

## **Aplicação das técnicas ACP (Análise por Componentes Principais), Kernel ACP e Kernel MAF (*Maximum Autocorrelation Factor*) em dados aerogamaespectrométricos da Província Mineral do Tapajós: resultados iniciais para estudos geológicos**

Thais Andressa Carrino <sup>1</sup>  
Adalene Moreira Silva <sup>1</sup>  
Nilson Francisquini Botelho <sup>1</sup>  
Marcelo Pinto <sup>2</sup>  
Luiz Carlos Pinheiro Clerot <sup>2</sup>

<sup>1</sup> Instituto de Geociências – Universidade de Brasília  
70910-900 – Brasília-DF, Brasil  
thais.carrino@gmail.com; {adalene, nilsonfb}@unb.br

<sup>2</sup> MAPEX - Mineração, Importação e Exportação Ltda.  
Av. Constantino Nery, 2552, Manaus-AM  
marcelopinto@mapex.com.br; lclerot@yahoo.com.br

**Abstract.** The Kernel ACP and Kernel Maximum Autocorrelation Factor (MAF) techniques were applied to an airborne geophysics dataset for the assessment of the geologic information in the area of the 12 de Outubro Prospect. These methods are characterized as other approach of the well known Principal Components Analysis technique that also was utilized in this study case. Based on the visual analysis, the results were compared and the performances of the application of the Kernel ACP and Kernel MAF techniques were of better quality than ACP technique for the recognition of geologic units in Tapajós Mineral Province. This case study shows that the use of the Kernel ACP and Kernel MAF has a great potential of application in multivariate geologic data.

**Palavras-chave:** Análise por Componentes Principais, Fator de Máxima Autocorrelação, método Kernel, geologia; Principal Components Analysis, Maximum Autocorrelation Factor, Kernel method, geology.

### **1. 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. Neste sentido, este artigo explora o uso das técnicas Análise por Componentes Principais (ACP), Kernel ACP e Kernel MAF (*Maximum Autocorrelation Factor*), visando o realce de unidades litológicas a partir de um conjunto de dados aerogamaespectrométricos do prospecto 12 de Outubro, localizado na Província Mineral do Tapajós (AM, PA). Esta região do Tapajós é caracterizada por rochas vulcânicas ácidas, como riolitos do Grupo Iriri que são hospedeiros de mineralizações de ouro, além de rochas vulcânicas intermediárias pertencentes à Formação Bom Jardim. Estas duas unidades geológicas possuem idade paleoproterozóica (~1,8 Ga), assim como as mineralizações ouro (Klein et al., 2001).

### **2. Dados gamaespectrométricos**

Os dados aerogamaespectrométricos que recobrem a área do prospecto 12 de Outubro foram adquiridos entre 17 e 30 de Janeiro de 2006 pela empresa Lasa Engenharia e Prospecções S.A. contratada da empresa Mapex Ltda. A altura de vôo foi mantida constante em 100 m, e as linhas de vôo (N-S) e de controle (E-W) variaram, respectivamente, em 400 m e 2300 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 (K/eTh, K/eU e eU/eTh) foram produzidas.

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

### **3. Técnicas de redução da dimensionalidade de dados multivariados: ACP, MAF, Kernel ACP e Kernel MAF**

A Análise por Componentes Principais (ACP) compreende uma abordagem 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 análise por MAF (*Maximum Autocorrelation Factor* – Fator de Máxima Autocorrelação) é comumente aplicada em geociências e em biologia (Larsen, 2002) para a redução de informações de um conjunto de dados para a detecção de mudanças (Nielsen et al., 1998; Nielsen & Canty, 2009; Nielsen & Olsson, 2010), para a caracterização de reservatório (e.g., Sancevero, 2007), assim como para a 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). A partir da análise por MAF, 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) é marcado por baixa autocorrelação. Desta forma, 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).

A análise por MAF 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 ACP encontra combinações lineares ortogonais com máxima variância, enquanto a técnica MAF visa a obtenção de combinações lineares ortogonais marcadas por máxima autocorrelação entre observações vizinhas (e.g., pixels). Por tal razão, ao se estudar imagens de sensoriamento remoto, o método MAF tende a possuir um desempenho melhor. Maiores informações sobre a fundamentação teórica da técnica MAF é mostrada em Larsen (2002).

As versões dessas duas técnicas usando o método kernel foram disponibilizadas por Canty (2010) em linguagem IDL (*Interactive Data Language*) para uso no *software* ENVI. Simplificadamente, para as técnicas Kernel ACP e Kernel MAF, é incluída uma fase de treinamento, marcada pela amostragem de pixels das imagens da ordem de  $n=1000$ , definindo-se os *dados de treinamento*, a partir dos quais são definidos os autovetores. A segunda etapa consiste na fase de generalização, em que cada vetor (pixel) das imagens é projetado ao longo dos autovetores encontrados na fase anterior. A aplicação de uma função kernel é usada a fim de tratar as não-linearidades dos dados 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).

A Figura 1 ilustra as etapas adotadas neste artigo para a análise dos resultados obtidos com o emprego das técnicas ACP, Kernel ACP e Kernel MAF, avaliando-se visualmente os resultados e composições coloridas, além da integração com dados geológicos coletados durante campanha de campo realizada em setembro de 2008.

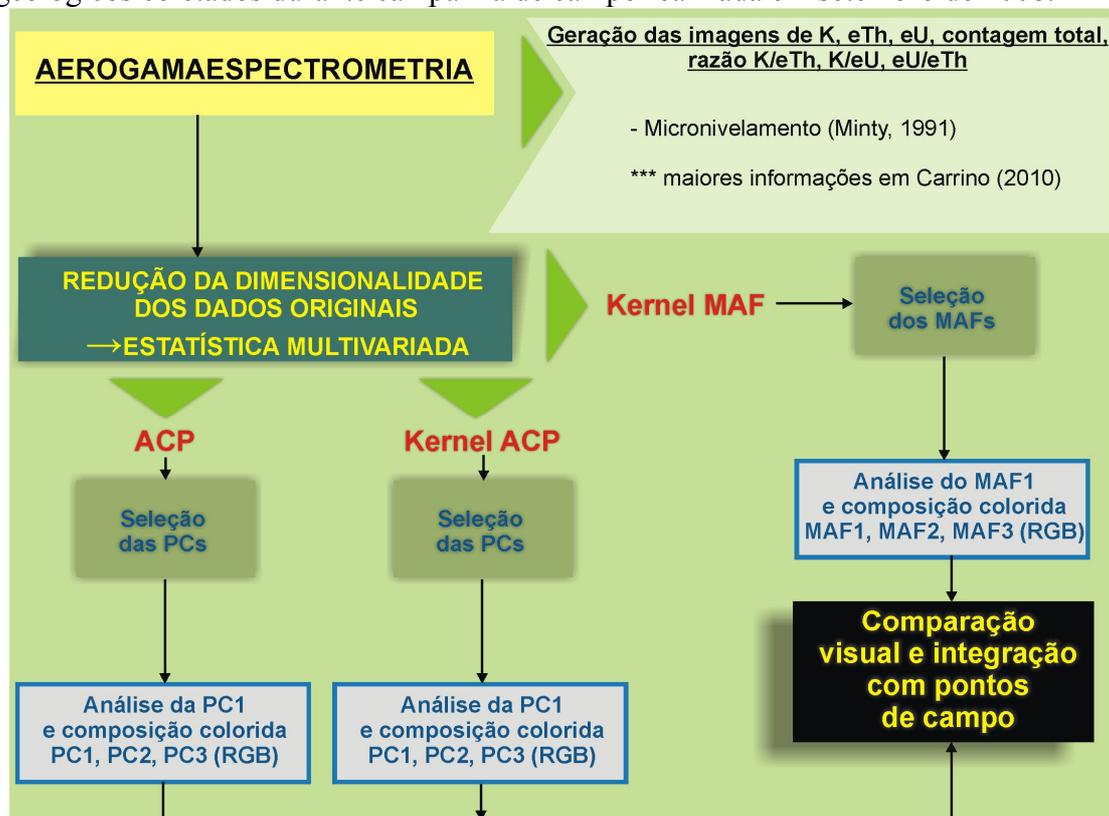


Figura 1. Resumo das etapas de aplicação das técnicas ACP, Kernel ACP e Kernel MAF nos dados aerogamaespectrométricos do prospecto 12 de Outubro.

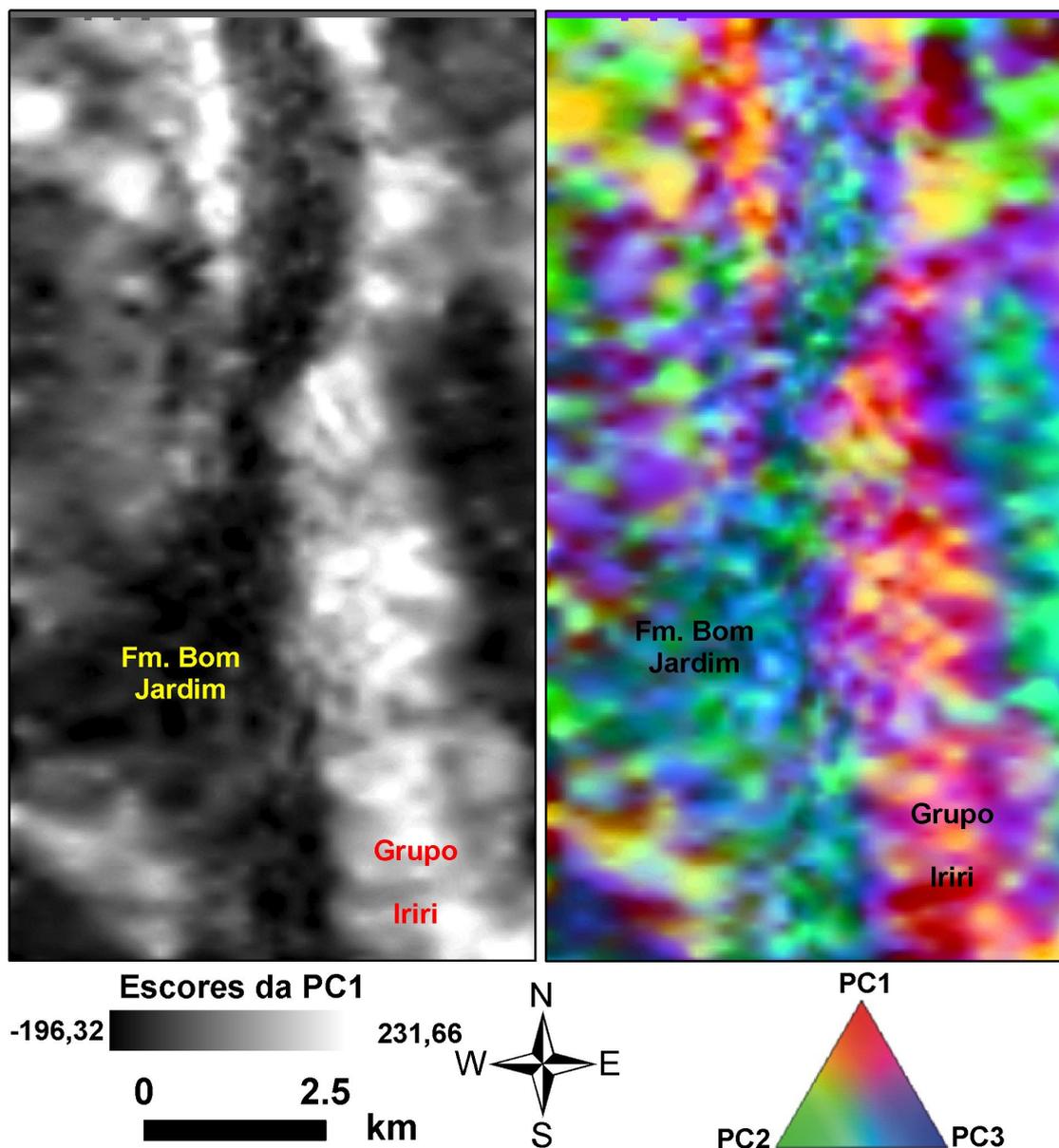


Figura 2. Imagem da primeira componente principal (PC1) e composição colorida no sistema de cores RGB da PC1, PC2 e PC3.

#### 4. Resultados

Os resultados gerados da aplicação das técnicas ACP, Kernel ACP e Kernel MAF estão apresentados nas Figuras 2, 3 e 4, respectivamente. Na Figura 5, os resultados produzidos pelas três técnicas supracitadas são comparados e integrados com dados geológicos coletados em campo.

Analisando-se as imagens da primeira componente principal (PC1) da Figura 2 com a PC1 invertida gerada pela técnica Kernel ACP (Figura 3) e o primeiro fator de máxima autocorrelação (MAF1) mostrado na Figura 4, é possível verificar que todas, em geral, permitem um reconhecimento similar de rochas vulcânicas ácidas do Grupo Iriri, caracterizadas por altos escores. Entretanto, a imagem do MAF1 compreende o resultado mais satisfatório para delinear as duas principais litologias da região do prospecto 12 de Outubro, a Formação Bom Jardim (baixos escores) e o Grupo Iriri, seguida da imagem da PC1 invertida, resultante da aplicação da técnica Kernel ACP,

pois ambas as técnicas geraram resultados mais contrastantes se comparados com a PC1 da Figura 2 (aplicação da técnica ACP).

Essa observação é mais facilmente analisada nas composições coloridas falsa-cor apresentadas nas Figuras 2, 3, 4 e 5. O resultado da ACP (Figuras 2 e 5) apresenta pouca definição dos limites das unidades Formação Bom Jardim e Grupo Iriri, diferentemente do produto gerado pela técnica Kernel ACP (Figuras 3 e 5). Ambos os resultados mostram ainda variações dentro de cada domínio geológico devido às ocorrências de espessos perfis de solos (principalmente nas rochas do Grupo Iriri) e presença de rochas máficas conjuntamente às rochas vulcânicas intermediárias da Formação Bom Jardim, o que foi verificado em campo, conforme mostrado na Figura 5.

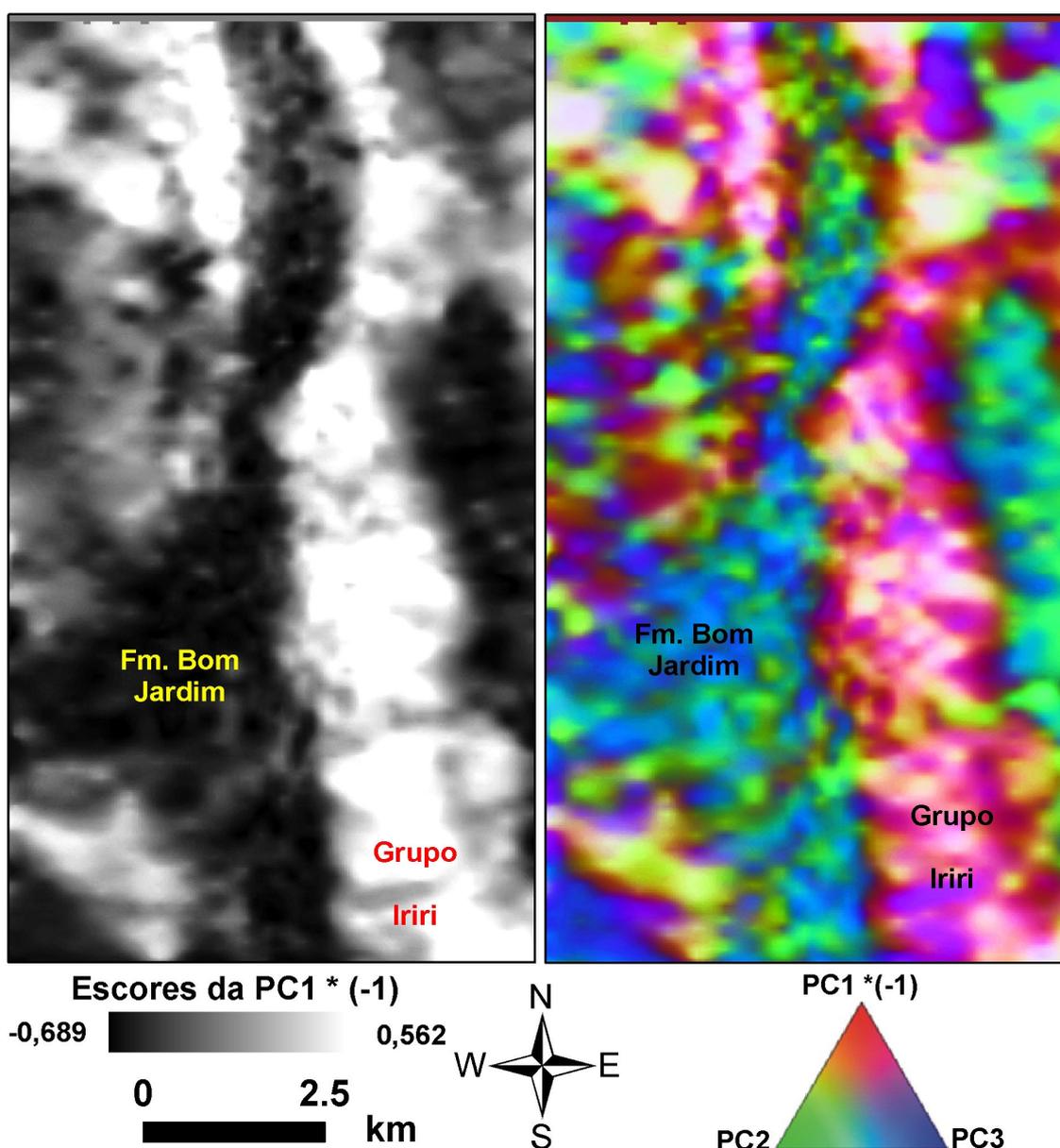


Figura 3. Imagem da primeira componente principal (PC1) gerada pela técnica Kernel ACP e composição colorida no sistema de cores RGB da PC1, PC2 e PC3.

Por sua vez, o resultado da composição colorida falsa-cor dos MAF1, 2 e 3 no espaço de cores RGB (Figuras 4 e 5) permite uma análise mais geral dos dois domínios

geológicos em questão, suprimindo de maneira mais drástica, a resposta de regolitos e/ou assembléias de rochas máficas de pequena extensão espacial. Este resultado é interessante por atenuar as respostas de solos que se desenvolvem em espessos perfis na região amazônica e que, devido à sua resposta superficial, é registrado por meio do uso de sensores gamaespectrométricos e passam a obliterar a resposta de rochas, algumas das quais podem estar mineralizadas (Wilford et al., 1997; IAEA, 2003).

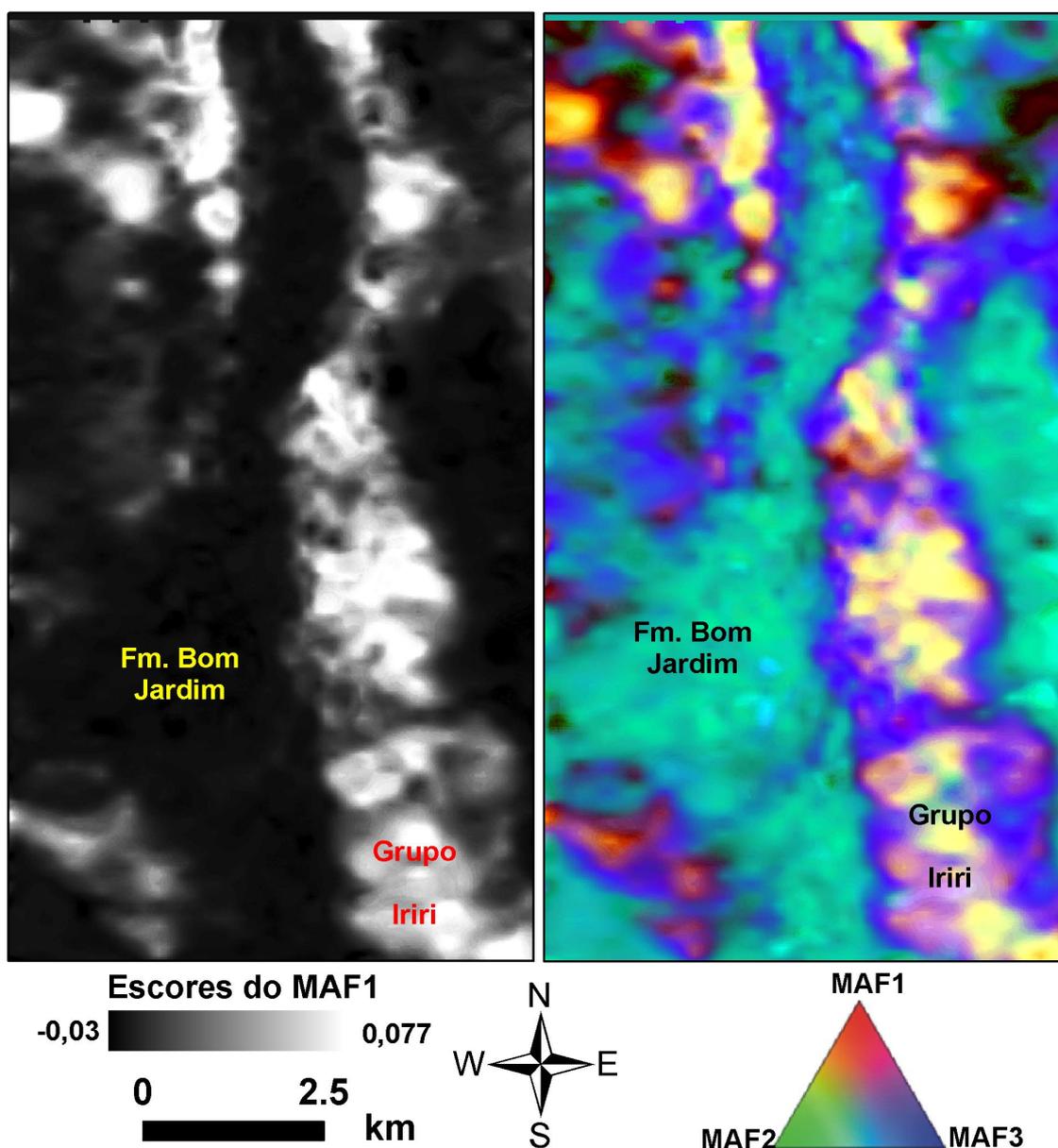


Figura 4. Imagem do primeiro fator de máxima autocorrelação (MAF1) e composição colorida no sistema de cores RGB do MAF1, MAF2 e MAF3.

## 5. Conclusões

A aplicação de técnicas estatísticas multivariadas em conjunto de dados, como os aerogamaespectrométricos, podem prover a simplificação dos dados originais visando uma melhor observação das variáveis analisadas. Neste presente artigo, utilizando-se as técnicas ACP, Kernel ACP e Kernel MAF fica evidente que os resultados gerados foram muito satisfatórios para as duas últimas, sendo que para a tão popular ACP não

ficou claro, visualmente, a distinção das duas principais unidades geológicas da região do prospecto 12 de Outubro (Formação Bom Jardim e Grupo Iriri).

O uso das técnicas Kernel ACP e Kernel MAF abre, portanto, novas perspectivas de processamento de dados multivariados no campo da Geologia, como o tratamento 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 analisar as respostas do retroespalhamento características de diferentes unidades geológicas.

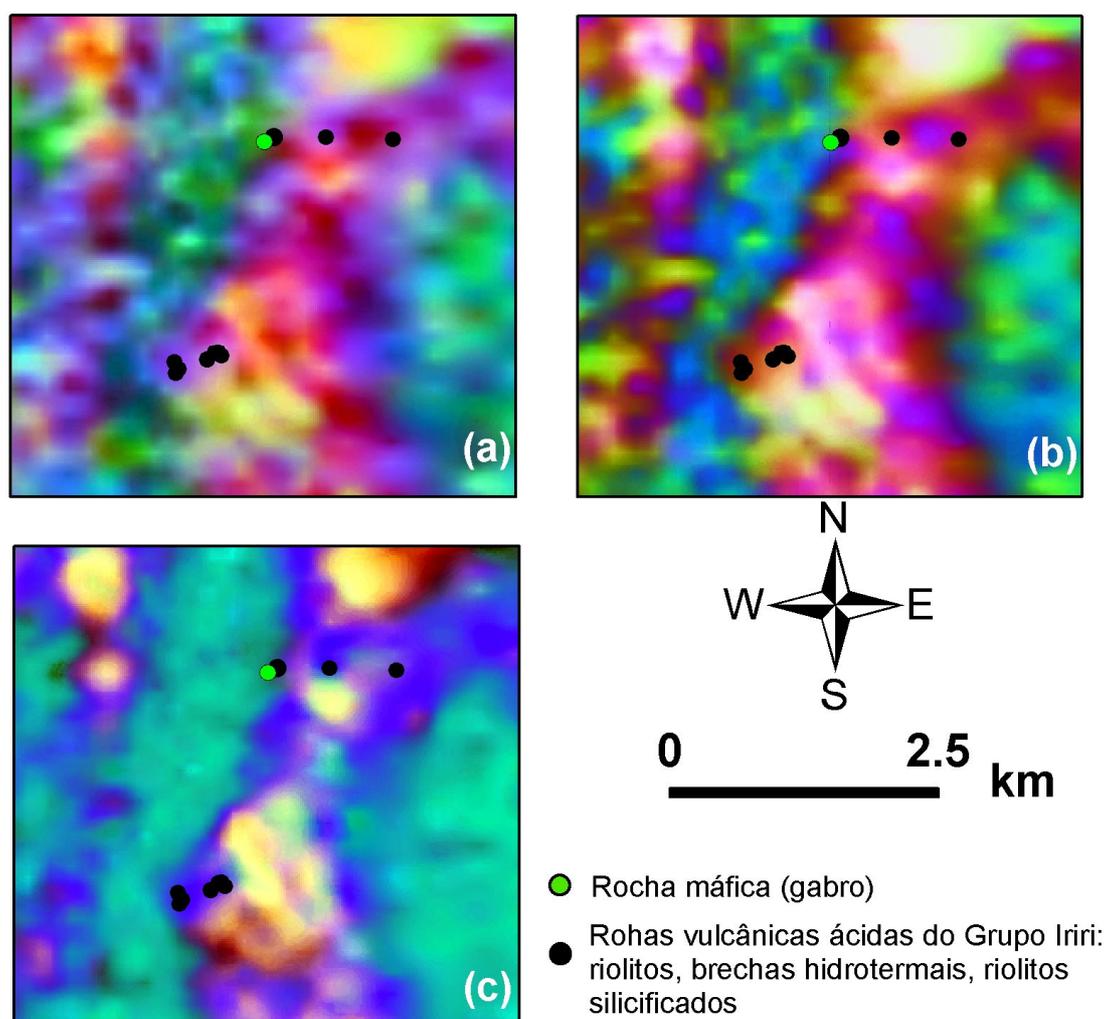


Figura 5. Integração das imagens em composição colorida RGB das três técnicas abordadas neste artigo com pontos de campo: (a) ACP (PC1, PC2, PC3 em RGB), (b) Kernel ACP (PC1 (invertida), PC2, PC3 em RGB), (c) Kernel MAF (MAF1, MAF2, MAF3 em RGB).

#### Agradecimentos

Agradecemos à UnB pelo suporte técnico e ao CNPq pelo projeto CT-Mineral (nº. 555063/2006-2). A.M. Silva e N.F. Botelho agradecem ao CNPq pelas respectivas bolsas de produtividade em pesquisa.

#### Referências bibliográficas

Canty M. Disponível em: <<http://mcanty.homepage.t-online.de>>. Acesso em 15 de novembro de 2010.

Carrino, T.A. **Geotecnologias aplicadas ao reconhecimento de áreas chaves à exploração aurífera na Província Mineral do Tapajós**. Dissertação de mestrado em Geologia, Universidade de Brasília, Brasília-DF, 192 p., 2010.

Hair J.F., Tatham R.L., Anderson R.E., Black W. 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, 1998.

IAEA (International Atomic Energy Agency) **Guidelines for radioelement mapping using gamma ray spectrometry data**. Viena, IAEA-TECDOC, 173p., 2003.

Klein E.L., Almeida M.E., Vasquez M.L., Bahia R.B.C., Santos M.L.E., Ferreira A.L. 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, 2001.

Lasa Engenharia e Prospecções S.A. 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, 2006.

Minty B.R.S. Simple micro-levelling for aeromagnetic data. **Exploration Geophysics**, v. 22, 591-592, 1991.

Larsen R. 2002. Decomposition using Maximum Autocorrelation Factor. **Journal of Chemometrics**, v. 16(8-10), 427-435.

Nielsen A.A., Conradsen K., Simpson J.J. 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, 1998.

Nielsen, A.A. & Canty, M. 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, 2009.

Nielsen, A. A. & Olsson, H. Change detection by the IR-MAD and Kernel MAF methods in Landsat TM data covering a Swedish forest region. **Forestsat**, 17p., 2010.

Sancevero S.S. **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., 2007.

Wilford P.N., Bierwirth P.N., Craig M.A. Application of airborne gamma-ray spectrometry in soil/regolith mapping and applied geomorphology. **AGSO Journal of Australian Geology & Geophysics**, v.17, n.2, 201-216, 1997.