



Using self-organizing maps to investigate dike swarms from airborne geophysical data in NE Brazil

Alanny Christiny Costa de Melo, Universidade Federal do Rio Grande do Norte (UFRN)
David Lopes de Castro, Universidade Federal do Rio Grande do Norte (UFRN)
Stephen James Fraser, Vectore Pty Ltd

Copyright 2019, SBGf - Sociedade Brasileira de Geofísica

This paper was prepared for presentation during the 16th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 19-22 August 2019.

Contents of this paper were reviewed by the Technical Committee of the 16th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

Abstract

Mafic dike swarms intruding in the Borborema Province is a part of a great magmatic event associated with the continental break-up that formed the Equatorial Atlantic. The Borborema Province is a complex mosaic of crustal blocks of different ages, origin and evolution, which was amalgamated during the Brasiliano Orogeny, at the end of the Neoproterozoic and beginning of the Phanerozoic, in the context of the construction of the supercontinent Gondwana. Analyzing a geological framework within a tectonically complex area is a multidisciplinary task, because traditional geological tools do not provide the desired answers. Airborne geophysical methods have been widely used due to the practicality and low cost of data acquisition to provide or improve geological mapping. This research presents a multivariate analysis of airborne geophysical data, using a Self-Organizing Map (SOM) approach. The visual interpretation of integrated products and the technique of spatial analysis SOM were used for the integration of magnetic and gamma-ray spectrometric data search to identify and separate the lithological units based on their geophysical signatures. The SOM investigation showed that the data could be clustered into seven domains. These results reveal the efficacy of using SOM as a tool for geophysical data analysis and for semi-automated mapping.

Introduction

Research in geosciences produces large amounts of geological, geochemical and geophysical data. However, such data are collected at a rate faster than it can be interpreted and integrated. The extraordinary volume of multidisciplinary data and the complexity of their interrelationships create a need for the use of methods and technologies to improve the efficiency and effectiveness of interpretation of these data. To fill this gap, we have used data-mining methods that work in hyperspace of multivariate parameters, one of these methods is the Self-Organizing Maps (SOM), which is an unsupervised technique for the integrated and automated analysis of multivariable data. This technique facilitates and accelerates the speed at which data can be analyzed and interpreted.

The study area is in the Precambrian Borborema Province (Fig.1). In a broader sense, the Borborema Province is a shield that consists of Archean to Proterozoic inliers amalgamated along volcano-sedimentary belts (Neves et al., 2000). This area was affected by a Cretaceous Large Igneous Province (LIP) formed during the Equatorial Atlantic Ocean opening and, therefore, related to the break-up of the West Gondwana supercontinent (Hollanda et al., 2019). Due to its geological complexity there is a large amount of data collected from this area, makes it favorable for the use of SOM. SOM provides means to analyze and interrelationships between distinct datasets, spatially localized in an interactive way.

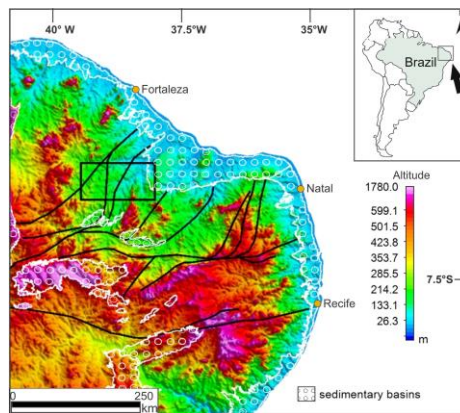


Figure 1. Schematic geological map of northeastern Brazil with sedimentary basins, major shear zones and the study area (polygon in black). Inset: The South American continent and location of study area.

SOM is a tool for analysis and visualization of n-dimensional data, based on principles of vector quantization and measures of vector similarity (Kohonen, 2001), which was proposed and disseminated in the 1980's by Finnish professor Teuvo Kohonen (Kohonen and Somervuo, 1998). Therefore, the SOM based techniques are also referred as Kohonen maps.

Based on the principles of vector quantization and measures of vector similarity (Fraser and Dickson, 2007), the SOM approach aims to extract information from the database by recognition of samples with similar characteristics and to find association between the variables through the formation of clusters.

The basic idea behind a SOM is to map the n-dimensional data patterns on to and 2D grid. This mapping attempts to preserve topological relationships. It means patterns that are close to the input space will be mapped to be close in the output space, and vice versa (Bação et al., 2005). The output data are a 2D map on a complex n-dimensional data set.

The great advantage is that SOM can be applied to both categorical and continuous variables. One limitation of the SOM is that it scales all the data to the range [0,1], it is not possible to set weights for the importance of different datasets, which would sometimes be useful (Torppa et al., 2015).

Data and Methods

The geophysical data were provided by the Brazilian Geological Survey (CPRM). The magnetic data were previously corrected for the diurnal variations and the International Geomagnetic Reference Field (IGRF) and leveling errors. The datasets were collected in 2010 (Tab. 1). All the geophysical processing (Fig.2) was performed using software Oasis Montaj 9.3.

Due to the dipole character of the magnetic field we applied the Reduction to the Pole filter (RTP), aimed to centralize the magnetic anomaly on its sources (Blakely, 1995). In addition, we applied the Analytical Signal Amplitude filter (ASA) to the RTP anomalies to produce a maximum on the magnetic contrasts, highlighting the edges of the magmatic bodies. The ASA anomaly is symmetrical and occurs directly over the entire length of the edges of the bodies and the center of the boundary (Nabighian, 1972).

Five geologically significant geophysical parameters were then selected for further analysis as input to the SOM procedure: RTP, ASA, K, eTh and eU.

Table 1. Airborne geophysical survey specifications.

nominal flight height (m)	150
flight-line direction	N-S
flight-line spacing (m)	2000
tie-line direction	E-W
tie-line spacing	20,000
magnetic field strength (nT)	25,670
magnetic field inclination	- 5.85°
magnetic field declination	- 8.78°
Interval between consecutive geophysical measurements	0.1 s (magnetometer) 1.0 s (spectrometer)

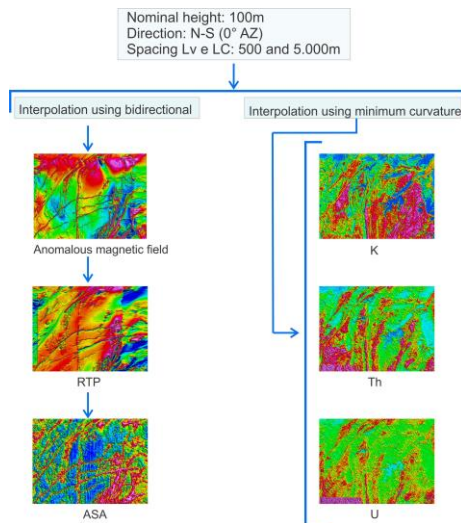


Figure 2. Workflow of magnetic (left) and gamma ray (right) processing. RTP – Reduction to The Pole, and ASA – Analytical Signal Amplitude filters.

The original dataset consisted of 294696 samples with six magnetic and gamma spectrometric variables (Fig. 2). A set of 'xy' files were created for both the magnetic and gamma spectrometric datasets and the datasets were then combined to make one CSV file where each point in space within the dataset had a data value for each data. Due to the magnitude of the analytic signal, logarithmic values were used to produce more definition in the analysis.

The input control user for SOM program are: hexagonal shape, with a toroid geometry and size of 70x50 to the SOM grid. In each case, SOM initialization was random, the shape was toroid, the grid was hexagonal, and the neighborhood function was Gaussian.

Self-Organizing Maps Analysis

We used the SOM implementation developed by the Commonwealth Scientific and Industrial Research Organization (CSIRO, Australia) to create a 2D SOM grid.

The technique is unsupervised, so no prior knowledge about the nature or number of clusters within the data is required, and analysis is based on measures of vector similarity. The SOM classifies data into distinct groups (or clusters), according to their similarities, which are represented as single vectors the BMUs (best-machine units) and represent nodes on the map (Fraser and Dickson, 2007).

The SOM nodes (values) are further clustered by a K-means clustering based on a pre-determined number of classes defined by a Davies-Bouldin (DB) analysis (Davies and Bouldin, 1979).

The SOM automatically calculates, for each input field data, the quantization error (QER). This error is a measure of the distance that a sample is from its node-vector (Fraser and Dickson, 2007). The greatest QE values represent anomalous data vectors. The high QER samples can also represent the edges or boundaries within a dataset, which for geology can be interpreted as lithological contacts or body borders (Bação et al., 2005).

Results

A Self-Organized Map analysis, using a 70x50 sized map, was carried out on the five input geophysical parameters. The U-Matrix resulting from this analysis is shown in Figure 3A. The U-Matrix shows three areas with high dissimilarity, which coincide with elevated contributions of RTP and ASA. High similarity in U-matrix is related with mediums contributions of K, eTh, eU.

To obtain the optimized number of clusters, we applied a Davies-Bouldin analysis (Davies and Bouldin 1979) and 7 clusters were chosen. Using K-means the 7 clusters were then produced (Figure 3B), based on the similarity between the node vectors (hexagons). Figure 3C shows the spatial distribution of input samples colour-code by their SOM-derived cluster number. The reference values derived from the total data set, to classify each SOM-derived cluster's component contribution, as shown in Table 2.

As a result, the sets of magnetic and gamma-ray data were separated into seven different clusters, and those associated to the magnetic dikes were presented in a map (Fig. 3C). According to their geophysics characteristics, the dikes swarms are assigned to clusters 1 and 6. The choice of these clusters was defined by the anomalous pattern expected for dikes (high magnetic amplitudes and short wavelength). These clusters also presented the highest values of QER.

The dikes interpreted from the SOM process were grouped into 3 different dikes swarms, Rio Ceará-Mirim (RCM), Riacho do Cordeiro (RC) and Canindé (C).

Possible contamination of SOM solutions by high frequency noise along the flight lines were duly considered in the interpretation of the results.

Validation of the Results

The results of this study were validated by field work of the above-mentioned techniques was done through the in an area preselected within the study area. We collected 66 rock samples of RCM dykes along the most expressive magnetic anomalies, with no visible weathering process, and measured their magnetic susceptibilities (MS).

Commented [DC1]: Esta sentença está muito confusa. Reescrever.
Falta mais uma sentença no final do parágrafo, informando qual o grau de confirmação dos resultados do SOM com os dados de campo.

Commented [DC2]: Seria interessante incluir um mapa geológico da área com os diques mapeados antes e as soluções do SOM sobrepostas.

Table 2 Influence of variables in each cluster from SOM analyses.

	Erro Q	RTP (nT)	ASA (nT)	K (%)	eTh (ppm)	eU (ppm)
Cluster 1	0.8470 (High)	-2.8398 (Low)	0.6374 (High)	3.3678 (High)	12.2782 (Medium)	1.5690 (Low)
Cluster 2	0.3735 (Low)	-46.2008 (Low)	0.0774 (Medium)	4.0106 (High)	11.7725 (Medium)	1.4937 (Low)
Cluster 3	0.3068 (Low)	-43.2815 (Low)	0.0754 (Medium)	1.3767 (Medium)	5.7564 (Low)	0.8070 (Low)
Cluster 4	0.4057 (Low)	3.7730 (Low)	0.0533 (Low)	2.6989 (Medium)	18.7306 (Medium)	2.0652 (Medium)
Cluster 5	0.3626 (Low)	68.3386 (Medium)	0.0866 (Medium)	1.9257 (Medium)	7.7591 (Low)	0.9813 (Low)
Cluster 6	1.1344 (High)	251.5091 (High)	1.0214 (High)	3.2211 (High)	11.9387 (Medium)	1.5603 (Low)
Cluster 7	0.6992 (Medium)	-10.1853 (Low)	0.0872 (Medium)	4.2914 (High)	32.5594 (High)	3.2306 (Medium)

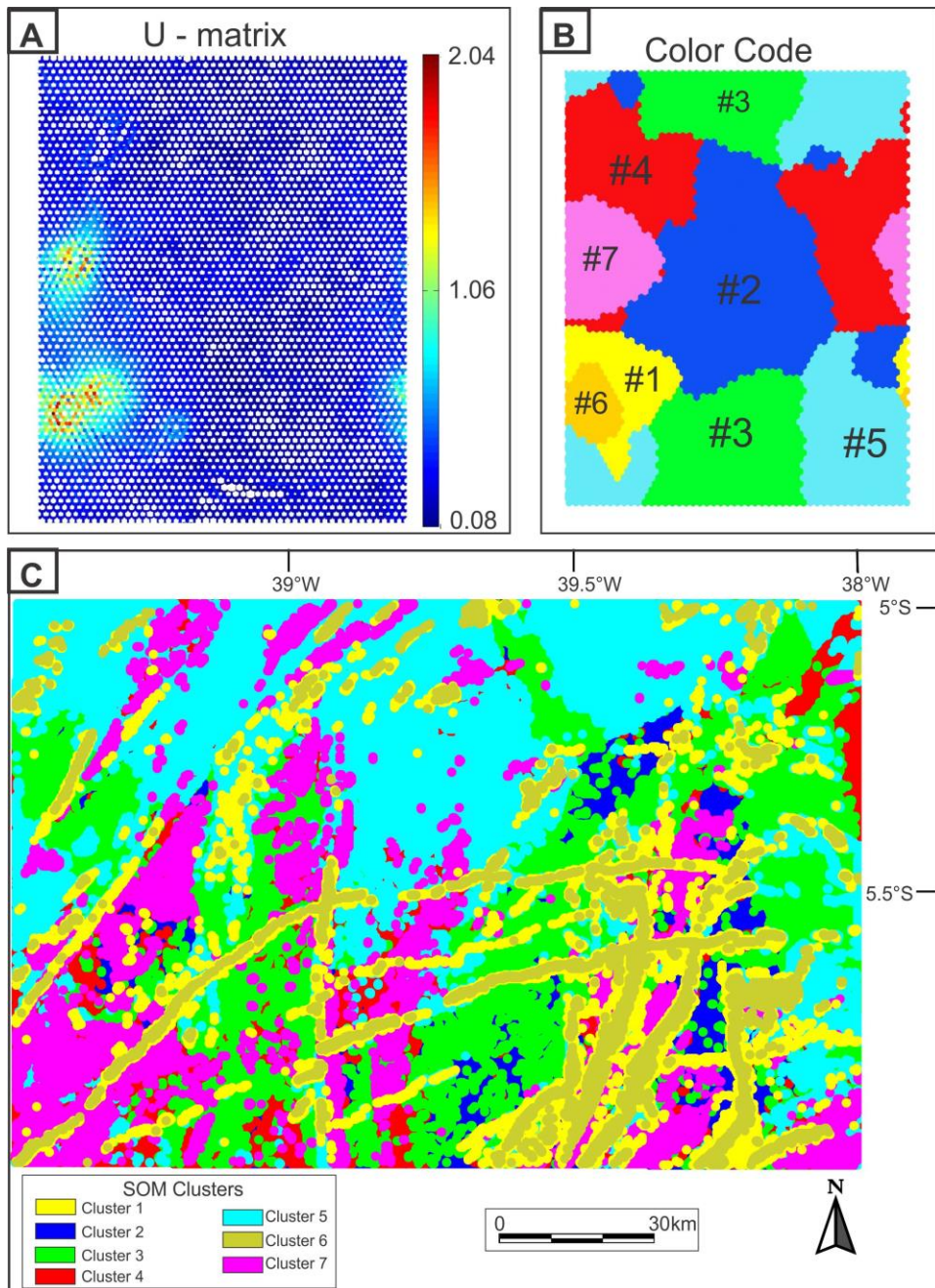


Figure 3. (A) U-Matrix, with the nodes coloured to represent their similarity; (B) K-means cluster classification of the BMU vector values on the self-organized map; (C) Spatial map displaying each sample coded by its cluster colour resultant from the SOM analysis

Conclusions

The application of the SOM method improves the delineation of the geophysical anomalies present in magnetic and gamma-ray maps and provides a higher resolution interpretive map. With the patterns of airborne geophysical data and the addition of SOM solutions directly related to the outcropped magmatic bodies, it was possible to estimate the distribution of these rocks and other buried bodies and thus to recognize the more realistic extension of the magmatic events in the Borborema Province.

However, an appreciation of the effects and influence of each input variable characteristics is necessary to understand the significance of the resultant SOM classes. The SOM analysis of airborne geophysical datasets produced 7 clusters in the studied area in Borborema Province. In this study, SOM has proven to be an effective tool to map dike swarms which two clusters (1 and 6) were associated with this dikes.

Acknowledgments

The authors would like to thank: Fundação Norte-Rio-Grandense de Pesquisa e Cultura (FUNPEC/UFRN/PETROBRAS/PRHPB 229 – 2013) (project 672013) for granting a PhD scholarship, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) for supporting a sandwich Ph.D. scholarship (Process 88881.188664/2018-01) to the first author and CNPq for DLC PQ grants; the Geologic Survey of Brazil (CPRM/Brazil) for the geophysical data; the Australian Commonwealth Scientific and Industrial Research Organization Division of Exploration and Mining (CSIRO/EM), located at the "Queensland Centre for Advanced Technologies" (QCAT), in Brisbane, for the facilities to execute the SOM analysis.

References

- Baçaõ, F., Lobo, V., and Painho, M.** 2005. The self-organizing map, the Geo-SOM, and relevant variants for geosciences: Computers and Geosciences, v. 31, p. 155–163.
- Blakely, R.J.** 1995. Potential Theory in Gravity and Magnetic Applications. Cambridge University Press, London, 441 p.
- Davies, D.L., Bouldin, D.W.** 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 1(2), p. 224–227
- Fraser, S.J., Dickson, B.L.** 2007. A New Method for Data Integration and Integrated Data Interpretation: Self-Organizing Maps. In: "Proceedings of Exploration 07: Fifth

Decennial International Conference on Mineral Exploration" edited by B. Milkereit, 2007, p. 907-910.

Hollanda, M.H.B.M., Archanjo, C.J.Filho, Macedo Filho, A.A., Fossen, H., Ernst, R.E., de Castro, D.L., Melo, A.C., Oliveira, A.L. 2019. The Mesozoic Equatorial Atlantic Magmatic Province (EQUAMP): A new large igneous province in South America. In: Srivastava, R.K., Ernst, R. E., Peng, P., editor/s. *Dyke Swarms of the World: A Modern Perspective*. Springer, p.87-110

Kohonen, T. 2001. *Self-Organizing Maps*, third extended edition, Springer Series in Information Sciences, Vol 30 (Springer: Berlin, Heidelberg, New York).

Kohonen, T., Somervuo, P. 1998. Self-organizing maps of symbol strings: *Neurocomputing*, v. 21, p. 19–30.

Nabighian, M.N. 1972. The Analytic Signal of Two-Dimensional Magnetic Bodies with Polygonal Cross-Section: Its Properties and Use for Automated Anomaly Interpretation. *Geophysics*, v. 37, p. 507-517.

Neves, S.P., Vauchez, A., Feraud, G. 2000. Tectono-thermal evolution, magma emplacement, and shear zone development in the Caruaru area (Borborema Province, Province, NE Brazil). *Precambrian Research*, v. 99, p.1–32.

Torppa, J., Middleton, M., Hyvönen, E., Lerssi, J., Fraser, S. 2015. A novel spatial analysis approach for assessing regional-scale mineral prospectivity In Northern Finland. In: Sarala, Pertti, editor/s. *Novel Technologies for Greenfield Exploration*. Geological Survey of Finland (GTK). p. 87-120.