

Aplicação de Redes Neurais Artificiais para Reconhecimento de Padrões em Solos

Fabiane Hilario dos Santos Costa¹
Carlos Roberto de Souza Filho¹

¹Universidade Estadual de Campinas
Caixa Postal 6152 CEP 13083-970
Campinas – SP – Brasil
fabianeusc@hotmail.com
beto@ige.unicamp.br

Abstract. The application of structures of artificial neural nets integrated to multisource data, allows establishing relations of synthesis between the variable that represents the accessible environment and environment phenomena accessible on different analytical levels. The space modeling, carried through the application of neural nets in multisource data, was made in the ArcSDM program (Arc Spatial Modeler Date), software that works as an extension of the ArcGis. By the GeoXplore module, it was applied three systems of training and classification of data: Radial Basis Functional Link Network (RBFLN) and Probabilistic Neural Network (PNN).

Palavras-chave: neural net, ArcSDM, modeling, geochemistry, redes neurais, ArcSDM, modelagem, geoquímica.

1. Introdução

As observações da Terra, no futuro, estão em aliar “inteligência” aos sistemas de coleta e análise de dados. O desenvolvimento da tecnologia espacial tem o intuito de adquirir informações sobre os processos e estruturas que ocorrem na superfície terrestre para fins de manejo de recursos naturais e planejamento ambiental.

As Redes Neurais Artificiais (RNA's) têm sido aplicadas para reconhecimento de padrões. Nessa perspectiva, Lee et al. (1990) *apud* Rojas (2004), relatam que as redes neurais desempenham bem esta tarefa ou melhor que as técnicas estatísticas, uma vez que não requerem que a natureza paramétrica da distribuição dos dados a ser classificados seja explícita. Sendo assim, as RNA's estão baseadas em uma arquitetura de elementos de processamento simples de entrada e saída, parecido aos neurônios no cérebro humano, e estão demonstrando sucesso em uma ampla variedade de aplicações, incluindo processamento e interpretação de dados de sensores remotos (Atkinson e Tatnall 1997).

Nesse contexto, esse trabalho utilizou um sistema de redes neurais artificiais (RNA) para a classificação supervisionada de mapas evidenciais multifonte, visando à análise de possíveis padrões ocultos contidos nesses dados que tenham associação espacial a locais ricos em Pb. Para tanto, pontos que exibem concentrações mais altas de Pb no solo e em sedimentos de corrente (valores anômalos em relação ao background geoquímico) foram utilizados para treinamento de RNAs.

1.1 Localização da Área de Estudo

A Bacia Hidrográfica do Rio Ribeira de Iguape (Fig. 1), o Complexo Estuarino Lagunar de Iguape-Cananéia-Paranaguá e as diversas bacias hidrográficas encaixadas entre esta e o Oceano Atlântico, genericamente denominada Vale do Ribeira, possuem uma área de

2.830.666 hectares (28.306 km²), abrangendo as regiões sul do estado de São Paulo (1.711.533 ha) e leste do estado do Paraná (1.119.133 ha), (ISA 1998).

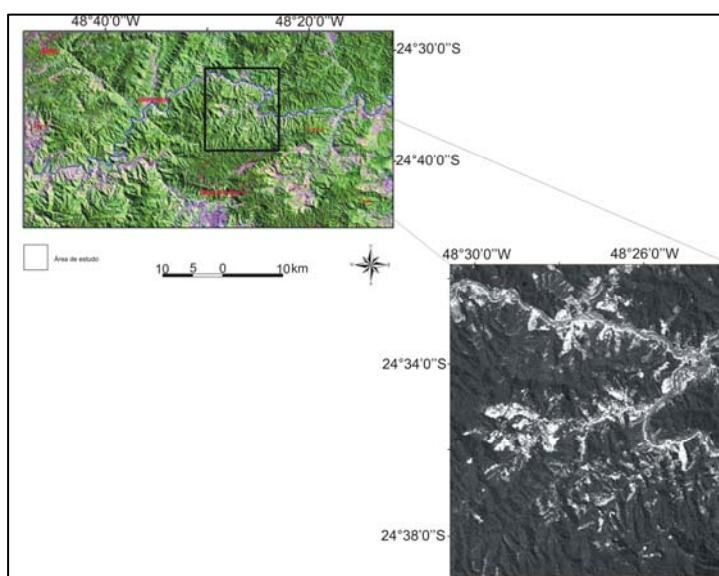


Figura 1: Localização da área de estudo.

Nesta região, a exploração de ouro data desde século XVI e um depósito de ouro primário, conhecido como Morro do Ouro, nas proximidades de Apiaí, foi explotado até o início dos anos 40. Com a descoberta do minério de chumbo, zinco e prata, ainda no século 19, iniciou-se a lavra desses minérios, que perdurou até a década de 90. A partir da década de 70, projetos de reconhecimento geoquímico regional desenvolvido pelo Serviço Geológico do Brasil (CPRM) indicaram anomalias significativas de cobre, chumbo e zinco que nortearam as pesquisas locais visando à prospecção de ouro e sulfetos associados (Silva e Câmara 1990).

2. Metodologia de Trabalho

Os dados utilizados como base para modelagem através de redes neurais artificiais compreendem: (i) geoquímica de solos e sedimentos de corrente para o elemento Pb (Fig. 2) e (ii) um conjunto de mapas evidenciais multifonte (Fig. 3).

O início do processo de modelagem deu-se com a seleção da área de estudo, onde o critério utilizado para escolha foi uma área onde houvesse a melhor cobertura dos dados geoquímicos, tanto de solo como de sedimentos. Após a escolha da área, foi desenvolvida a primeira etapa de processamento dos dados que se deu através da combinação dos mapas evidenciais multifonte, como são dados provenientes de naturezas distintas, é necessário combiná-los num mapa (*raster*) único, que contém o conjunto de evidências agrupadas (*feature vector map*).

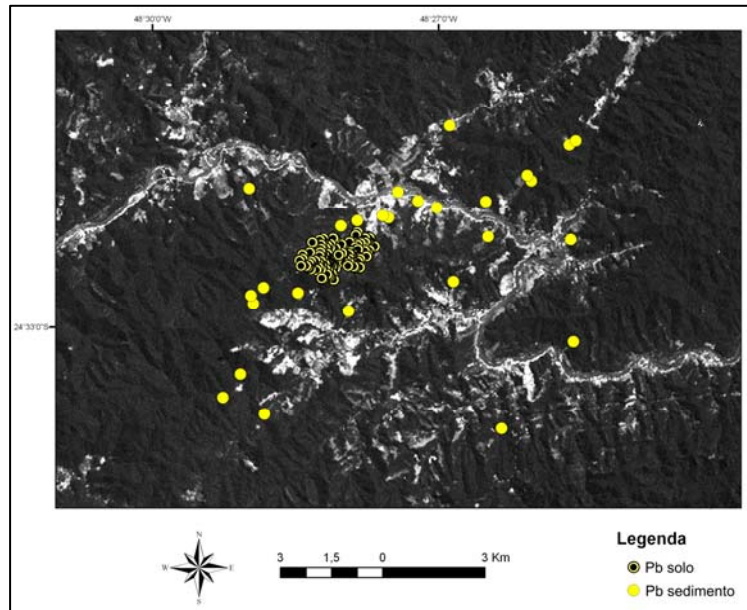


Figura 2: Mapa de localização dos pontos de amostragem de chumbo em solos e sedimentos de corrente na área de estudo. A imagem de fundo é a banda 3 do Landsat ETM+.

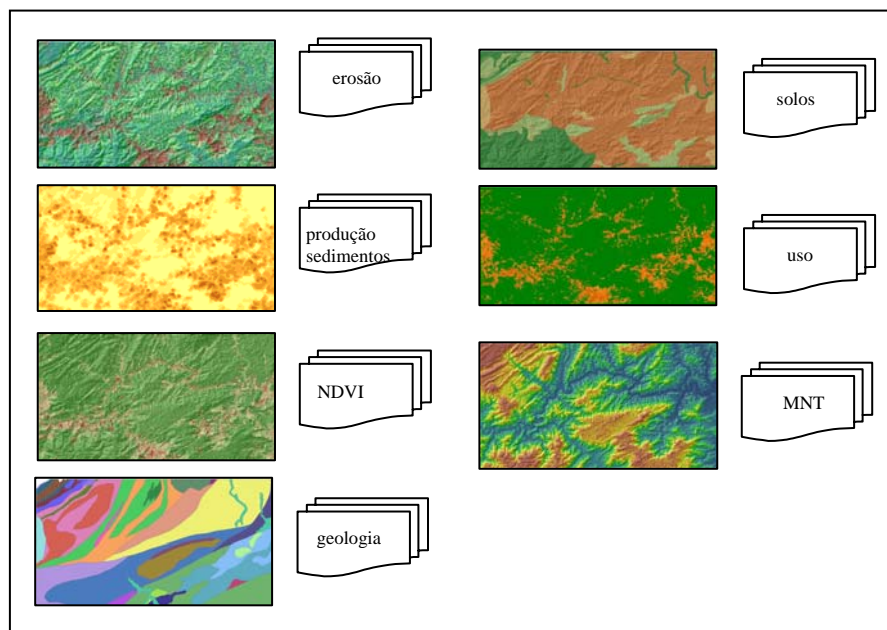


Figura 3: Variáveis de entrada para modelagem através de Redes Neurais Artificiais.

Os pontos de treinamento utilizados foram os de dados geoquímicos de Pb em solos e sedimentos de corrente provenientes do banco de dados do Serviço Geológico do Brasil. Após esta etapa de pré-processamento, deu-se início a modelagem por RNAs (Fig.4) pelo módulo GeoXplore. Como produto deste processamento, obteve-se dois mapas de classificação com valores ordenados, cujos limites de classes foram ajustados pelo método de *quantis*, para uma melhor visualização dos resultados, eliminando-se os valores não associados às regiões anômalas.

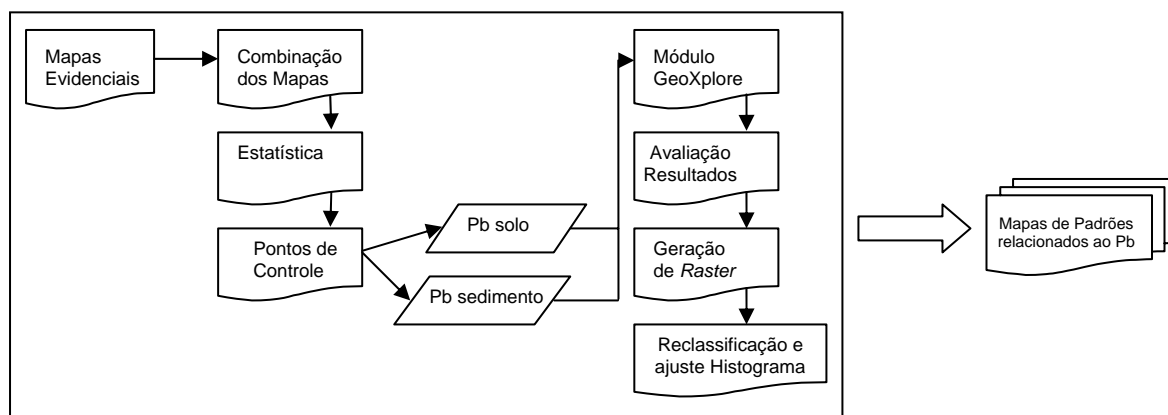


Figura 4: Fluxograma da metodologia aplicada para modelagem espacial utilizando RNAs.

3. Resultados e Discussões

As técnicas de modelagem espacial têm como premissa a orientação pelos dados, onde os pontos de treinamento em quantidade e qualidade são fundamentais para reconhecimento de padrões. O trabalho aqui desenvolvido utilizou dados de análises geoquímicas de Pb em amostras de solo (com concentrações mais altas, acima de 94 ppm) como pontos de treinamento na classificação de dados multi-fonte, através de redes neurais.

O sistema de classificação *Probabilistic Neural Network* (PNN) foi o que apresentou o melhor resultado. Através da aplicação dessa rede neural foi obtido um mapa com áreas classificadas que possuem um padrão inter-mapas evidenciais (Figura 5). Isso demonstra que há uma associação espacial entre os pontos que exibem teores mais elevados de Pb no solo e um conjunto de informações temáticas que foram detectadas como um determinado padrão pelas RNAs, na forma de áreas.

Objetivando verificar uma provável repetição dos padrões mapeados para uma região mais ampla em relação aquela onde foi realizada a amostragem de solos, uma nova classificação foi testada com a rede PNN, mas utilizando-se dados regionais de geoquímica de Pb derivados de sedimentos de corrente. Os resultados dessa classificação são ilustrados na Figura 6.

Esses resultados permitem as seguintes ponderações:

(i) o resultado da classificação obtido por RNA (Figura 5), mostra que a assinatura extraída dos dados multi-fonte aparece tanto em setores da área de estudo para os quais se tem amostras de geoquímica de solos (utilizadas como pontos de treinamento), como também em setores aonde inexistem amostras de solos coletadas e analisadas.

(ii) a análise conjunta do resultado da classificação por RNAs apresentadas nas Figuras 5 e 6 mostra que as áreas que apresentam, simultaneamente, geoquímica de solos e de sedimentos de corrente como pontos de treinamento na classificação, foram mapeadas como zonas equivalentes, ou seja, com o mesmo padrão nos mapas multi-fonte.

(iii) outro fato interessante é que áreas onde não se tem geoquímica de solos para uso no processo de classificação, mas que foram mapeadas com base nos pontos de treinamento de geoquímica de solo, coincidem com as áreas mapeadas na classificação através de pontos de treinamento de sedimentos de corrente.

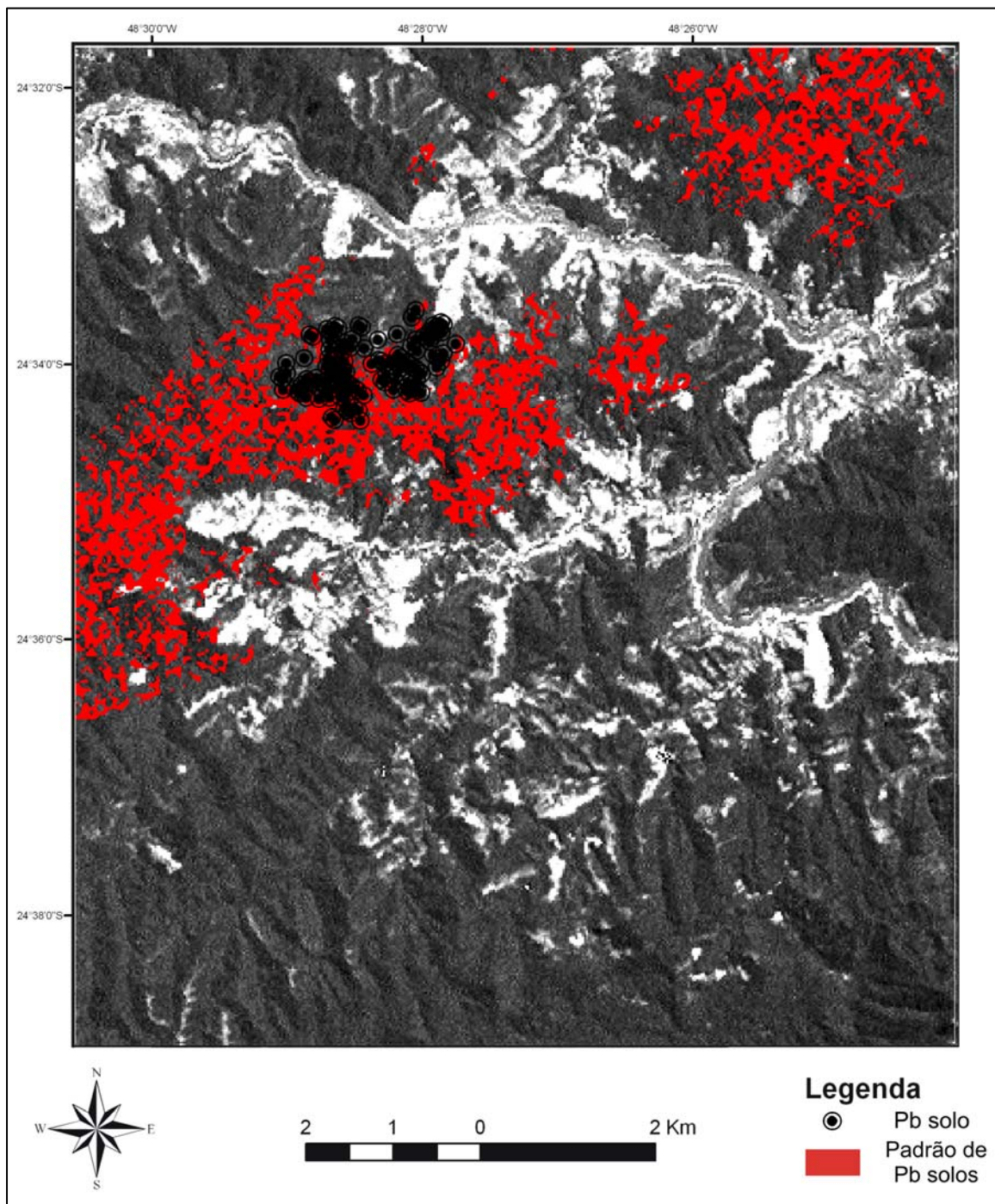


Figura 5: Mapa de padrões anômalos de Pb em solo como resultado da classificação pelo sistema *Probabilistic Neural Network* (PNN) na área de estudo. A imagem de fundo é a banda 3 do Landsat ETM+.

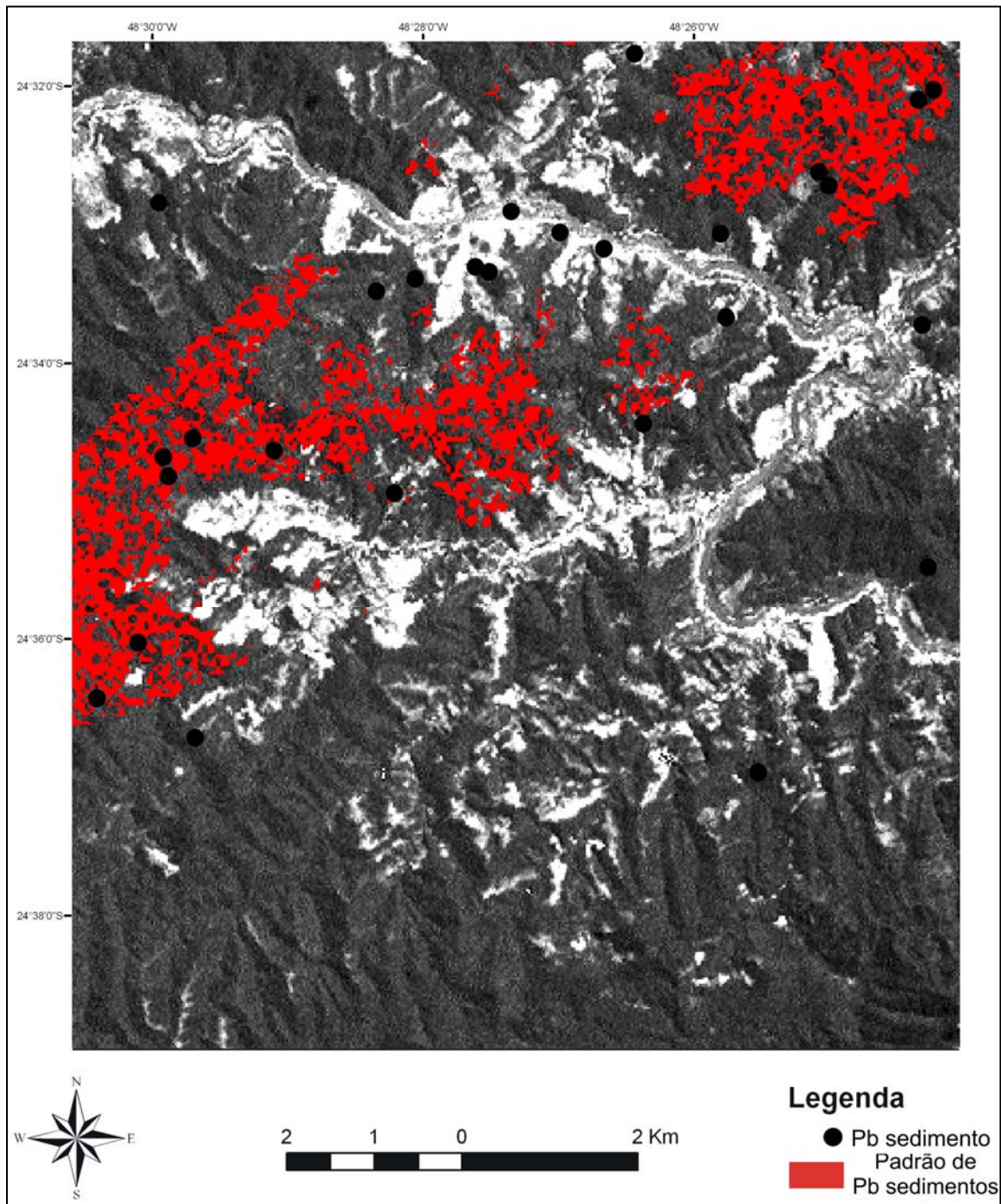


Figura 6: Mapa de padrões anômalos de Pb em sedimentos como resultado da classificação pelo sistema *Probabilistic Neural Network* (PNN) na área de estudo. A imagem de fundo é a banda 3 do Landsat ETM+.

A interpretação desses resultados tem desdobramentos interessantes, entre os quais se destaca os seguintes:

(i) possivelmente há uma correlação entre a geoquímica de solos e a de sedimentos de corrente nessa área, independente de época e condições de levantamento. Alba & Souza Filho (2008), analisaram dados geoquímicos de solos e sedimentos de corrente no médio Vale do Ribeira e demonstraram, através de outros métodos de análises de dados, a existência dessa

mesma correlação. Este fato deve-se a diversos aspectos, tais como relevo, tipo de vegetação, litologia, pedologia ou uma combinação parcial ou total desses elementos.

(ii) considerando essas coincidências de resultados, é plausível que seja possível utilizar um ou outro levantamento como pontos de treinamento para se mapear padrões nos dados multifonte que sejam espacialmente compatíveis com zonas de maior concentração de Pb.

(iii) na medida em que esse comportamento seja bem equacionado, pode-se estender a classificação por RNA para outras áreas vizinhas, sem nenhum levantamento geoquímico.

A qualidade dos resultados obtidos pela modelagem por redes neurais associa-se aos atributos dos dados de entrada, de forma que determinadas variáveis podem contribuir mais (ou menos) no resultado da classificação. Analisando-se os ponderadores que a rede PNN calculou para cada mapa evidencial utilizado, observa-se que o mapa geológico, o mapa de uso e o NDVI são as evidências com maior peso e são as que mais influenciaram no modelo final.

4. Conclusões

O processo de modelagem espacial aqui realizado através de RNAs, visando o reconhecimento de áreas que possam ter um padrão multi-fonte associado a anomalias geoquímicas de Pb, foi realizado através da extensão *Arc Spatial Data Modeller (ArcSDM)* no *software Arc Gis 9.2*. A modelagem compreendeu testes com dois sistemas supervisionados, nos quais o PNN mostrou melhor *performance* ao detectar uma assinatura pouco ambígua nos dados multifonte, que possui ampla coerência espacial aos pontos de treinamento (anomalias de Pb em solos).

A validação, das áreas classificadas com base em geoquímica de Pb em amostras de solos, foi feita com os dados geoquímicos de Pb em sedimento de corrente. Esses dados mostram superposição de áreas com mesmo padrão, sendo assim, os resultados apresentam-se similares entre si. É importante salientar que os dados geoquímicos, sob o aspecto de concentração, são semelhantes, o que favorece o treinamento da rede na busca de padrões nos mapas fontes.

O modelo proposto é ainda experimental. Para sua melhor validação, serão necessárias novas etapas de campo, seguidas de análises geoquímicas de amostras extraídas de locais onde os mapas indicam áreas que possuem o mesmo padrão. Neste caso, a área alvo escolhida seria a região nordeste da região de estudo.

Referências Bibliográficas

Atkinson, P. M.; Tatnall, A.R.L.,1997. Neural networks in remote sensing. **International Journal of Remote Sensing**, v. 18, n. 4, p. 699-709.

Instituto Socioambiental (ISA), 1998. **Diagnóstico Socioambiental do Vale do Ribeira: Documento Síntese**. São Paulo.

Rojas, E.H.M., 2004. **Síntese Genética de Redes Neurais Artificiais ART2 na Classificação de Imagens ASTER para Mapeamento de Uso e Cobertura da Terra na Região Norte do Mato Grosso**. Dissertação de Mestrado, Instituto Nacional de Pesquisas Espaciais, São Jose dos Campos – SP, 124 p.

Silva, C.R. E Câmara, M.M., 1990. Depósito de ouro e prata do Ivaporunduva, município de Eldorado Paulista (SP). In: SBG, Congresso Brasileiro de Geologia, 36, Natal. **Anais...**