Copper and Gold Mineral Potential Modeling through Machine Learning in the Northern Copper Belt, Carajás Mineral Province - PA

Junny Kyley Mastop de Oliveira[1,2]; Adalene M. Silva[2]; Felipe M. Tavares[1]; Iago S. L. Costa[1]; Catarina L.B. Toledo[2]

[1] SGB – Serviço Geológico do Brasil

[2] UnB – Universidade de Brasília

The Northern Copper Belt (NCB), located in the world class Carajás Mineral Province, hosts important Cu and Au deposits, the largest being the Salobo deposit (1,133.4 Mt@0.62% Cu and 0.35 g/t Au) as well as others of smaller tonnages (e.g. Furnas). Nonetheless, the region still holds high potential for new discoveries. There are at least two important mineral systems for Cu and/or Au from different metallogenetic epochs : a) in the Neoarchean (2.72 to 2.55 Ga), a large-scale hydrothermal system, chronologically and spatially related to magmatism of the Old Salobo type, generated IOCG-type deposits (MSM1), for example the Salobo mine; b) in the Paleoproterozoic, contemporary to A-type magmatism of 1.88 Ga, when different polymetallic deposits (MSM2), related to granites, were emplaced, like the Gameleira deposit (Cu-Au-Fe-Mo-Co-REE's). In recent years, advances in exploration targeting techniques and artificial intelligence allowed several improvements regarding the development of mineral potential maps (MPMs) and Mineral potential modeling. Yet, only a few studies have been carried out on district scale in the NCB, concerning resource assessment and mineral potential modelling. In this context, we present this work, with the objective of mapping the mineral potential of both mineral systems for Cu and Au, using machine learning techniques. District-scale MPMs for each system (MSM1 and MSM2) and a final integrated model (FIM) were generated using spatial data modelling techniques. The development was carried out in the ArcGisPro software, while the model validation was made in the Orange Data Mining software. We used a data-driven method based on the Random Forest algorithm (RF). The adopted algorithm is useful for data classification and regression, combining several decision trees and making repeated predictions of the same phenomenon represented by the training data set. RF has been used for accurate predictions in various fields. The data used for training consisted of 25 deposit and 25 non-deposit points (30% applied for testing), with the non-deposit points being randomly drawn, but always controlled/checked with known geological points, considering a minimum distance of 1km between the non-deposits. The datasets were rasterized and normalized geological and geophysical data. The structural framework and aeromagnetic geophysical data were the most important prospective vectors, while neoarchean rocks were less relevant to the model. According to cross-validation parameters in 10 folds, the model had an accuracy of 86%, which demonstrates good performance in predicting prospective zones. The MIF, with 6 potentiality classes, pointed to several favorable areas for copper and gold, some indicating continuity of known targets and others in isolated sectors with no known deposits. It is worth highlighting the southeast (eastward extension

of the poju-game-gfunda trend) and northwest (northeast of the GT46 deposit) sectors of the NCB with extensive areas of moderate to very high potential (>10 km).