

## **Building a Better Geochemical Picture of the Australian Landscape**

P. T. MAIN<sup>1\*</sup>, D. C. CHAMPION<sup>1</sup>, J. WILFORD<sup>1</sup>, E. N. BASTRAKOV<sup>1</sup>

<sup>1</sup>Geoscience Australia, Canberra, Australia (\*correspondence: philip.main@ga.gov.au)

A vast quantity of surface sediment geochemistry data has been collected for the Australian continent by the mining industry and government over the last 50 years. Commonly, these datasets have been used in isolation, limiting their potential for regional interpretation and uses beyond the original intent. In order to maximise value and utility of this information across multiple scales, the data need to be levelled into one seamless dataset. Although variation between datasets may be due to difficult to correct differences (e.g. sample media, and targeted landscape features), many of the surveys are standardised such that variation is largely due to analytical factors (different laboratories, methods, and instruments). Accordingly, multiple surveys can be successfully merged by levelling to minimise analytical variation. To achieve this aim we have used statistical procedures based on analytical results for certified reference materials, internal project standards, and subsets of re-analysed survey samples. This was applied to nine geochemical surveys, totalling 15,605 samples, to produce a seamless levelled dataset for northern Australian (~2,100,000 km<sup>2</sup>).

Levelled samples were used to train a machine learning algorithm (RandomForests) to generate a suite of predictive geochemical maps. Machine learning provides a distinct advantage over traditional interpolation techniques, e.g. kriging, in that a large number of regional and national covariate datasets (e.g., aeromagnetic, radiometrics and Landsat data) can be used to elucidate relationships with the target data. This results in high resolution (up to the resolution of co-variate datasets, ~90 m) and physically meaningful predictions relating geochemistry to geology and geographical features. Machine learning modelling predictions can be used for identification of true data outliers (geochemical anomalies) that might be expressions of mineral systems. By identifying areas of model under-fitting and determining that they are not related to geological, geographical, and geomorphological factors, maps of anomalous points can be generated for follow-up analytical and statistical examination. Such maps are ambivalent to environmental background variations as elevated element concentrations in naturally “low” areas are not masked by higher background values elsewhere. In this presentation, we discuss the specific results obtained for total concentrations of Cu in transported sediments of the North Australian Craton.