Principal Component and K-means Analysis of Airborne Gamma Ray Spectrometry Surveys

Rafael Pires de Lima[1,2] and Dr. Kurt J. Marfurt[1]*
The University of Oklahoma[1] and CPRM-Geological Survey of Brazil[2]
rlima@ou.edu, kmarfurt@ou.edu

The University of Oklahoma
Mewbourne College of Earth and Energy

- Airborne geophysical data are integrated with ground control to map:
  - Uranium and potash deposits (Zhang et al., 1998), and
  - Shallow geology (Carneiro et al., 2012; Cracknell and Reading, 2014; Harris et al., 2015).
- Traditionally, the Geological Survey of Brazil (CPRM) interpreters manual integrate airborne geophysical data with surface outcrop and geochemical data to better map the natural resources and risks of the country. However,
  - Interpreter integration of airborne gamma-spectrometry maps with geologic outcrop data can be very time-consuming, and
  - The resulting maps can be quite subjective, depending on the experience of the interpreter.
- We apply two techniques to accelerate and standardize airborne gamma ray geophysical interpretation:
  - K-means clustering algorithm (Hartigan and Wong, 1979)
  - Principal component analysis (PCA)
- Using total count (TC), potassium (K), equivalent thorium (eTh), and equivalent uranium (eU) we find that
  - TC, K, eTh, and eU K-means clusters generates maps that are similar but more noisy than traditional interpreter-generated geologic maps
  - Principal components 1 and 2 K-means clusters provide less noisy maps that are better correlated to those generated using traditional interpreter-generated methods.

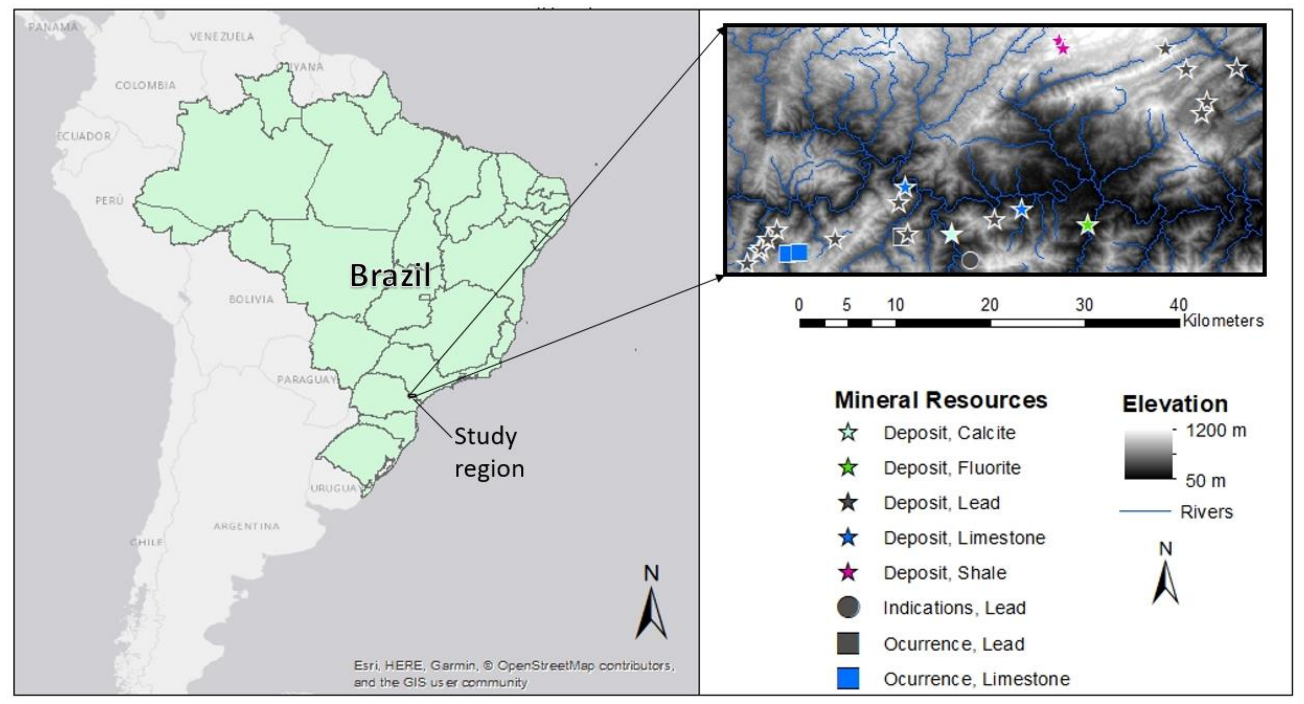

Figure 1: Area of study with mapped mineral resources. The reference system used in this paper is the geographic coordinate system SIRGAS 2000. Background location map from OpenStreetMap, Brazilian state boundaries provided by The Brazilian Institute of Geography and Statistics, geologic mineral resources information courtesy of CPRM, elevation model courtesy of NASA's Shuttle Radar Topography Mission (SRTM).

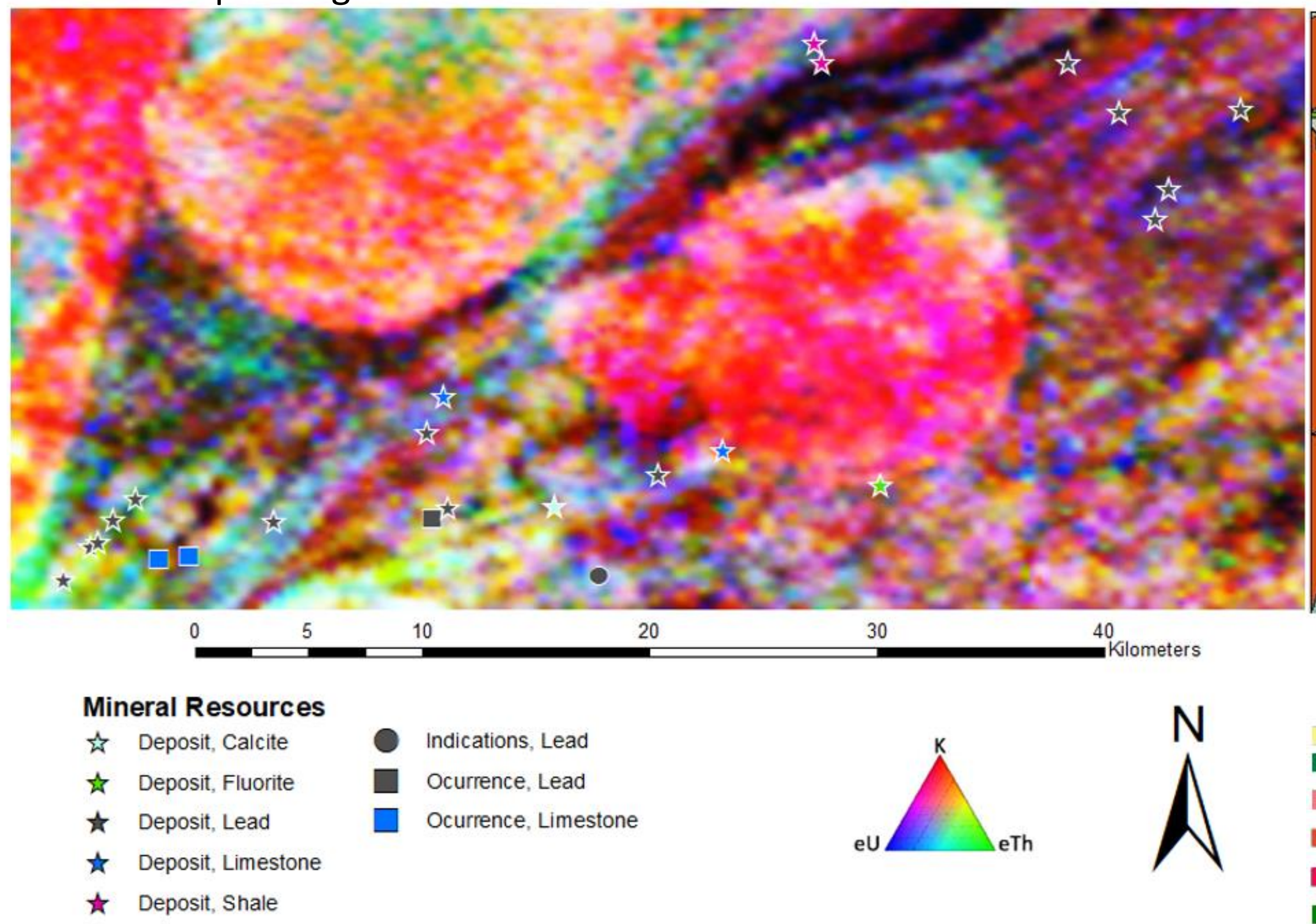

Figure 2: Ternary image of the study area created by displaying potassium (K) against red, equivalent thorium (eTh) against green, and equivalent uranium (eU) against blue. Stars indicate commercial and potential mineral resources. When all elements are present, the map location appears white. When none of the elements are present the map location appears black.

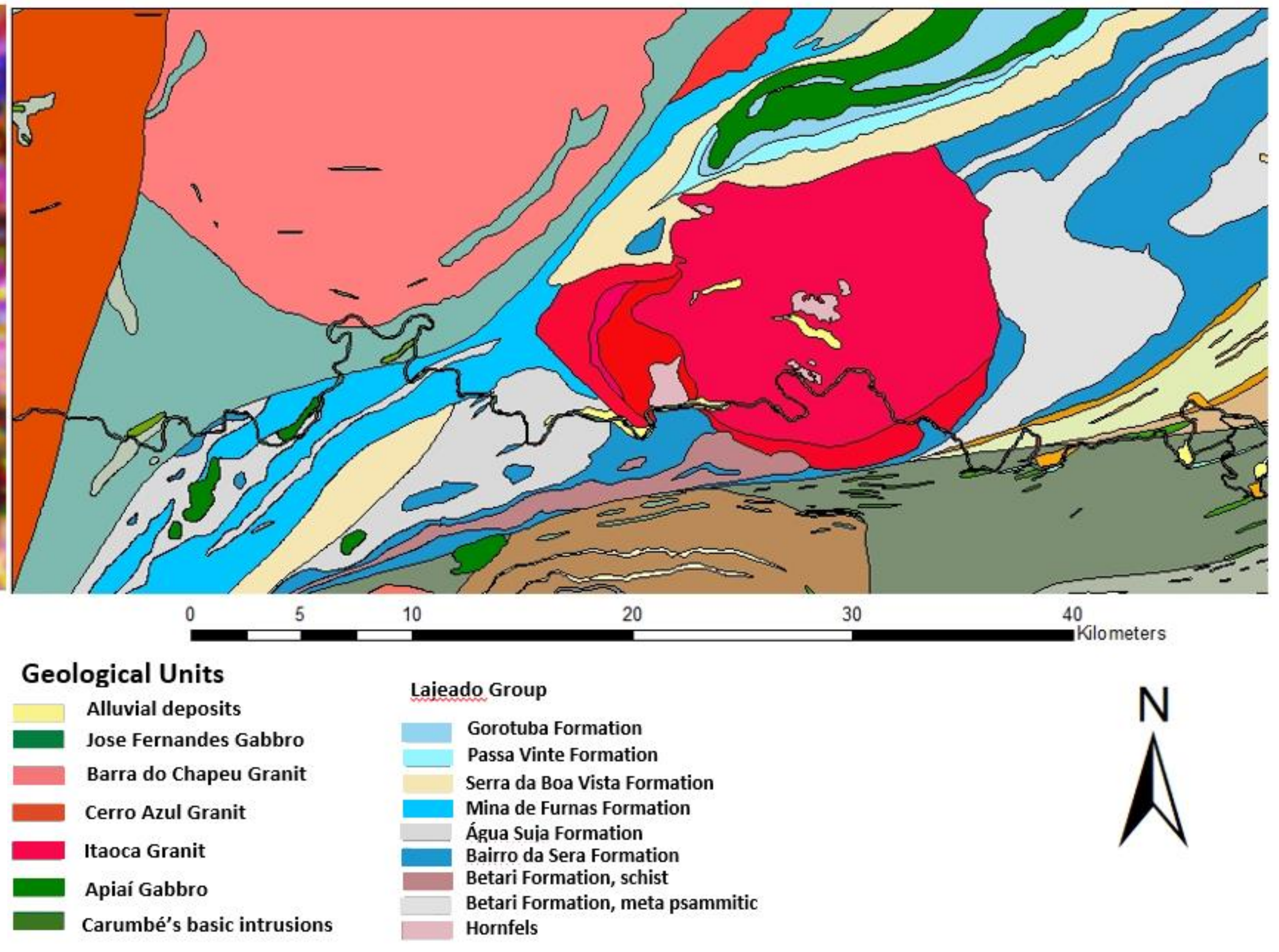

Figure 3: Geologic map of the study area simplified from Lopes et al. (2017). Note the visual correlation with Figure 2, such as the Itaoca Granite ("red" in both images), and the Barra do Chapeu Granite ("pink" in both images).

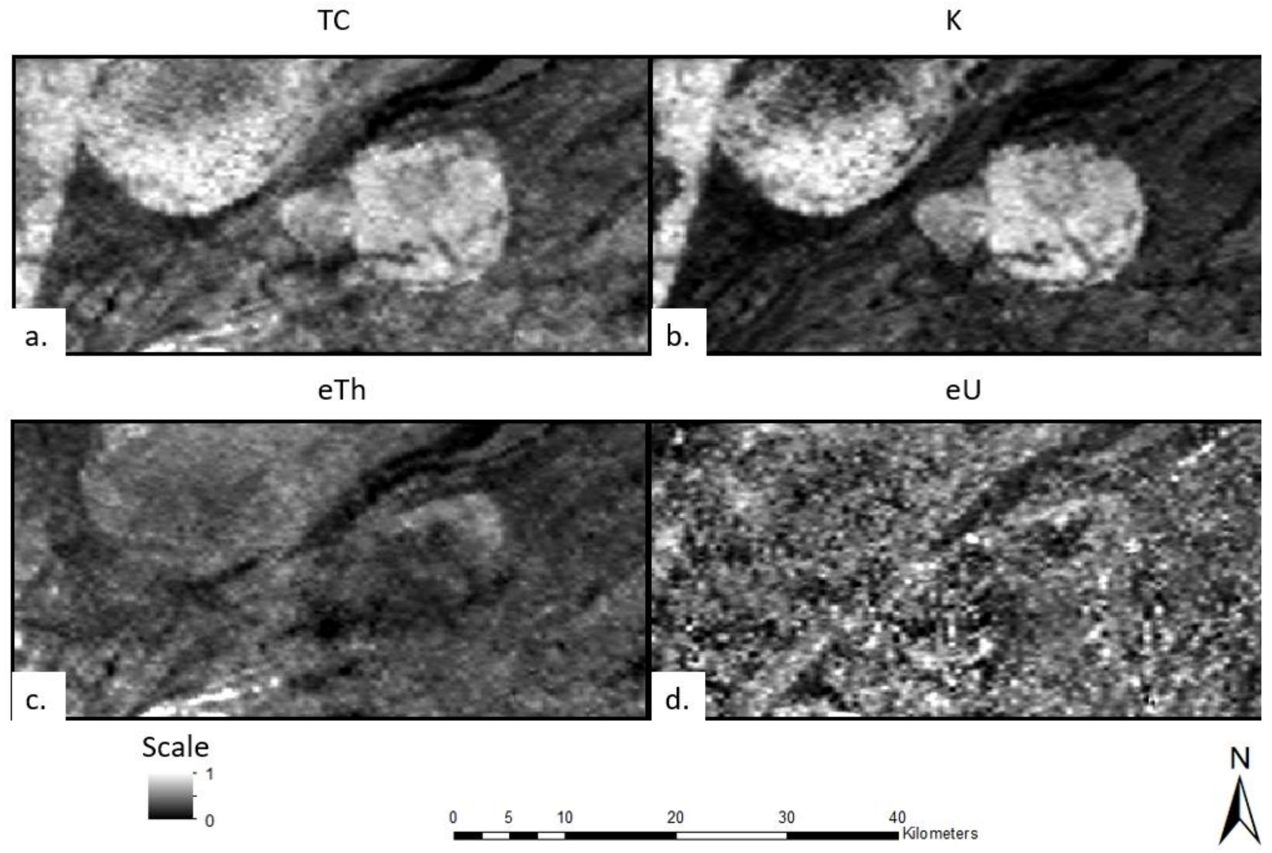

Figure 4: Linearly scaled (a) TC, (b) K, (c) eTh, and (d) eU. Note the variability between each display, with several anomalies present in all four images and that the eU panel exhibits a higher level of short wavelength noise than the other images. The color bar is the same for all panels as the data was scaled to range from 0 to 1.

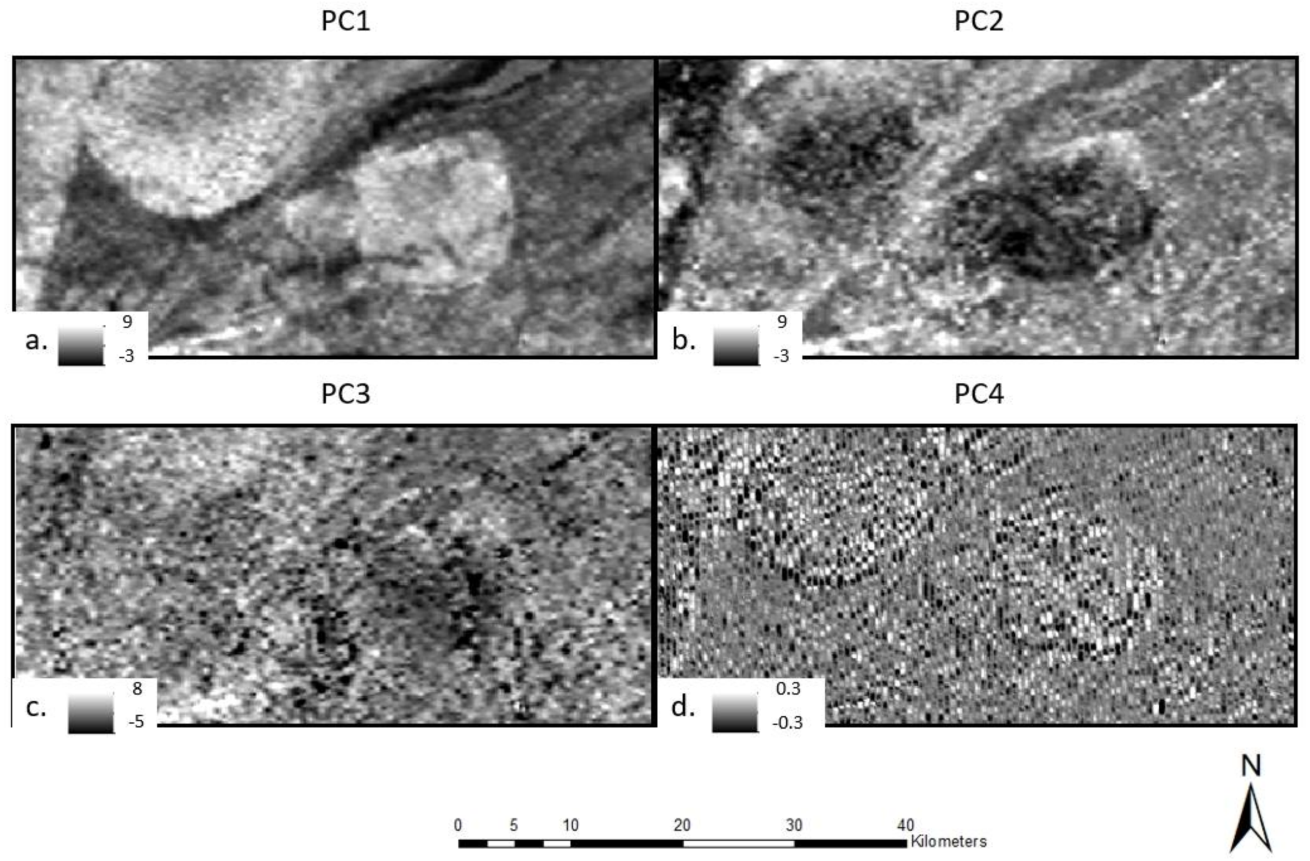

Figure 5: Principal components (a) PC1, (b) PC2, (c) PC3, and (d) PC4 computed from the data shown in Figure 4. The PC1 result most resembles the traditional geologic map shown in Figure 3 and the ternary K-eTh-eU shown in Figure 2. PC2 still has some possible geological information left. PC3 and PC4 contain mostly random and acquisition footprint noise such as the N-S flightline trend in (d).

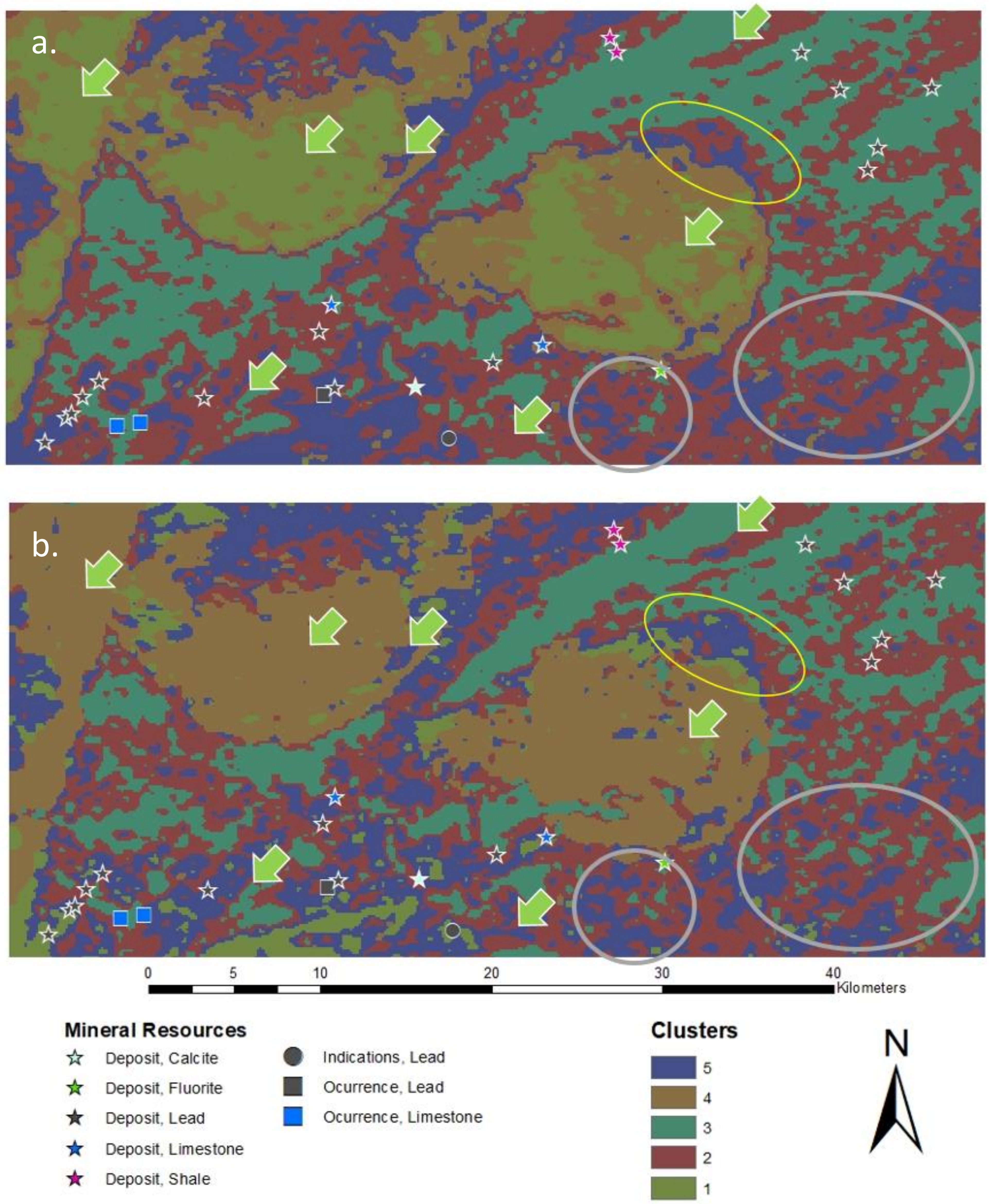

Figure 6: K-means clustering results. a. uses TC, K, eTh, and eU rasters as input while b. uses PC 1 and PC 2 as input . Although the clusters are arbitrarily set during the calculations and have no direct relation between them, it is possible to compare the results obtained in both panels. Green arrows indicate improvements in clustering (either less noisy or better separation) between the results in b. Gray ellipses show some examples of mapped geological boundaries that are inaccurately imaged by the gamma-spectrometry survey, and therefore are not properly clustered using either technique. The yellow ellipse highlights a region with different spectral characteristics that has not been identified as being a separate lithology through geologic surface mapping. Such variations may indicate different geological processes that have been previously overlooked.

References: Carneiro, C. de C., S. J. Fraser, A. P. Crósta, A. M. Silva, and C. E. de M. Barros, 2012, Semiautomated geologic mapping using self-organizing maps and airborne geophysics in the Brazilian Amazon: Geophysics, **77**, no. 4, K17–K24, doi:10.1190/geo2011-0302.1.

Cracknell, M. J., and A. M. Reading, 2014, Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information: Computers & Geosciences, **63**, 22–33, doi:10.1016/J.CAGEO.2013.10.008.

Harris, J. R., E. Grunsky, P. Behnia, and D. Corrigan, 2015, Data- and knowledge-driven mineral prospectivity maps for Canada's North: Ore Geology Reviews, **71**, 788–803, doi:10.1016/J.OREGEOREV.2015.01.004.

Hartigan, J. A., and M. A. Wong, 1979, Algorithm AS 136: A K-Means Clustering Algorithm: Journal of the Royal Statistical Society. Series C (Applied Statistics), **28**, 100–108, doi:10.2307/2346830.

Lopes, A. P., L. M. de A. L. Ribeiro, E. D. Salvador, M. Pavan, and A. D. R. da Silva, 2017, Áreas de relevante interesse mineral Vale do Ribeira: Mineralizações polimetálicas (Pb, Ag, Zn, Cu e Au - "Tipo Panelas") em zonas de cisalhamento rúptil, cinturão Ribeira Meridional, SP-PR: CPRM. http://rigeo.cprm.gov.br/jspui/handle/doc/18958

Rosa, J. W. C., J. W. C. Rosa, and R. A. Fuck, 2014, Geophysical structures and tectonic evolution of the southern Guyana shield, Brazil: Journal of South American Earth Sciences, **52**, 57–71, doi:10.1016/J.JSAMES.2014.02.006.

Zhang, Y., S. Xiong, and T. Chen, 1998, Application of airborne gamma-ray spectrometry to geoscience in China: Applied Radiation and Isotopes, **49**, no. 1–2, 139–146, doi:10.1016/S0969-8043(97)00222-4.

Acknowledgments: Principal Component Analysis and K-means clustering were generated using R. Maps generated with ArcGIS 10.5. Thanks to Luiz Pinto, Rafael Ribeiro and Jairo de Andrade for helping with the data. Rafael acknowledges CNPq (grant 203589/2014-9) for graduate sponsorship and CPRM for granting absence of leave. Thanks to CPRM-Brazil for providing different airborne geophysical data, as well as geological information on different parts of Brazil. Thanks STRM for providing elevation models.