

# Mapas Auto-Organizáveis Aplicados ao Mapeamento do Potencial Mineral na Província Mineral de Carajás, Norte do Brasil

Emilson Pereira Leite<sup>1</sup>  
Carlos Roberto de Souza Filho<sup>1</sup>

<sup>1</sup> Universidade Estadual de Campinas – UNICAMP/IG  
Caixa Postal 6152 – 13083-970 – Campinas-SP, Brasil  
{emilson, beto}@ige.unicamp.br

**Abstract.** A preliminary Self-Organizing Map (SOM) was designed with the aim of integrating and searching for patterns in airborne geological and geophysical gamma-spectrometric and magnetic data of the Serra Leste region, Carajás Mineral Province. SOM is an unsupervised Artificial Neural Network method that performs a non-linear mapping from a high-dimensional data space to a 2-dimensional grid, whereas preserving the topological relations in the original data. The SOM grid can be efficiently used in an integrated visualization and understanding of the internal relationships in the data. The K-means algorithm is applied to the SOM grid to reduce the number of mapped patterns so as to facilitate interpretation. Unfolding of the clustered SOM grid associates each mapped pattern with the spatial position of each data point. Texture information is incorporated to the SOM classified map by integrating it with the First Principal Component of a multi-polarized R99B-SAR (SIVAM/SIPAM) image of the study area, thereby producing an end-product that can be useful for further detailed investigations.

**Palavras-chave:** Carajás Mineral Province, airborne geophysics, synthetic aperture radar, self-organizing maps, Província Mineral de Carajás, Aerogeofísica, Synthetic Aperture Radar, mapas auto-organizáveis.

## 1. Introdução

Mapas Auto-Organizáveis (SOM - *Self-Organizing Maps*; Kohonen, 1995) têm sido utilizados para análise e detecção de padrões em dados multi-dimensionais, tais como geoquímicos (e.g. Penn, 2004); multi-espectrais (e.g. Saveliev e Dobrinin, 1999); hiper-espectrais (e.g. Penn, 2002); geofísicos (e.g. Klose, 2006), entre outros. Métodos de estatística multi-variada tradicionais normalmente não são adequados para modelar dados com relações não-lineares, distribuições não-normais, dados esparsos ou descontínuos. A formação e localização de depósitos minerais são o resultado da interação entre diversos processos físicos que deixam assinaturas na forma de feições geológicas associadas aos depósitos. Tais feições são caracterizadas por respostas típicas em um ou mais conjuntos de dados (e.g. geofísicos, multi-espectrais, geológicos). A relação entre esses diferentes tipos de dados e depósitos minerais conhecidos pode ser extremamente complexa para ser adequadamente modelada por aproximações lineares. Além disso, os dados podem ser esparsos, descontínuos e formarem distribuições não-normais, tornando ineficiente a aplicação de métodos de estatística multi-variada tradicionais. SOM e outros métodos de Redes Neurais Artificiais (e.g. Leite e Souza Filho, 2008) são mais apropriados nestes casos, sendo que o SOM tem a vantagem de não requerer de dados de treinamento.

O objetivo deste trabalho é demonstrar, de forma preliminar, a potencialidade da utilização de SOM na interpretação conjunta de dados aerogeofísicos, geológicos e posterior integração com dados de SAR (Synthetic Aperture Radar), no intuito de auxiliar na identificação de padrões associados às mineralizações presentes na região de Serra Leste, Província Mineral de Carajás (PMC), norte do Brasil (Figura 1).

A área de estudo está localizada entre os paralelos 634.000 E e 666.000 E e os meridianos 9.325.000 N e 9.347.400 N, Zona 22S. Representa uma área chave na PMC devido à diversidade de depósitos minerais ali presentes. Compreende diversos tipos de depósitos de classe mundial incluindo: (i) o depósito de Serra Pelada (Au–Pd–Pt) hospedado pelo grupo Rio Fresco e pela Formação Águas Claras (~ 2.68 Ga; Cabral et al. 2002a; 2002b); (ii) depósitos de Cu–Mo–Au (e.g., Serra Verde; Villas and Santos, 2001) em rochas meta-

vulcânicas do Grupo Rio Novo, próximas ao contato com o Granito Estrela de 2.76 Ga (Marschik et al. 2005); (iii) depósitos de ferro em rochas arqueanas de 2.76 Ga do Supergrupo Itacaiúnas (e.g., Serra Leste); (iv) depósitos de manganês (e.g., Sereno; Coelho e Rodrigues, 1986) hospedado pelo Grupo Rio Fresco; e depósitos de cromo-EGP (Elementos do Grupo da Platina; e.g., Luanga) associados com complexos máfico-ultramáficos (Suita, 1988).

## 2. Dados

Os dados aero-gamaespectrométricos foram coletados em 1993 pela empresa GEOMAG, utilizando detectores de cristais de NaI ativados por Tálcio, a 60 m de altitude, ao longo de linhas de vôo orientadas na direção E-W, perpendicular ao *trend* de rochas do Grupo Rio Novo. Ao longo dessas mesmas linhas de vôo, foram coletados os dados aero-magnéticos pela Geoterrex-Dighem, em 1999, a 120 m de altitude.

Os dados geofísicos foram pré-processados para remoção de dados espúrios e interpolados utilizando algoritmo de curvatura mínima (Briggs, 1974), formando grids de concentração de Th, K, U e do campo magnético total, com espaçamento de 62,5 m. O IGRF-1995 (International Geomagnetic Reference Field) (IAGA, 1996) foi removido do campo magnético total e os dados foram micronivelados para remoção de ruídos remanescentes ao longo das linhas de vôo. Posteriormente, foram calculadas as derivadas direcionais do campo magnético anômalo e amplitude do sinal analítico (ASA) (Roest et al. 1992). Os mapas das derivadas direcionais em  $-x$  e  $-y$  serviram para traçar eixos magnéticos a partir dos quais foi elaborado um mapa de *buffers* de 200 m ao redor desses eixos, até um máximo de 2000 m.

O mapa de litologias foi utilizado como dado categórico, no qual uma sigla representando a litologia está associada a cada célula do *grid*. Todos os *grids* são compostos por 511 x 360 células com espaçamento de 62,5m entre elas.

## 3. Mapas Auto-Organizáveis (SOM - *Self-Organizing Maps*)

O algoritmo do SOM, também conhecido como Rede de Kohonen, foi desenvolvido pelo Físico Finlandês Teuvo Kohonen (Kohonen, 1995). SOM consiste em uma matriz de elementos chamados neurônios, nós ou unidades, usualmente arranjados em matrizes de baixa dimensionalidade (1-D ou 2-D), o que facilita a interpretação. Os elementos da matriz podem ser representados por hexágonos, retângulos ou até mesmo formas irregulares justapostas. Associado a cada unidade existe o vetor protótipo, com a mesma dimensionalidade em relação aos vetores de entrada. A imagem de um vetor de entrada  $\mathbf{x}$  na matriz SOM, também conhecida como matriz-U, é a unidade  $\mathbf{m}_c$  que mais se aproxima de  $\mathbf{x}$ . Esta unidade é chamada de unidade de melhor ajuste (*Best Matching Unit* - BMU) ou unidade vencedora. Usando uma medida de dissimilaridade  $d(\mathbf{x}, \mathbf{m}_i)$ , esta unidade é representada pelo índice:

$$c = \arg_i \min \{d(\mathbf{x}, \mathbf{m}_i)\}, \quad (1)$$

e os vetores protótipo são atualizados iterativamente segundo a equação:

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + \alpha(t) h_{ci}(t) [\mathbf{x}(t) - \mathbf{m}_i(t)]. \quad (2)$$

$h_{ci}(t)$  representa uma função de núcleo de vizinhança que atua dentro de um raio pré-definido, geralmente grande no começo do treinamento e que decresce com o aumento das iterações, definindo quais unidades e o quando elas serão afetadas por cada um dos vetores de entrada.  $\alpha(t)$  é a razão de aprendizado e também decresce durante o processo de treinamento.

Deste modo, os vetores protótipos são atualizados de forma a seguirem os vetores de entrada de uma maneira ordenada. Como resultado, os vetores protótipos de unidades

próximas uma das outras são similares e, portanto, a matriz SOM preserva a relação topológica dos vetores de entrada. O erro de quantização é dado por:

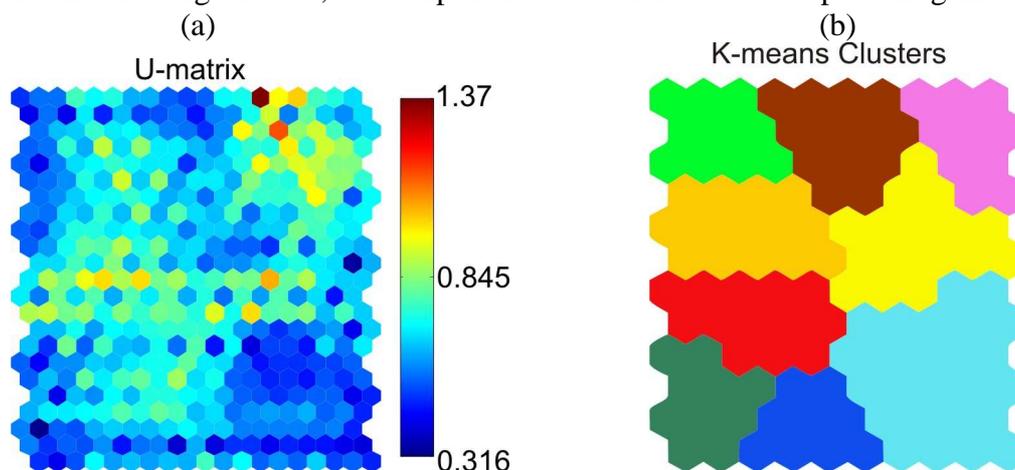
$$qer = \frac{1}{N} \sum_{t=1}^N \|\mathbf{x}(t) - \mathbf{m}_i(t)\|, \quad (3)$$

onde  $N$  é o número de vetores de entrada que forma o vetor protótipo  $\mathbf{m}_i$ . Para agrupar as unidades da matriz SOM em classes similares, foi utilizado o algoritmo *K-means*, onde o número de classes  $K$  foi definido com auxílio do índice de Davies-Bouldin (Davies e Bouldin, 1979).

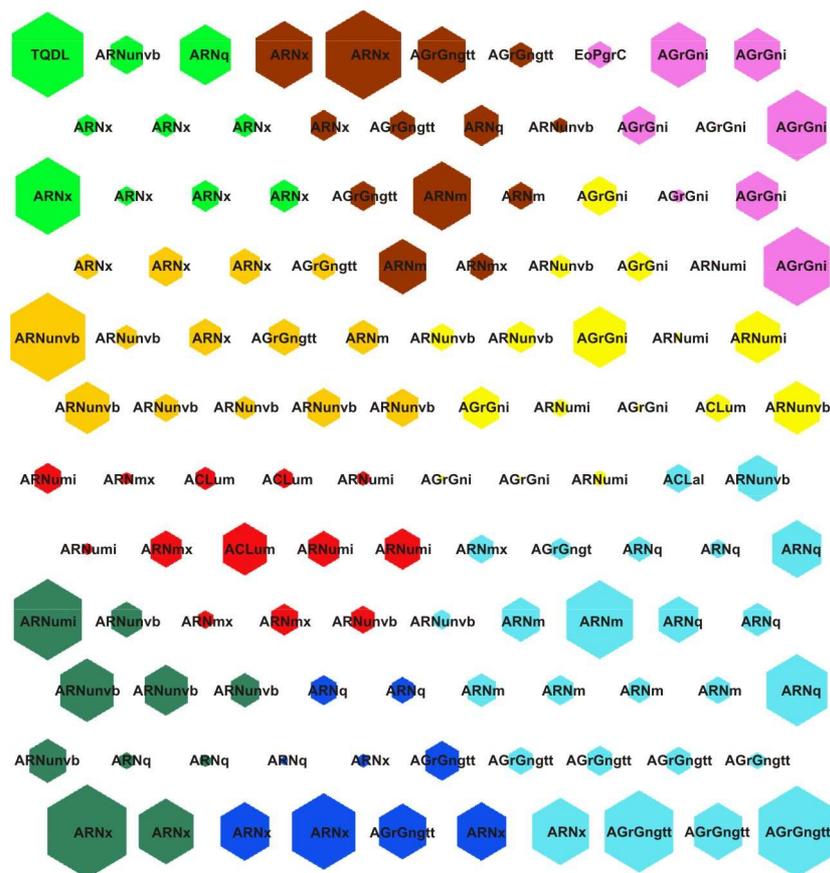
### 3. Resultados

A Figura 1a apresenta a matriz SOM, obtida utilizando-se 12 x 10 unidades. Os menores valores nesta matriz indicam as menores dissimilaridades entre os vetores protótipo e os vetores de entrada e são representados por cores frias. Por outro lado, os maiores valores indicam as maiores dissimilaridades e são representados por cores quentes. A matriz SOM foi reclassificada em 9 classes pelo algoritmo *K-means* (Figura 1b). O programa SiroSOM© v 2.0 foi utilizado para o cálculo do mapa SOM e posterior reclassificação.

Na Figura 2, cada hexágono da matriz SOM foi re-escalonado de acordo com a quantidade de vetores de entrada utilizados para gerar o vetor protótipo correspondente. Quanto maior o hexágono, maior a quantidade de vetores de entrada utilizada. A Figura 3 apresenta diagramas de dispersão das unidades do mapa SOM, indicando as cores das classes SOM da Figura 1b. As contribuições individuais de cada variável de entrada para a matriz SOM são apresentadas nos planos das componentes (Figura 4). As Figuras 5 e 6 mostram, respectivamente, a distribuição espacial do mapa SOM re-classificado e a distribuição espacial do erro de quantização. O histograma da distribuição do erro de quantização apresenta média de 1,07 (Figura 7). O mapa SOM re-classificado foi sobreposto à imagem da Primeira Componente Principal dos dados SAR R99-B/SIPAM-SIVAM, nas polarizações HH, HV e VV, através da transformação *Hue-Saturation-Value*, onde a componente *Hue* foi representada pelo mapa SOM reclassificado, a componente *Saturation* foi representada por um valor constante e igual à 0.6, e a componente *Value* foi substituída pela imagem SAR.

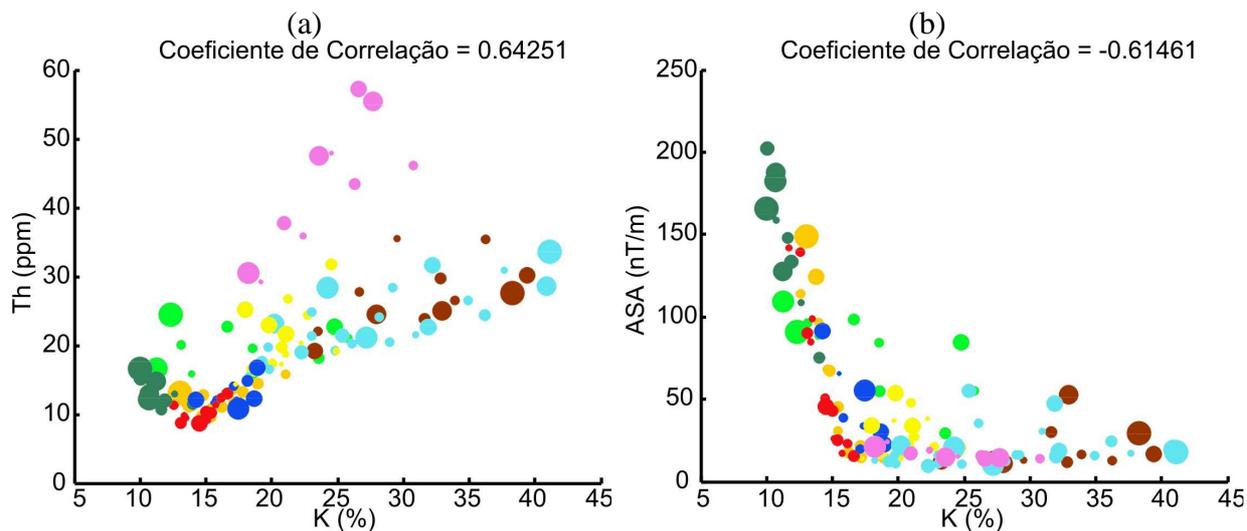


**Figura 1.** (a) Matriz SOM. Cada hexágono representa um vetor protótipo. Para melhor visualização e contraste entre as unidades, a matriz é preenchida com valores indicatrizes em linhas e colunas adjacentes àquelas que contêm dados. (b) Matriz SOM reclassificada em 9 classes definidas pelo algoritmo *K-means*.

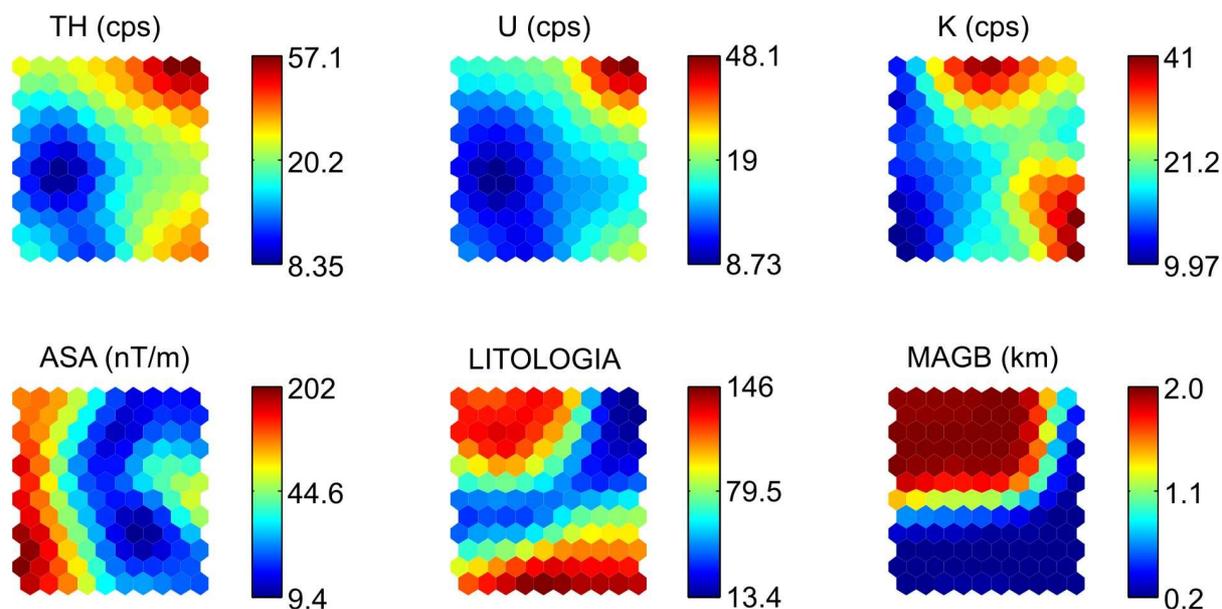


- TQDL:** coberturas lateríticas quaternárias
- EoPgrC:** granito cigano
- ACLum:** dunitos e peridotios
- ACLgg:** gabros de granulação grossa
- ACLg:** gabros
- ACLal:** anortositos e leucogabros
- AgrE:** granite estrela
- ARNx:** sericita-clorita-xistos e talco-xistos;
- ARNvb:** rochas vulcânicas básicas
- ARNunvb:** rochas máficas e vulcânicas básicas
- ARNumi:** rochas ultra-máficas não-diferenciadas
- ARNmx:** rochas máficas e xistos
- AgrGngt:** granites tonalíticos e gnaisses
- AgrGngtton:** tonalitos e granitos plutônicos
- AgrGni:** granitos não-diferenciados e gnaisses;
- ARNq:** quartzitos
- ARNm:** meta-pelitos e meta-siltitos

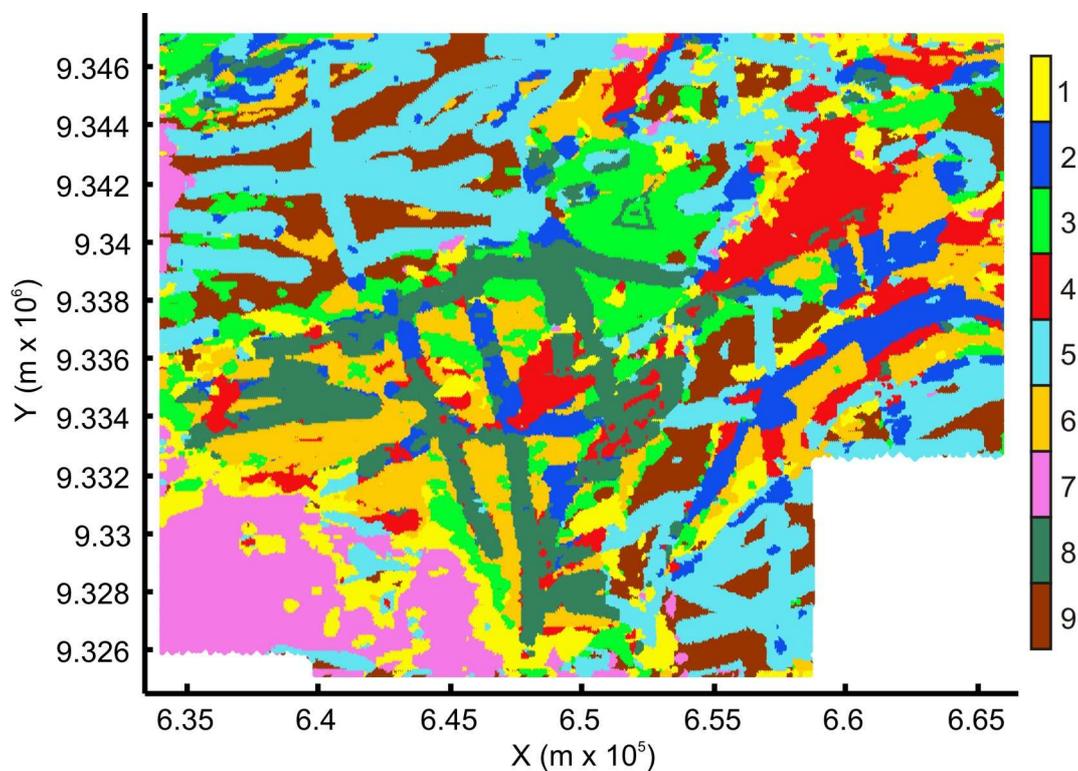
**Figura 2.** Mapa SOM reclassificado com legendas das litologias sobrepostas. O tamanho dos hexágonos é proporcional à quantidade de vetores de entrada associados aos respectivos vetores protótipos. As cores são equivalentes às da Figura 1b.



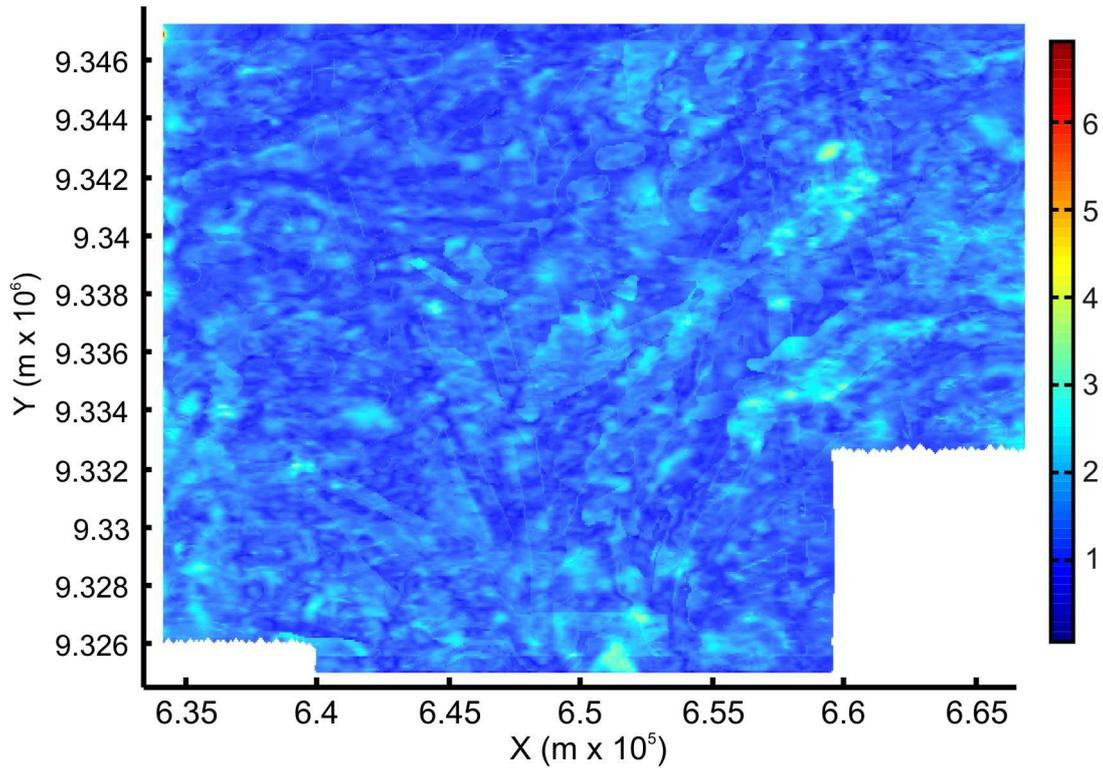
**Figura 3.** Diagramas de dispersão das unidades SOM coloridas com as cores do mapa SOM reclassificado da Figura 1b. O tamanho de cada círculo é proporcional à quantidade de vetores de entrada utilizados para a determinação do vetor protótipo representante da respectiva unidade. (a) Concentração de Tório contra concentração de Potássio. (b) Amplitude do Sinal Analítico (ASA) do campo magnético anômalo contra a concentração de Potássio.



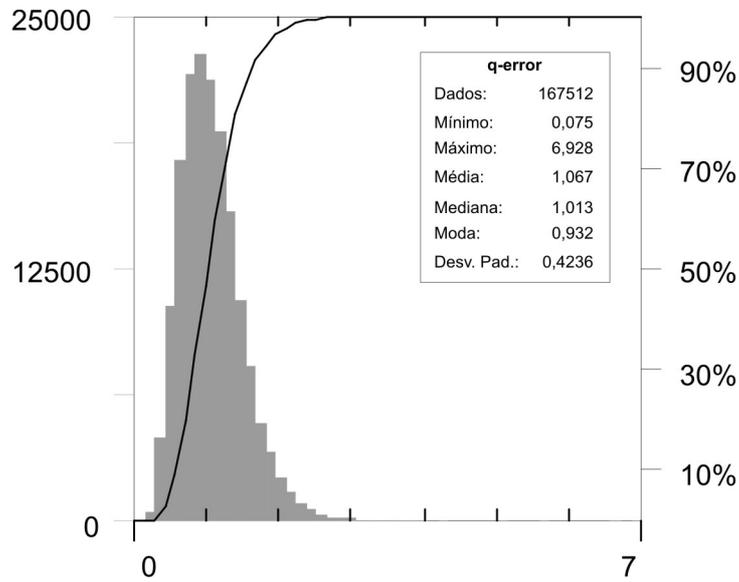
**Figura 4.** Planos de componentes de cada variável de entrada. TH = concentração de Tório, U = concentração de Urânio, K = concentração de Potássio, ASA = Amplitude do Sinal Analítico, MAGB = *Buffers* em torno de eixos magnéticos. As litologias foram codificadas numericamente de forma não-ordenada.



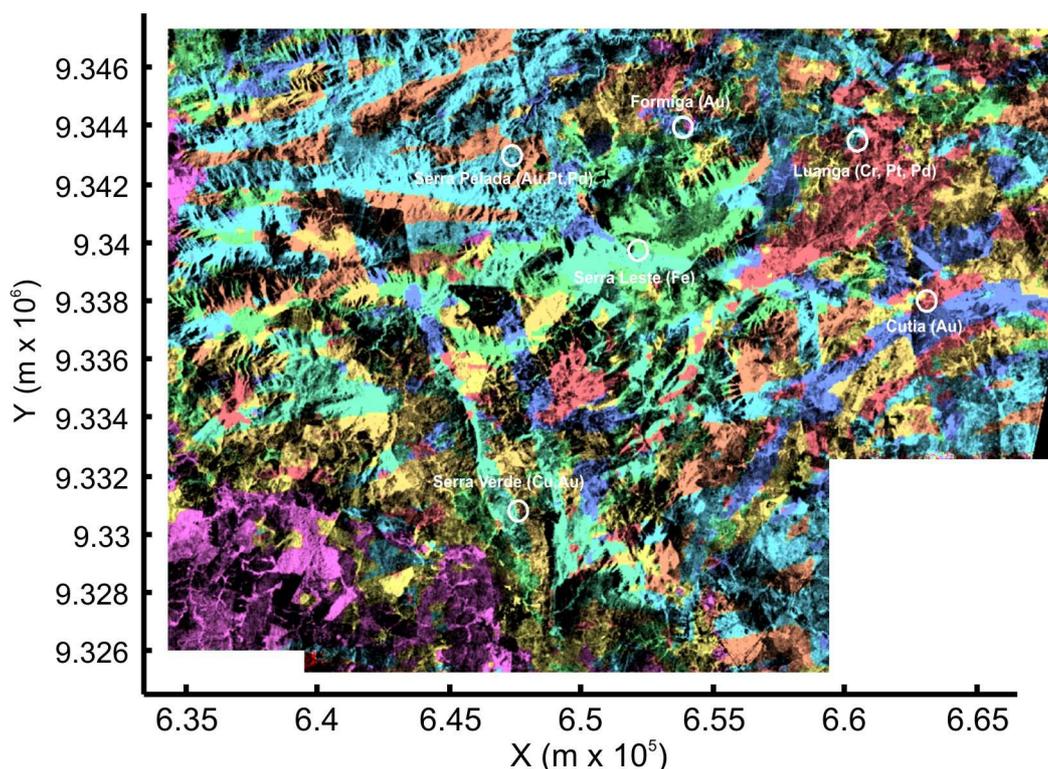
**Figura 5.** (a) Distribuição espacial do mapa SOM reclassificado sobre um *grid* de tamanho e espaçamento equivalente aos *grids* de entrada.



**Figura 6.** Erro de quantização associado à cada vetor protótipo da matriz SOM.



**Figura 7.** Histograma da distribuição do erro de quantização com estatística básica incluída.



**Figura 8.** Mapa SOM reclassificado sobreposto à imagem da Primeira Componente Principal dos dados R-99B/SAR(SIPAM/SIVAM), nas polarizações HH, HV e VV. As ocorrências minerais reconhecidas na área são indicadas por círculos.

#### 4. Discussão e Conclusões

O mapa da Figura 2 é uma excelente forma de analisar o mapa SOM e compará-lo com as litologias da área de estudo. As coberturas lateríticas estão distribuídas, em sua totalidade, na classe 3 (Figura 5), região onde localiza-se o depósito de Fe de Serra Leste. As rochas do complexo máfico-ultramáfico Luanga estão associadas à classe 4 (Figura 5), relacionada ao depósito de Cr-EGP presente neste complexo. A classe 5 (Figura 5) mapeou os meta-pelitos e meta-siltitos, estando também associada à valores intermediários da concentração de Tório, valores baixos do campo magnético e valores intermediários a altos da concentração de Potássio. Nesta classe encontra-se o depósito de Au-Pt-Pd de Serra Pelada.

Os resultados gerados neste trabalho mostram a potencialidade do SOM na integração e interpretação conjunta de dados remotos multi-dimensionais, aplicados ao mapeamento do potencial mineral em áreas de densa cobertura vegetal, como a Província Mineral de Carajás. Em uma etapa futura, a classificação textural prévia da imagem R-99B/SAR(SIPAM/SIVAM) permitirá a incorporação de classes texturais na elaboração do mapa SOM.

#### Agradecimentos

E.P. Leite e C.R. Souza Filho agradecem a Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) pela bolsa de pós-doutorado (Proc. N.: 05/04453-3) e pelo suporte financeiro (Proc. N. 03-09916-6), respectivamente.

#### Referências Bibliográficas

Briggs I.C., 1974. Machine contouring using minimum curvature. **Geophysics** 39, 39–48.

Cabral A.R., Lehmann B., Kwitko R., Cravo C.H.C., 2002a. The Serra Pelada Au–Pd–Pt deposit, Carajas Mineral Province, northern Brazil; reconnaissance mineralogy and chemistry of very high grade palladian gold mineralization. **Econ Geol** 97, 1127–1138.

Cabral A.R., Lehmann B., Kwitko R., Cravo C.H.C., 2002b. Palladium and platinum minerals from the Serra Pelada Au–Pd–Pt deposit, Carajás Mineral Province, Northern Brazil. **Can Mineral** 40, 1451–1463.

Davies D. and Bouldin D., 1979. A Cluster Separation Measure, **IEEE Trans. Pattern Anal. Mach. Intell.** 1(2) 224-227.

IAGA Division 5, Working Group 8, 1996. International geomagnetic reference field, 1995, revision. **Geophysical Journal International** 125, 318–321

Klose, C.D., 2006. Self-Organising Maps for Geoscientific Data Analysis: Geological Interpretation of Multi-dimensional Geophysical Data. **Computational Geosciences**, 10, 265-277, doi:10.1007/s10596-006-9022-x.

Kohonen, T., 1995. **Self-Organizing Map**, 2nd ed., Springer-Verlag, Berlin, pp. 113.

Leite, E.P. and Souza Filho, C.R., 2008. Probabilistic neural networks applied to mineral potential mapping for platinum-group elements in the Serra Leste region, Carajás Mineral Province, Brazil. **Computers & Geosciences**, Article in Press, doi: 10.1016/j.cageo.2008.05.003.

Liu, Y., R. H. Weisberg, and C. N. K. Mooers, 2006. Performance evaluation of the self-organizing map for feature extraction. **Journal of Geophysical Research**, 111, C05018, doi:10.1029/2005JC003117.

Penn, S.B., 2002. Using self-organizing maps, histograms, and standard deviation to detect anomalies in hyperspectral imagery data. In: **Proceedings of the Fifth International Airborne Remote Sensing Conference**. 22–24, Miami, Florida.

Penn, S.B., 2004. Using self-organizing maps to visualize high-dimensional data. **Computers & Geosciences**, 31(5), 531-544.

Roest W.R., Verhoef J., Pilkington M. 1992. Magnetic interpretation using the 3-D analytical signal. **Geophysics** 57, 116–125.

Saveliev, A. A. and Dobrinin, D. V., 1999. Hierarchical multispectral image classification based on self organized maps. In: IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99, 5, 2510–2512, Piscataway, NJ.

SiroSOM© v 2.0 – CSIRO Exploration & Mining. – <http://em.csrio.au/> - [acessado em 15/11/2008]

Suita M.T.F., 1988. Geologia da área Luanga com ênfase na petrologia do complexo básico-ultrabásico Luanga e depósitos de cromita associados, Pará. Dissertação de Mestrado, Brasília, UnB, 83 pp.

Villas R.N., Santos M.D., 2001. Gold deposits of the Caraja's Mineral Province: deposit types and metallogenesis. **Mineralium Deposita** 36, 300–333.