



Comparação do desempenho das técnicas Análise por Principais Componentes e Fator de Máxima Autocorrelação para processamento de dados aerogeofísicos: investigação teste num prospecto da Província Mineral do Tapajós

Thais Andressa Carrino, IG-UnB; Adalene Moreira Silva, IG-UnB; Nilson Francisquini Botelho, IG-UnB; Marcelo Pinto, MAPEX Mineração; Luiz Carlos Pinheiro Clerot, MAPEX Mineração.

Copyright 2011, SBGf - Sociedade Brasileira de Geofísica

This paper was prepared for presentation during the 12th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, August 15-18, 2011.

Contents of this paper were reviewed by the Technical Committee of the 12th 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

The MAF (Maximum Autocorrelation Factor) technique, more precisely the kernel version, was applied to an airborne geophysics dataset for the assessment of the geologic information in the area of the Maués Prospect (Brazilian Amazon). This method is characterized as another approach of the well known Principal Component Analysis technique that also was utilized in this case study. Based on the visual and statistical analysis, the results were compared for the determination of the main lithologies in this portion of the Tapajós Mineral Province. This case study shows that the use of the (Kernel) MAF has a great potential for applications in multivariate geologic data.

Introdução

O tratamento de dados multivariados em Geologia é importante para a extração da informação desejada em meio a um conjunto de dados muitas vezes redundante ou caracterizado por inúmeras variáveis para serem analisadas individualmente. Para tal, existem várias técnicas de análise multivariada que, conforme sintetizado em Hair et al. (1998), caracterizam-se pelo objetivo de “medir, explicar e prever o grau de relacionamento entre variáveis estatísticas”.

Com a finalidade de realçar unidades litológicas, imagens aerogamaespectrométricas do Prospecto Maués, localizado na Província Mineral do Tapajós, foram processadas por meio da Análise por Principais Componentes (APC) e do Fator de Máxima Autocorrelação (MAF - *Maximum Autocorrelation Factor*), esta última disponibilizada gratuitamente por Canty (2010) como uma extensão do aplicativo de *software* ENVI 4.7 em linguagem IDL (*Interactive Data Language*).

O Prospecto Maués é caracterizado por rochas hospedeiras de veios de quartzo sulfetados e mineralizados em ouro equivalentes a metassedimentos do Grupo Jacareacanga, de idade paleoproterozóica (~2,1 Ga). Granitos da Suíte Intrusiva Parauari (~1,8 Ga) estão em contato com as rochas metassedimentares do Grupo Jacareacanga, e falhas NW-SE são bem

marcadas nesta região, e estão relacionadas à instalação do *trend* aurífero Patrocínio-Espírito Santo (Coutinho, 2008).

Esse prospecto foi selecionado como área teste da aplicação comparativa das técnicas APC e MAF em função da cessão de dados aerogeofísicos pela empresa Mapex Mineração.

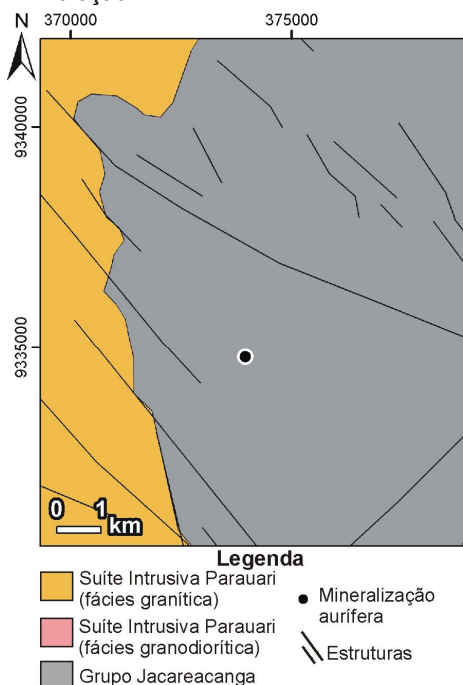


Figura 1. Mapa geológico abrangendo a área teste de estudo (modificado de Klein et al., 2001).

Técnicas de redução de dados selecionadas

A APC compreende uma técnica estatística usada para analisar inter-relações entre um grande número de variáveis e explicar essas variáveis em termos de suas dimensões inerentes comuns. O objetivo consiste em encontrar um meio de condensar a informação contida nas variáveis originais em um conjunto menor de variáveis estatísticas com perda mínima de informações (Hair et al., 1998).

As componentes principais são apresentadas na ordem decrescente de variabilidade, sendo que a última componente caracteriza-se pela menor contribuição que dá para a explicação da variabilidade total dos dados originais.

O foco da APC para a redução da dimensionalidade dos dados torna possível identificar padrões nos dados anteriormente não visualizados por possuírem alta correlação entre os mesmos.

Os autovetores representam a contribuição de cada imagem original a cada componente principal (PC) e podem ser positivos ou negativos devido ao reposicionamento da origem da distribuição das N variáveis. Como cada PC resultante representa uma adição ponderada pelo respectivo coeficiente da matriz de autovetores das variáveis amostradas, é possível selecionar a PC que contém a informação sobre o alvo analisado, e que está associada a valores mais altos dos autovetores específicos, com sinais positivos ou mesmo negativos.

A técnica MAF é comumente aplicada em geociências e em biologia para a redução de informações de um conjunto de dados visando, por exemplo, a detecção de mudanças (Nielsen & Canty, 2009; Nielsen & Olsson, 2010), a caracterização de reservatório (e.g., Sancevero, 2007), além da modelagem da variação da forma de objetos biológicos, como ossos (e.g., Larsen, 2002). A premissa básica compreende a existência de ordenamento ou estruturação de observações (variáveis), como em sensoriamento remoto, onde cada pixel consiste no espectro de reflectância organizado em duas dimensões (X e Y). Com base nesta técnica, procura-se uma transformação que maximiza a autocorrelação entre os pixels das imagens (Larsen, 2002), sabendo-se que o sinal de interesse caracteriza-se por alta autocorrelação (no presente estudo de caso associada à informação geológica) enquanto o sinal de menor interesse (ruído) possui baixa autocorrelação. Assim, o MAF1 compreende a combinação linear das variáveis originais que maximiza a autocorrelação; o MAF2 é a combinação linear de variáveis originais que maximiza a autocorrelação, sendo ortogonal ao MAF1; os demais MAFs são definidos de forma similar (Nielsen et al., 1998).

O Fator de Máxima Autocorrelação necessita da estimação da matriz de covariância dos dados e da matriz de covariância da diferença entre os dados originais e a versão espacial deslocada dos dados. Observa-se que a APC procura uma combinação linear que maximiza a variância, enquanto a técnica MAF visa a obtenção da máxima autocorrelação entre observações vizinhas. Por tal razão, ao se estudar imagens, o método MAF tende a possuir um desempenho melhor. Já a APC falha, por vezes, pelo fato de alguns sinais de interesse com alta autocorrelação terem pequena variância. Além disso, no caso da APC, a escolha de quantas PCs devem ser usadas no processo interpretativo ou classificatório é considerado subjetivo, sendo necessária a análise do gráfico de autovalor a fim de verificar até que PC a maior parte da variabilidade dos dados originais está representada. No caso dos MAFs, a questão da representação da máxima autocorrelação dos dados será sempre associada ao primeiro fator (MAF1).

A versão dessa técnica usando o método kernel foi utilizada. Simplificadamente, é inclusa uma fase de treinamento, marcada pela amostragem de n pixels das

imagens, definindo-se os dados de treinamento, a partir dos quais são determinados os autovetores. A segunda etapa consiste na fase de generalização, em que cada vetor (pixel) das imagens é projetado ao longo dos autovetores da matriz kernel encontrados na fase anterior. A aplicação de uma função kernel é usada a fim de tratar as não-linearidades dos dados originais através de uma transformação dos mesmos num espaço de grande dimensionalidade. Desta forma, torna-se possível realizar a análise linear dos dados naquele espaço (Nielsen & Canty, 2009).

Dados gamaespectrométricos

Os dados aerogamaespectrométricos que recobrem a área do Prospecto Maués foram adquiridos entre 17 e 30 de janeiro de 2006 pela empresa Lasa Engenharia e Prospecções S.A. contratada da empresa Mapex Mineração. A altura de voo foi mantida constante em 100 m, e as linhas de voo (N-S) e de controle (E-W) consistiram, respectivamente, em 400 m e 4500 m. A aeronave utilizada foi uma Cessna, modelo 208-*Grand Caravan*, com o prefixo PT-MEP. Corresponde a um modelo turbo-hélice monomotor, com velocidade média de 255 km/h. O sistema de navegação abrangeu o uso do GPS Novatel OEM2 (Lasa Engenharia e Prospecções S.A., 2006).

O gamaespectrômetro empregado consistiu no *Exploranium*, modelo GR-820, composto por 256 canais espectrais. Dois conjuntos de cristais voltados para baixo de NaI de 1024 polegadas cúbicas para cada um foram selecionados, além de dois cristais voltados para cima de 256 polegadas cúbicas cada um. Uma leitura foi realizada a cada segundo (Lasa Engenharia e Prospecções S.A., 2006).

Os dados de K (%), eTh (ppm), eU (ppm) e contagem total ($\mu\text{R/h}$) foram interpolados utilizando-se o método da curvatura mínima (5000 iterações), considerando-se um tamanho de célula de 100 m, e micronivelados segundo a metodologia de Minty (1991). A partir das imagens geradas, razões dos radioelementos (eU/K, eU/eTh e eTh/K) foram produzidas, além da normalização do canal

do K $\left[K_n = \frac{K}{(K + eTh + eU)} \right]$ e o cálculo do parâmetro F.

Estas duas últimas imagens foram geradas, pois realçam a unidade metassedimentar (Grupo Jacareacanga), hospedeira de mineralizações auríferas.

Todas as imagens geradas foram re-escaladas para o formato 8 bits e agrupadas em um único arquivo a fim da aplicação das técnicas de estatística multivariada supracitadas para a redução da dimensionalidade destes dados.

Dados auxiliares

Uma faixa de voo do radar de abertura sintética (SAR) da plataforma R99B foi empregada ao estudo. Esta imagem faz parte do acervo de dados da Missão Terra do Meio

Oeste de 2006, e caracteriza-se por 8 looks, ângulo de incidência variando de 45° a 81,10°, visada para oeste, resolução espacial em alcance e azimute de 18 m e tamanho de pixel de 15 m.

Também foram gentilmente disponibilizados pela empresa Mapex Mineração, 166 dados de geoquímica de ouro em rocha.

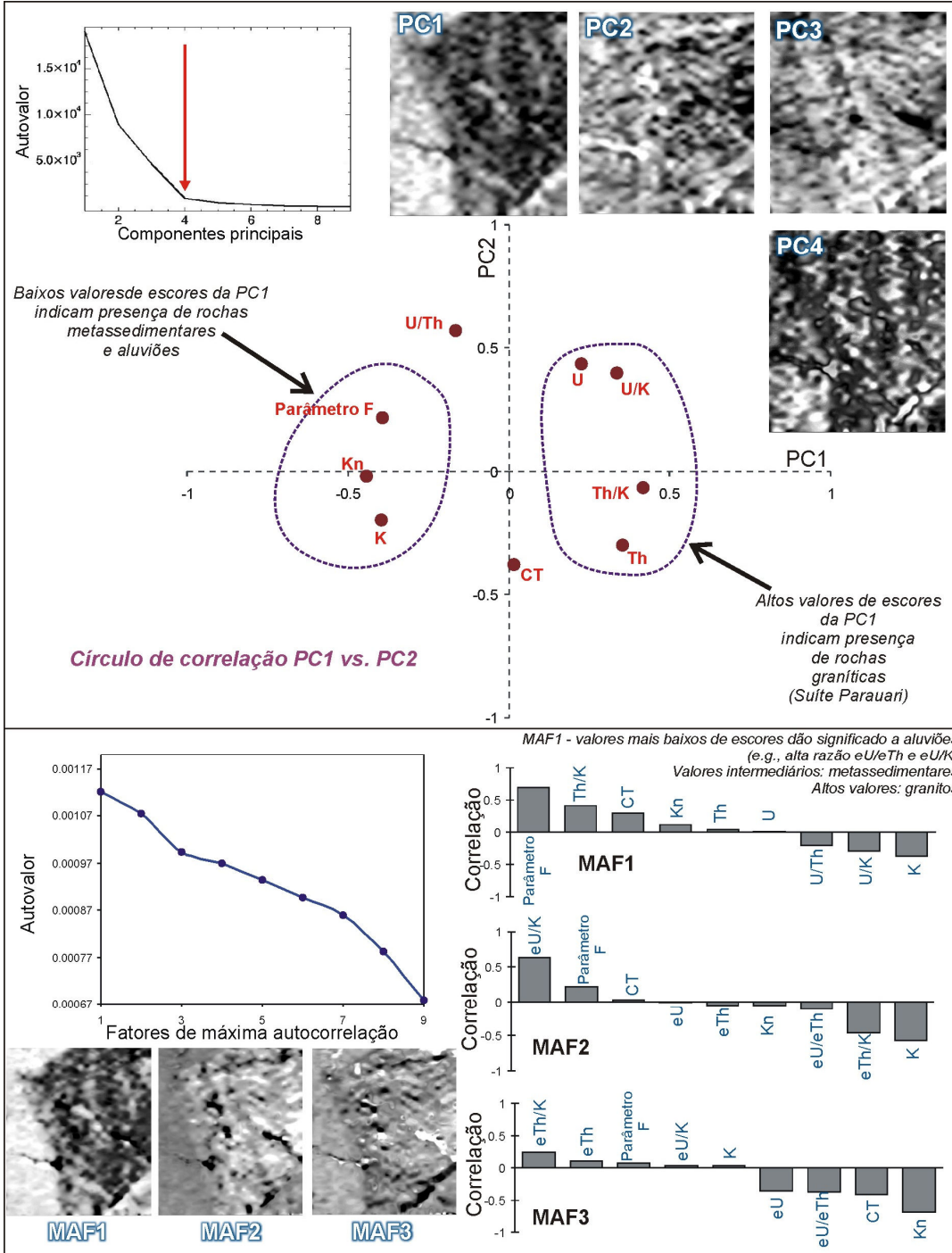


Figura 2. Resultados do processamento das imagens gamaespectrométricas por meio da ACP e MAF. Na parte superior, são apresentadas as imagens das quatro primeiras componentes principais, o gráfico de autovalor e o círculo de correlação da PC1 versus a PC2, sendo possível determinar, na PC1, um domínio de rochas metassedimentares e das aluviões e outro relacionado com granitos Parauari. Na parte inferior, são mostradas as imagens dos três primeiros MAFs, o gráfico de autovalor, e gráficos de correlação das imagens originais com as imagens do MAF1, MAF2 e MAF3.

Aplicação da APC e MAF

As nove imagens aerogamaespectrométricas citadas anteriormente foram processadas pelas técnicas APC

(usando matriz de covariância) e MAF (versão kernel, aplicando-se na etapa de treinamento, uma subamostragem de 2000 pixels). A partir destes

procedimentos, foram gerados gráficos de autovalores e os autovetores (Fig. 2).

Observando os gráficos de autovalores referentes à aplicação das técnicas APC e MAF, é possível verificar que a partir da PC4 o gráfico passa a caracterizar-se por uma curva com inclinação mais suave, indicando que as quatro primeiras PCs são marcadas por maior variabilidade dos dados originais (97,33%) e, assim, tendem a explicar melhor a informação geológica contida nas 9 imagens gamaespectrométricas originais, com perda mínima de informação. Diferentemente, ao se analisar o gráfico de autocorrelação dos MAFs, observa-se que o MAF1 é marcado pela maior correlação das imagens originais, e tende a traduzir o contexto geológico da região do Prospecto Maués.

Analisando-se a imagem da PC1 e, com base no gráfico do círculo de correlação da PC1 vs. PC2, observa-se claramente o domínio dos metassedimentos e aluviões (menores valores de escores) em relação à ocorrência de granitos pós-colisionais (maiores valores de escores). Comparativamente, no caso do MAF1, este contraste de domínios é ainda maior, principalmente para a distinção de aluviões (valores de escores mais baixos) em relação aos metassedimentos (também caracterizados por baixos valores de escores, mas não sendo os mais inferiores). A região dos granitos apresenta-se melhor homogeneizada, se comparada com a imagem da PC1. A variação de DNS (escores) neste domínio na imagem da PC1 reflete a possível variação da composição desta litologia ou mesmo a possível presença de solos que são usualmente bem espessos no Tapajós e caracterizados por altos valores de eTh (eTh/K, etc), pois estão associados a minerais mais resistentes ao intemperismo (e.g., monazita, zircão etc).

Estas observações acima são mais facilmente entendidas quando analisadas as composições coloridas no sistema de cores RGB da PC1, PC2 e PC3 e do MAF1, MAF2 e MAF3. Esclarece-se que, para uma melhor comparação, as imagens das componentes principais e do MAF1 tiveram seus histogramas invertidos no intuito de que as imagens coloridas reproduzissem em tonalidades similares as mesmas unidades geológicas. Estes resultados estão mostrados na Fig. 3. Visualmente, o desempenho da composição colorida MAF1 (-1), MAF2 e MAF3 em RGB é superior ao das PCs visando a caracterização de granitos Parauari (tonalidades azuladas e esverdeadas), de aluviões (predomínio de tonalidades vermelha e magenta) e de metassedimentos do Grupo Jacareacanga (predomínio de cores esbranquiçadas, mas havendo variações de tons em função da composição metassedimentar).

Uma outra maneira de se analisar a informação geológica após a redução de dimensionalidade de dados multivariados, no caso gamaespectrométricos, é a aplicação de um método classificador ou um agrupamento de classes, conforme utilizado por Sancevero (2007) na caracterização de uma seção sísmica. Assim como utilizado por este autor, o classificador não-supervisionado K-médias (MacQueen, 1967) foi selecionado para o processamento das quatro

primeiras componentes principais geradas, e do primeiro MAF, conforme indicado na Fig.3. Neste processamento, foram determinadas a geração de 2 classes para traduzir, de maneira genérica, a ocorrência de metassedimentos do Grupo Jacareacanga e dos granitos Parauari. Analisando-se a Fig. 3, verifica-se que o resultado derivado da aplicação da técnica K-médias na imagem do MAF1 permite uma melhor definição do contato dos granitos com as rochas metassedimentares paleoproterozóicas. Este resultado, mais uma vez, reforça a potencialidade da técnica MAF para aplicação geológica num conjunto de dados aerogamaespectrométricos.

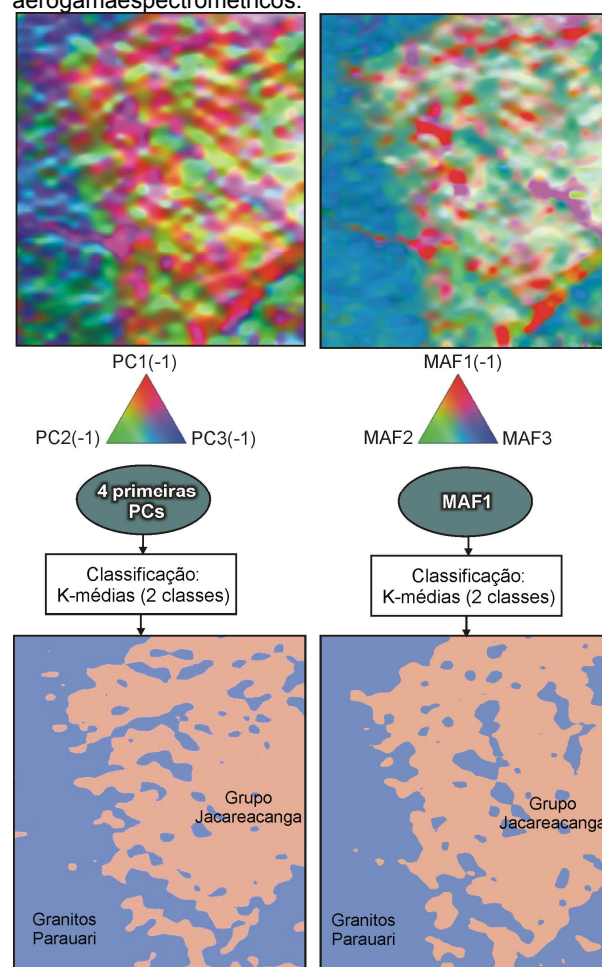


Figura 3. Composições coloridas das três primeiras componentes principais e dos três primeiros fatores de máxima autocorrelação (parte superior). Na parte inferior, são apresentados os resultados da aplicação do classificador K-médias nas quatro primeiras componentes principais e na imagem do MAF1, permitindo a distinção das duas principais unidades litológicas da região do Prospecto Maués.

Selecionando-se a composição colorida MAF1(-1), MAF2 e MAF3 no espaço de cores RGB, procedeu-se à integração com a imagem de radar de abertura sintética da plataforma aérea R99B, banda L, polarização HH

usando-se o método IHS. Este resultado, indicado da Fig. 4, permite analisar conjuntamente as informações litológicas derivadas dos MAFs e da macrotopografia proveniente da imagem de radar. Nota-se que as anomalias de ouro localizam-se em rochas do Grupo Jacareacanga e, preferencialmente, próximas à interface dos granitos Parauari.

Também, com base no produto integrado da Fig. 4, é possível interpretar que lineamentos de drenagem e, conseqüentemente, alinhamentos de drenagem de direção NE-SW são bem marcados nos domínios de rochas metassedimentares, e pouco visíveis nos granitos Parauari, estes caracterizados por macrotopografia mais suavizada e menos estruturada. Este *trend* NE-SW pode ser interpretado como um registro (posteriormente reativado?) da fase tectono-evolutiva de ~2,0 Ga sugerida no PROMIN-Tapajós (Klein et al., 2001). A estruturação NW-SE, posterior (1,99 a 1,60 Ga, de acordo com o PROMIN-Tapajós), também encontra-se bem marcada em toda a área de estudo e, de acordo com Coutinho (2008), faz parte do registro paleoproterozóico do *trend* aurífero Patrocínio-Espírito Santo.

Todas estas caracterizações advindas da aplicação da técnica MAF, somando-se às interpretações conjuntas com a imagem de radar, permitem melhorar o entendimento geológico da região do Prospecto Maués, em relação ao mapa geológico disponível para esta área (cf., Fig. 1).

Conclusões

A aplicação de técnicas de estatística multivariada em conjunto de dados, como os aerogamaespectrométricos, pode prover a simplificação dos dados originais visando um melhor entendimento da informação contida nas variáveis analisadas. Neste artigo, utilizando-se as técnicas ACP e MAF fica evidente que os resultados gerados foram muito satisfatórios para a distinção litológica no que tange este último método, tanto por análise visual de composições coloridas, quanto pela aplicação de métodos como a classificação não-supervisionada K-médias e a transformação IHS com imagem de radar imageador. O uso da técnica MAF abre, portanto, novas perspectivas de processamento de dados multivariados no campo da Geologia, como o tratamento de imagens aerogeofísicas, podendo ser estendido ao processamento do grande número de imagens/bandas espectrais de sensores hiperespectrais, de dados de geoquímica de rocha e solo, e de radar de abertura sintética multipolarizado em terrenos geologicamente expostos, a fim de se analisar as respostas do retroespalhamento características de diferentes unidades geológicas, por exemplo.

Agradecimentos

À UnB pelo suporte técnico e ao CNPq (CT-Mineral, processo n°. 555063/2006-2). A. M. Silva e N. F. Botelho agradecem ao CNPq pelas respectivas bolsas de

produtividade em pesquisa. T.A. Carrino agradece a S.S. Sancevero por propor o uso da técnica MAF.

Referências bibliográficas

- Canty M. 2010. Disponível em: <<http://mcanty.homepage.t-online.de>>. Acesso em 15 de novembro de 2010.
- Coutinho M.G.N. (ed.) 2008. *Província Mineral do Tapajós: geologia, metalogenia e mapa preliminar para ouro em SIG*. Rio de Janeiro, CPRM, 402p.
- Hair J.F., Tatham R.L., Anderson R.E., Black W. 1998. Chapter 1 - Introduction. In: Hair J.F., Tatham R.L., Anderson R.E., Black W. (Ed.). *Multivariate Data Analysis*. 5 ed., Nova Jersey, Pearson Education Inc., 23-45.
- Klein E.L., Almeida M.E., Vasquez M.L., Bahia R.B.C., Santos M.L.E., Ferreira A.L. 2001. Geologia e recursos minerais da Província Mineral do Tapajós. Folhas Vila Mamães Anã (SB.21-V-D), Jacareacanga (SB.21-Y-B), Caracol (SB.21-X-X), Vila Riozinho (SB.21-Z-A) e Rio Novo (SB.21-Z-C). Estados do Pará e Amazonas. Escala 1:500.000. Brasília: CPRM/DIEDIG/DEPAT, 2001, 81 p.
- Larsen R. 2002. Decomposition using Maximum Autocorrelation Factor. *Journal of Chemometrics*, v. 16(8-10), 427-435.
- Lasa Engenharia e Prospecções S.A. 2006. Levantamento aeromagnetométrico gradiométrico e aerogamaespectrométrico. Blocos Bandeirante, Maués e Doze de Outubro: relatório final de aquisição e processamento de dados. Volume 1. Textos técnicos e anexos. 41 p.
- MacQueen J. 1967. Some methods for classification and analysis of multivariate observations. In: University of Blerkely, Blerkely Symposium on Mathematical Statistics and Probability, 5th, Blerkely, 325-350.
- Minty B.R.S. 1991. Simple micro-levelling for aeromagnetic data. *Exploration Geophysics*, 22: 591-592.
- Nielsen A.A., Conradsen K., Simpson J.J. 1998. Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: new approaches to change detection studies. *Remote Sensing of Environment*, v. 64, 1-19.
- Nielsen, A.A. & Canty, M. 2009. Kernel principal component and maximum autocorrelation factor analyses for change detection. *Image and Signal Processing for Remote Sensing XV*, SPIE, vol. 7477, 74770T-1-74770T-6.
- Nielsen, A. A. & Olsson, H. 2010. Change detection by the IR-MAD and Kernel MAF methods in Landsat TM data covering a Swedish forest region. *Forestsat*, 17p.
- Sancevero S.S. 2007. Estudo de aplicação de métodos quantitativos em dados sísmicos no processo de caracterização integrada de reservatórios. Tese de

doutorado em Geociências, Universidade Estadual de Campinas, Campinas-SP, 220p.

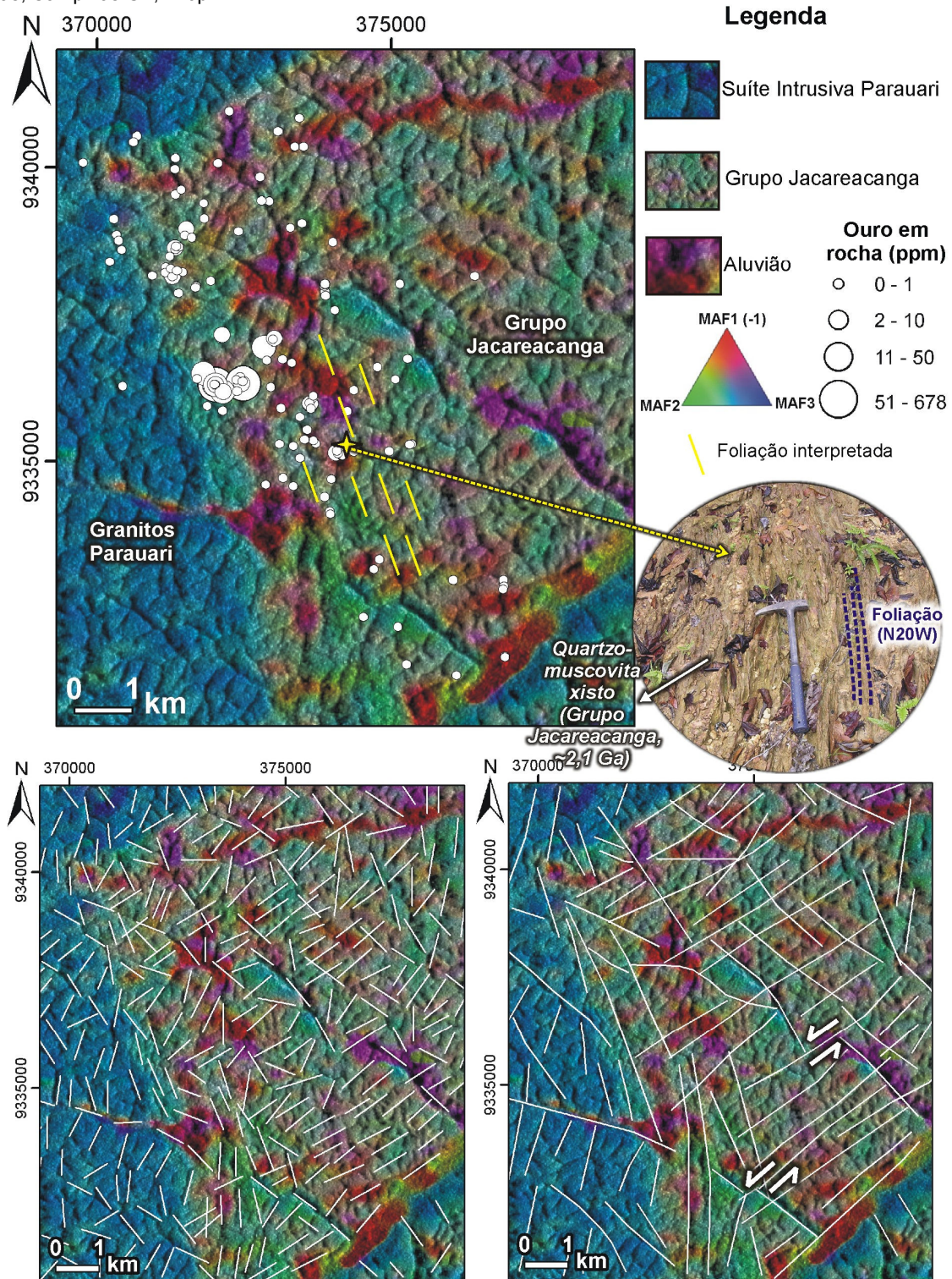


Figura 4. Composição colorida do MAF1 (histograma invertido), MAF2 e MAF3 em RGB integrada com a imagem de radar R99B, band L, polarização HH. As principais unidades geológicas e principais lineamentos (imagem inferior esquerda) e alinhamentos de drenagem (imagem inferior direita) estão indicados.