-Ais and 2816 Oranjemund, Ministry of Mines and Energy, Windhoek.

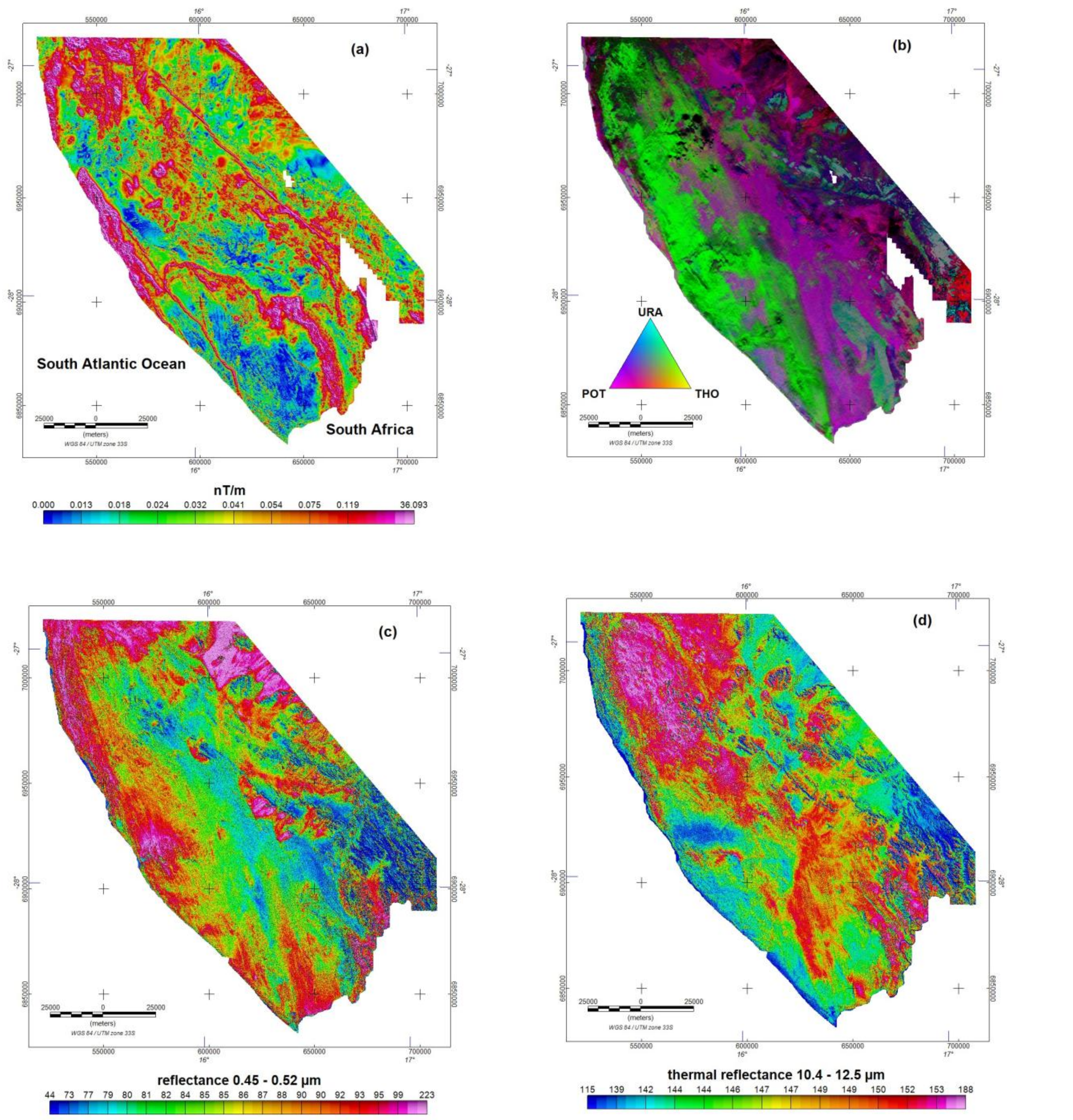
Paasche, H. and Eberle, D., 2009, Rapid integration of large airborne geophysical data suites using a fuzzy partitioning cluster algorithm: A tool for geological mapping and mineral exploration targeting: Exploration Geophysics, 40, 277-287, doi: 10.1071/EG0828.

Paasche, H. and Eberle, D., 2011, Automated compilation of pseudo-lithology maps from from geophysical data sets: A comparison of Gustafson-Kessel and fuzzy c-means cluster algorithms: Exploration Geophysics, 42, 275-285; doi: 10.1071/EG11014.

Paasche, H. and Tronicke, J., 2007, Cooperative inversion of 2D geophysical data sets: A zonal approach based on fuzzy c-means cluster analysis: Geophysics, 72, 3, A35-A39; doi:10.1190/1.2670341.

Note: figures or tables too large to be interspersed with text in column format should be placed here at the end of the paper.

The size limit for a paper in the conference volume is 10 MByte



.

Figure 2. High resolution airborne geophysical and satellite imagery data sets to enter fuzzy clustering. (a) Analytical signal of the magnetic field; (b) Ternary image of the radioelements 40K, Th and U; (c) Landsat 7, band 1; (d) Landsat 7, band 6.