EXPLORE
Newsletter for the Association of Applied Geochemists
www.appliedgeochemists.org

NUMBER 198 MARCH 2023

Old soils, new targets: Reassessing historic soil surveys with Ultra-Fine+® in a machine-learned landscape context

Anicia Henne[1], Ryan Noble[1], Dave Cole[2], Morgan Williams[1], Fang Huang[1] and Alfredo López Alfageme[3]
[1]CSIRO Mineral Resources, ARRC Building, 26 Dick Perry Avenue, Kensington, WA 6151, Australia
[2]CSIRO's Data61, Research Way, Clayton, VIC 3168, Australia
[3]CSIRO Chile International Centre of Excellence, Avda Apoquindo 4700 piso 9, Las Condes, Santiago, Chile

https://doi.org/10.70499/YCQK8307

INTRODUCTION

Greenfields mineral exploration in covered terrain is often hindered by a failure to detect, understand, and evaluate near-surface geochemical anomalies. This is due to several challenges including in the detection of relevant trace metals in cover (e.g., due to dilution or differential concentration), the availability of sufficient information (e.g., additional path-finder elements and soil properties), and an understanding of landscape context to enable exploration based on informed decision-making processes.

Between 2012 and 2018, the UltraFine+® method was designed specifically to address the challenges of detecting a suite of 52 elements in shallow transported cover (Noble *et al*. 2020a) by Australia's National Science Agency, the Commonwealth Scientific and Industrial Research Organisation (CSIRO). Over the past three years the CSIRO, with the support of many industry sponsors and Australian State and Territory geological surveys, developed a machine learning workflow to create landscape context for these results. Here we highlight some of the main developments of this approach to assist future research and development in soil geochemical analysis and interpretation.

Surface geochemical signatures of mobile metals relating to mineralisation below (or within) cover are often diluted and close to, or below, analytical detection limits (Anand *et al.* 2016) and improvement in soil sample analysis has commonly focussed on lowering detection limits by developing more sensitive extraction methods or partial extractions on bulk samples (e.g., Chao 1984; Bajc 1998; Gray *et al.* 1999). The UltraFine+® method focusses on extracting the clay-sized fraction (<2 µm; Noble *et al.* 2020a) from a standard soil sample which hosts the bulk of useful indicator elements, as these are preferentially adsorbed on clay-sized particles and other "scavenging" phases with large surface areas (e.g., organic compounds and various oxides/oxyhydroxides; Hall 1998). By removing the majority of the coarse-grained "barren" (usually silica-dominated) portion of a soil sample, the geochemical signal of mobile trace metals is increased, effectively enhancing measured abundances by up to 100 – 250 % of elements such as Au, Cu and Zn (Noble *et al.* 2018).

Regardless of the advanced method of extraction, significant value can be added by improving the interpretation of surface geochemical surveys in landscape context. In mineral exploration, the composition of the sampled regolith material, its position within the landscape, its genetic relationship to the bedrock hosting potential mineralisation (in situ vs. transported cover) and the depth of this cover (Anand *et al.* 2016) can affect the way metals move through the environment and impact how we interpret the data. However, these aspects are not always appropriately considered during surface exploration. Recently, the CSIRO has developed a machine learning workflow referred to as "Next Gen Analytics", to delineate landscape context for surface geochemical survey results. This approach uses exclusively publicly-available spatial features derived from remotely sensed data with the goal to normalise geochemical concentrations by landscape type. This enables the comparison of samples across vast regions with varying landscapes in greenfields exploration settings.

Here we examine historically collected samples from a greenfields soil survey over 3600 km² in Western Australia collected in 1999, and illustrate some key advances that UltraFine+® delivered in 2018, and the Next Gen Analytics approach delivered in 2023 (Fig. 1). The comparison shows the benefits of improved sensitivity of gold and other elements for analyses (Fig. 1B) with the addition of multielement interpretation and the ability to highlight geochemical anomalies within different landscape types (Fig. 1C). Importantly, Next Gen Analytics identifies anomalies in transported cover types where the measured elemental abundances are commonly lower compared to the whole survey, but anomalous relative to soils collected in similar landscape settings (Fig. 1C). While landscape/regolith/landform maps are available for many regions in Australia, these are often produced at coarse resolutions, are derived from aerial photography and, in some cases, supplemented by limited on-the-ground observations, and do not take depth of cover into consideration. Above all, they are prone to human inconsistencies in interpretation. Using machine learning to derive landscapes from remotely sensed spatial features allows this approach to be employed in most regions, even where traditional map products are not available or only available at a coarse resolution (as is often the case when explorers advance into true greenfields settings).

continued on page 5

Old Soils, New Targets: ... *continued from page 1*

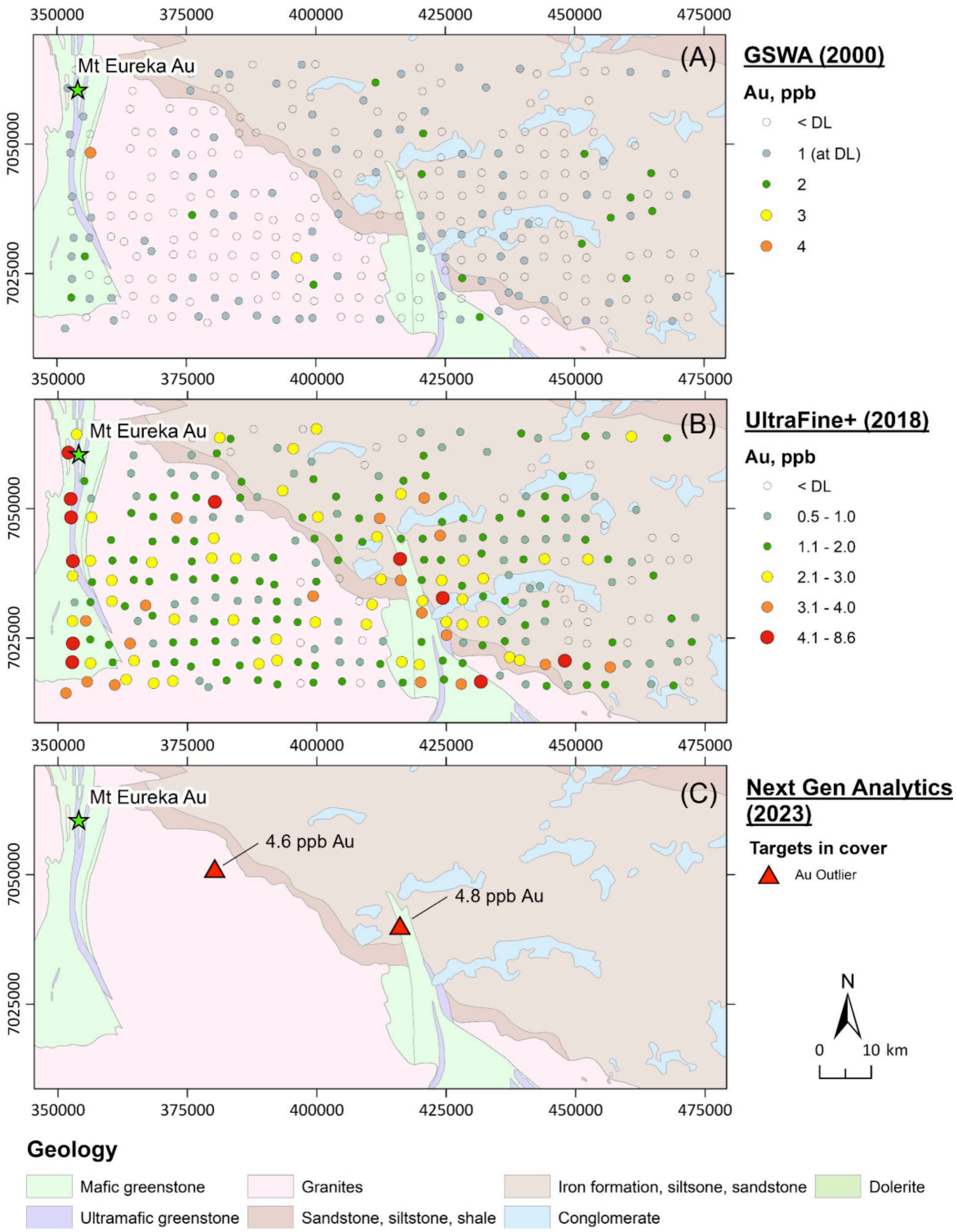


Fig. 1: Key developments in soil sample analysis and interpretation over the past two decades on the example of Au in the Kingston project area in Western Australia. The same soil samples were analysed in 2000 and in 2018, and later interpreted in machine learning derived landscape context in 2023. Mt Eureka, the only mined Au occurrence within the survey area, is indicated with a green star. (A) Gold results analysed by the Geological Survey of Western Australia in 2000 (Pye et al., 2000) via a traditional soil sample analytical fire assay fusion method from <2 mm to > 0.45 mm sieved soil that was milled to a nominal <75 µm size. Only 111 samples out of 302 returned Au above the detection limit with the traditional soil analysis method. Of these, 93 samples were at the detection limit (1 ppb) and only 18 samples showed appreciable Au concentrations. (B) Gold analytical results of the same 302 archived soil samples via the UltraFine+® analytical method in 2018 (Noble et al. 2020b). The same samples returned detectable Au in most samples with the Ultra-Fine+® method – 38 samples were below the detection limit (0.5 ppb), and 2 samples were at the detection limit. (C) Gold outliers identified in the 2018 UltraFine+® dataset, where outliers are derived for each proxy landscape type with the Next Gen Analytics machine learning workflow (see details in text below). Two new potential targets have been identified in transported cover which show relatively low concentrations (4.6 and 4.8 ppb Au) compared to the highest values (up to 8.6 ppb).

continued on page 6

Old Soils, New Targets: … *continued from page 5*

Study Site – Kingston

Our case study covers the southern half of the Kingston 1:250,000 topographic map sheet in the Northern Goldfields of Western Australia. The area straddles the north-eastern edge of the highly prospective Yilgarn Craton (dominated by granites and greenstones) and extends into the Earaheedy Basin (dominated by shales interbedded with iron formation of the Frere Formation; Fig. 2). Surface geochemical exploration within this area has largely focused on orogenic Au and komatiite-hosted Ni mineralisation, both hosted within the greenstone belts (https://minedex.dmirs.wa.gov.au/web/home) and, to date, only the Mount Eureka Au mine has been developed within the area (Fig. 2). The small deposit produced 941 t of ore from 1932 – 1937, with quartz-vein associated mineralisation hosted in silicified talc-carbonate schist within mafic greenstone (Pye *et al*. 2000). In recent years, industry has shown interest in VMS-style Cu-Zn deposits within the region (https://www.roxresources.com.au/projects/mt-fisher-gold-project/).

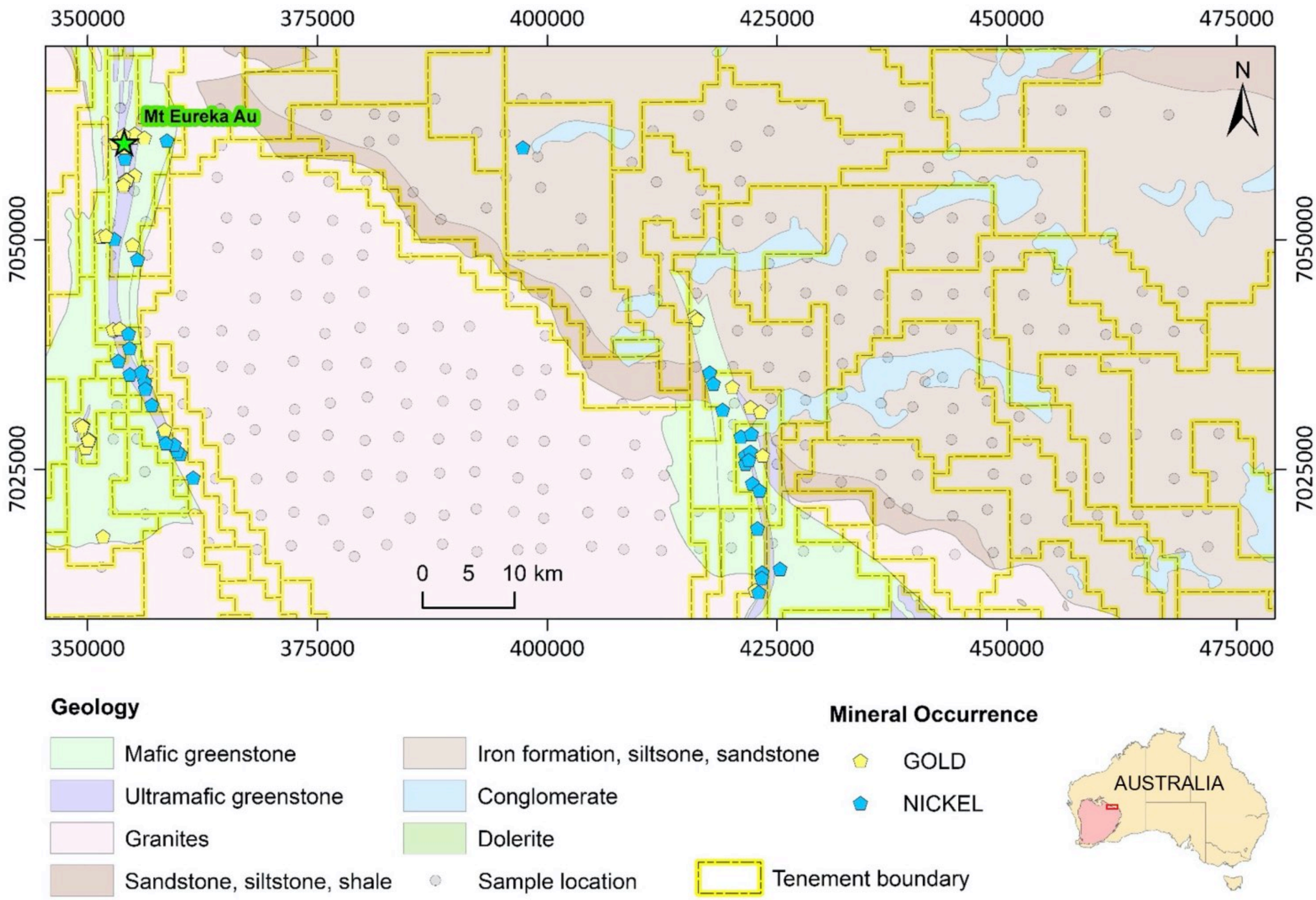


Fig. 2: Location of case study site, samples collected by the GSWA in 1999 and reanalysed via the UltraFine+® method in 2017, and simplified geology of part of the Kingston map sheet (after Martin et al. 2014). Tenement distribution indicates current (as of 20 February 2023) exploration activities within the area are focused on greenstone and sedimentary units. Recorded mineral occurrences are concentrated within the ultramafic and mafic greenstone units.

The bedrock in the area is largely concealed by transported cover over almost 85 % of the area with few exposed or residual weathered regolith materials (Pye *et al*. 2000; Fig. 3A). The transported cover is dominated by broad sheetwash plains and valleys incised by channels, sandplain and colluvial materials with minor lacustrine (playas), alluvial and eolian deposits (Fig. 3B). The cover materials have been characterised in detail in a study by the Geological Survey of Western Australia (GSWA) in 1999 noting remnants of major drainage systems that are either calcrete-dominated or characterised by dunes and alluvial channels associated with playa-lake systems (Pye *et al*. 2000). In the same study, the GSWA also analysed almost 1000 samples across the Kingston map sheet which were archived at the GSWA's core storage facility and subsequently made available for reanalysis via the UltraFine+® analytical method (published in Noble *et al*. 2020a). The availability of these analytical results from a variety of different regolith materials, in conjunction with the detailed regolith mapping carried out by the GSWA, made the Kingston map sheet a useful location to trial and improve the UltraFine+® method as well as the Next Gen Analytics approach.

continued on page 8

Old Soils, New Targets: … *continued from page 6*

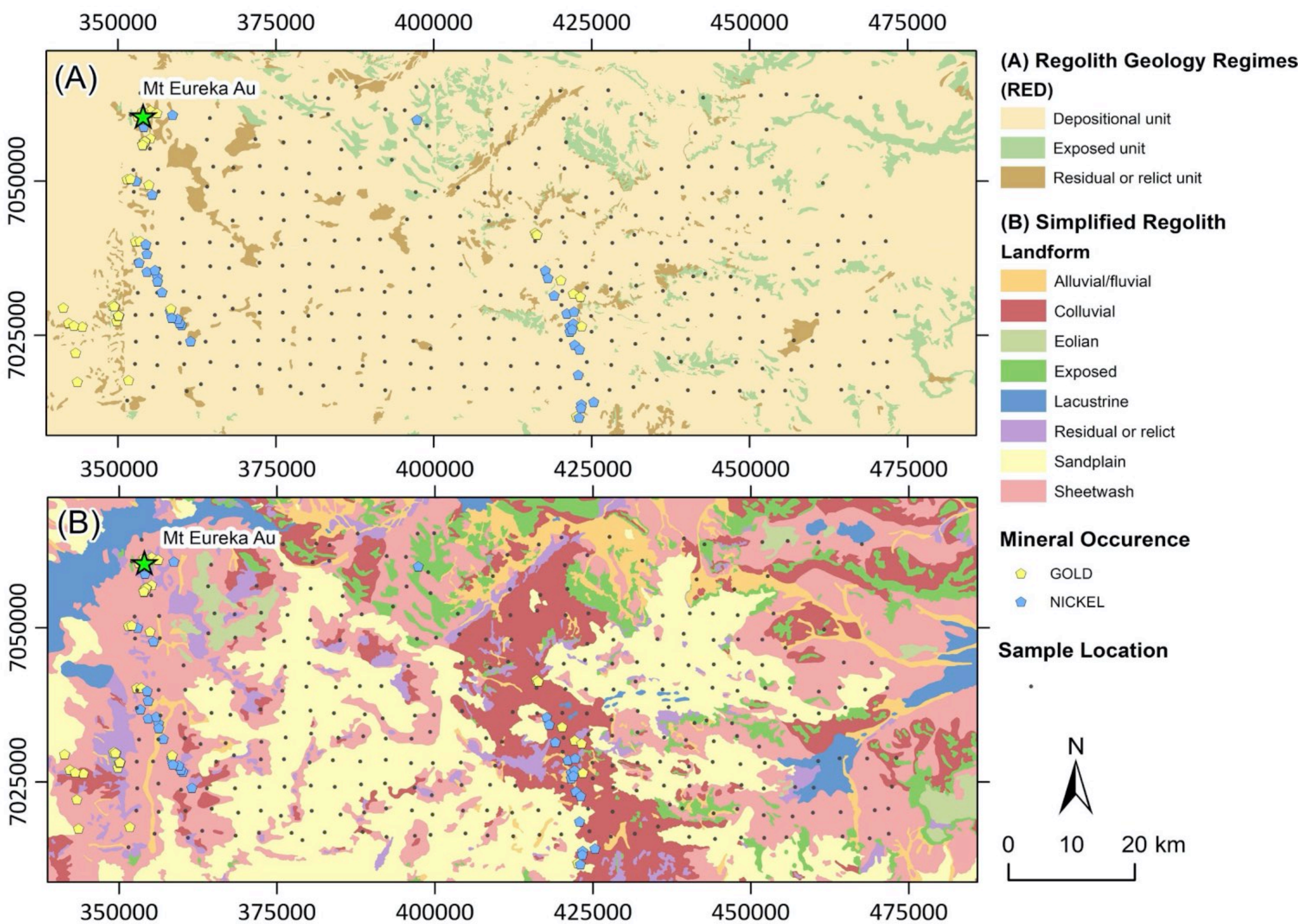


Fig. 3: Regolith maps over the Kingston area. (A) Regolith geology regimes map (de Souza Kovacs and Jakica 2021) based on the RED scheme of Anand et al. (1993) where regolith landform units are classified into three major regimes (residual, erosional and depositional) relating to their composition and landscape position. (B) Simplified regolith landform map (Jakica et al. 2020) which expands the RED regimes into more detail regarding dominant mechanisms of formation or parent material.

2000 – Historic soil sampling survey and regolith mapping

The GSWA collected almost 1000 regolith samples from stream sediments, sheetwash/soil, lake sediments and sandplain materials in the Kingston map sheet as part of an extensive regional study in 1999. These samples were sieved to < 2 mm to >0.45 mm, pulverised and analysed using seven different analytical methods to derive 49 geochemical and physicochemical parameters including major and trace elements, pH, and conductivity. The methods and analytical results are documented in Pye *et al*. (2000). The GSWA also provided a regolith map, laboriously generated by combining information from topographic data, black-and-white aerial photographs (dated 1974 and 1993), Landsat Thematic Mapper images (dated 1994), previous geological maps (Bunting 1980, 1986; Myers and Hocking 1998) as well as geochemical regolith survey results and field observations (one per 16 km²) collected by six geologists and six field assistants using two helicopters. Airborne radiometric and magnetic data were also consulted (Pye *et al*. 2000). The resulting map is available in McGuiness and Pye (2000).

The work of the GSWA characterised regolith materials as well as indicated the abundance of metals of interest on the Kingston map sheet (Pye *et al*. 2000). Given the analysed size fraction was >0.45 mm, it was expected that the results would show lower concentrations than analyses via UltraFine+®. Indeed, for Au, 94% of the data was at or below the detection limit (1 ppb). Only 18 of the 302 samples returned Au above the detection limit with the traditional soil analysis method. Where detected, concentrations were generally low (Au ≤4 ppb) with the highest values recorded in cover over mafic and ultramafic (greenstone) rocks (Fig. 1A). Some higher values of Bi and As were observed near the most westerly greenstone belt that hosts the Mt Eureka mine, while most Ag concentrations were below the detection limit (0.1 ppm). Higher values of Ni, Pb and Zn were also observed in proximity to greenstone belts within the area, while higher Cu values were only observed near or over the most westerly greenstone belt that hosts the Mt Eureka mine.

continued on page 10

Old Soils, New Targets: ... *continued from page 8*

The work by Pye *et al*. (2000) indicates that regolith chemistry in the Kingston map sheet is at least partly controlled by underlying bedrock geochemistry, differentiating regolith that overlays greenstone and granitic bedrock. However, Pye *et al*. (2000) also note the control of regolith type on geochemistry in soil samples, such as a general depletion of many analytes in lake sediments with low clay content (inferred from major element analysis), as well as the prevalence of SiO_2 in sandplain materials and the potential effect of dilution of other analytes by eolian-derived quartz sands that may contribute material to otherwise locally-derived sandplain materials. They indicate that both bedrock and regolith type exert controls on bulk geochemistry, and it is therefore important to assess geochemical analyses within this context.

While the regolith materials map derived from this work (MacGuiness and Pye 2000) is detailed, compiling the map required multiple soil analyses, extended on-the-ground labour and an in-depth understanding of regolith processes. Pye *et al*. (2000), also note that discrepancies arose when regolith codes were assigned from field observations compared to those assigned from maps compiled from remotely sensed data.

2018 – Old soils re-assayed with UltraFine+®

In 2017, 302 archived sheetwash and sandplain samples, initially collected by the GSWA, were reanalysed via the UltraFine+® soil analytical method (MAR-04) at LabWest Pty Ltd, Perth, Australia. This method was developed by the CSIRO in collaboration with a commercial laboratory and sponsored by nine industry and state government partners (Noble *et al*. 2018), and was designed to capture the mobile element concentration in cover. The complete UltraFine+® method is based on separating the ultrafine (<2 μm) size fraction via suspension in de-ionised water and a dispersant followed by centrifugation, and uses a microwave-assisted aqua regia digestion in a closed Teflon tube to derive a multi-element suite using ICP-MS and ICP-OES. The full method is described in detail in Noble *et al*. (2020a). However, detection limits have since been improved, and additional elements I, Br and Pd have been added to the standard analysis suite since its publication (https://www.labwest.net/ultrafine-dl/). Additional rare earth elements (REE) are also available for some commercial analytical packages.

The effect of various small size fractions on elemental analysis results in mineral exploration has been investigated, among others, by Scott and van Riel (1999), Morris (2013), Anand *et al*. (2014), Arne and MacFarlane (2014), Baker (2015) and Sader *et al*. (2018), and included test work on Au concentrations in different size fractions of sand dune samples, which indicated that most of the Au was contained in the <2 μm size fraction (Noble *et al*. 2013). Based on this work, the UltraFine+® soil analytical method was developed to separate the <2 μm "ultrafine" soil fraction for multielement analysis, effectively concentrating the phases of interest followed by an aggressive digestion (Noble *et al*. 2020a, 2020b). The improved recovery of mobile elements from the ultrafine fraction results in higher absolute measured concentrations on average, effectively reducing the number of results below the detection limit. This improves the resolution of concentrations near the detection limit, which enables the delineation of subtle geochemical enrichments for these elements (e.g., Au; Fig. 1A, B). The detection of these subtle variations is particularly relevant for exploration through transported cover. As an example, the exploration relevant elements Au, As, Cu, Ni, Pb and Zn in the Kingston survey show a higher median abundance in the UltraFine+® results compared to the previous survey results via mixed-acid digestion and fire assay (Fig. 4).

It is important to note that the ultrafine method does not increase concentrations or lower detection limits below other current analytical methods, but rather removes the diluting effect of trace element-poor phases in the bulk sample. The different approach to sampling and analysing is designed to capture more of the mobile phases derived via dispersal mechanisms in regolith (shallow, <30 m, the mobile phase in transported as well as weathered *in situ*) materials, while more immobile or resistate elements are likely better recovered by standard digestions such as four-acid (Henne *et al*. 2022). It is therefore not surprising that the greatest overall measured abundances are not always recovered with the UltraFine+® method (e.g., in the case of gold nuggets or, in this study, As, Ni and Pb; Fig. 4B, D, E). However, the UltraFine+® results show smaller interquartile ranges which increases confidence that the method consistently measures phases relating to the same geological processes (the clay fraction). This indicates that the resulting outlier definition is more dependable in a regolith material context than the more variable analyses via standard soil analytical methods (compare red and green box plots in Fig. 4).

continued on page 11

Old Soils, New Targets: ... *continued from page 10*

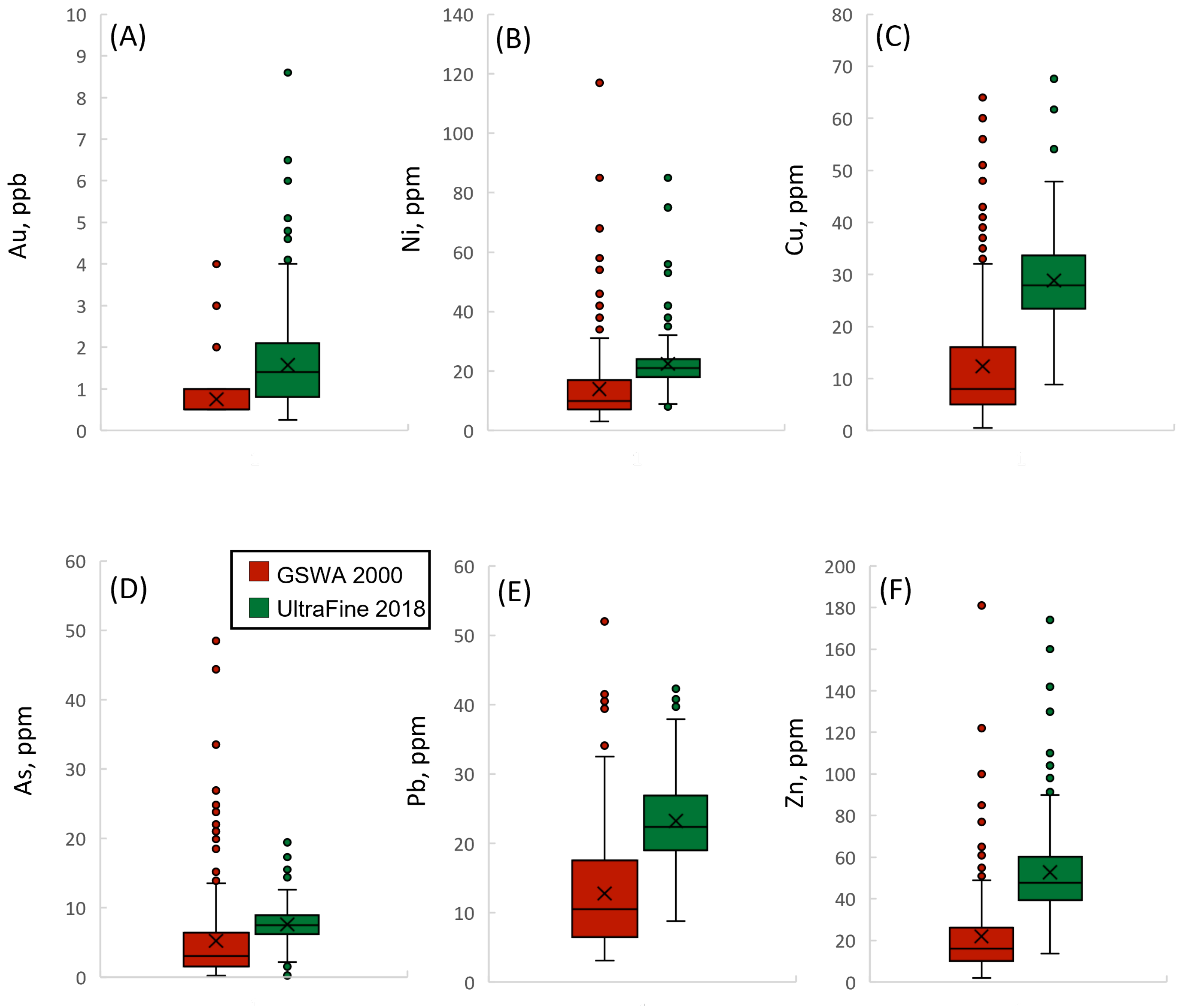


Fig. 4: Comparison of exploration relevant pathfinder element analyses between the historic soil analyses and the UltraFine+® analyses. Boxplots include values below the detection limit (replaced by half the respective detection limit); excluding the median; n=302. The 2000 survey results via mixed acid-digestion (except for Au which was analysed via fire assay) are displayed in red and the 2018 results from re-analysis via UltraFine+® are displayed in green. (A) Au in ppb, detection limits GSWA 2000 = 1 pbb, UltraFine+® 2018 = 0.5 ppb. (B) Ni in ppm, all values above the detection limits. (C) Cu ppm, detection limits GSWA 2000 = 1 ppm (all values above the detection limit for UltraFine+® 2018). (D) As in ppm, detection limits GSWA 2000 = 0.4 ppm (all values above the detection limit for UltraFine+® 2018). (E) Pb in ppm, all values above the detection limits. (F) Zn in ppm, all values above the detection limits.

2023 Next Gen Analytics – Adding Machine Learning to delineate landscape context for surface geochemistry

Over the period of 2020 to 2023, the CSIRO developed an unsupervised machine learning workflow to generate proxies for landscape types from spatial feature layers to provide context for surface geochemical data interpretation. The complete Next Gen Analytics workflow includes a variety of outputs such as principal component analysis, exploration indices, dispersion and source direction and outputs for non-geochemical soil property data. Here, we only present parts of this workflow (Fig. 5) and highlight the outputs of one landscape model and the resulting geochemical outliers in landscape context. An overview of the complete Next Gen Analytics data package and several How-to guides can be accessed via https://research.csiro.au/ultrafine/.

Methods

The spatial feature layers that were used in the landscape model presented here for the Kingston case study included

continued on page 12

Old Soils, New Targets: ... *continued from page 11*

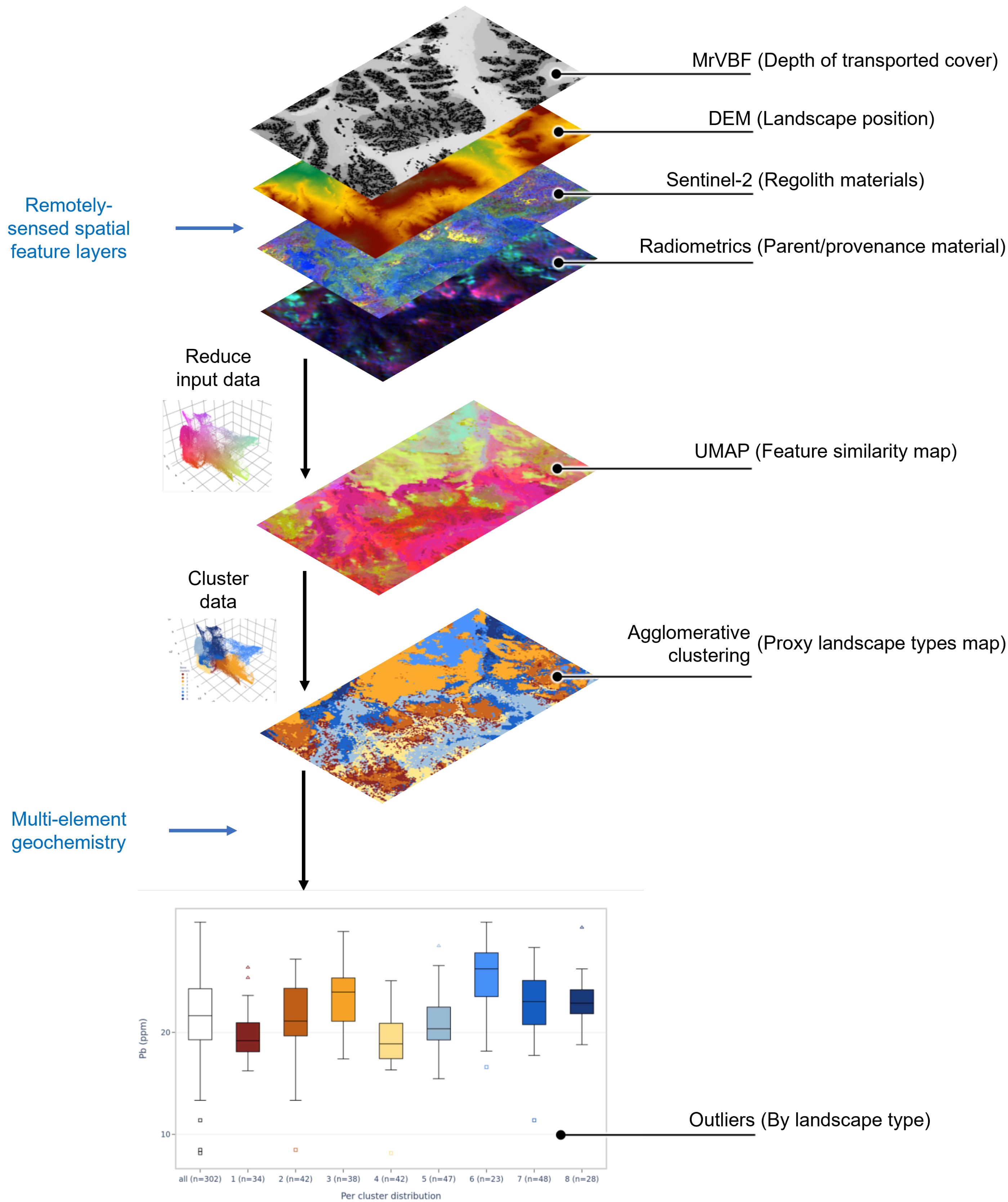


Fig. 5: Simplified Next Gen Analytics machine learning workflow with key components applied to the Kingston case study to demonstrate the main outputs (proxy landscape maps and outliers by landscape type for each analysed element). The boxplots resulting from this workflow break down the soil sample dataset into groups according to a machine-learned landscape cluster. See text for more details.

continued on page 13

Old Soils, New Targets: ... *continued from page 12*

a Digital Elevation Model (Copernicus GLO-30), Multi-resolution Valley Bottom Flatness (MrVBF; a proxy of depth of cover developed by Gallant *et al*. 2012), Radiometric K (%), Th (ppm) and U (ppm) (Poudjom Djomani and Minty 2019a-c) and regolith band ratios derived, after a method by Gozzard (2005), from Sentinel-2 multispectral imagery (Geoscience Australia's barest earth product; Wilford and Roberts 2021). The spatial feature layers were resampled to a common 100 m² grid and the roads within the area were masked (and appear as white lines in outputs). The dimensionality reduction algorithm Uniform Manifold Projection and Approximation (UMAP; McInnes *et al*. 2018) was used to project the pixel (grid cell) data into a three-dimensional latent space. After which, a multi-step clustering process was used to assign a proxy landscape type to each data point (i.e., pixel). First, a random subset of 20,000 pixels was used to fit an agglomerative clustering model (scikit-learn implementation in Python; Pedregosa *et al*. 2011). Next, these same samples and their cluster IDs were used to train a random forest classifier. Finally, this classifier was used to predict cluster IDs for all pixels across the model area. Pixels were classified into eight clusters based on similar spatial feature properties, to provide comparable complexity to the eight units on the existing state-wide Simplified Regolith Landform map (Fig. 3B). Pixels in each cluster were assigned a colour based on the ordinal rank of the mean MrVBF for each cluster.

The resulting proxy landscape map was used to group geochemical data according to the landscape cluster corresponding to each soil sample location. Elemental outliers were then calculated for each of these clusters, after a log-transformation. Results below detection limits were first replaced by half the detection limit. Here, outliers are defined as values that are greater than 1.5•IQR beyond the first and third quartiles (where IQR is the interquartile range, Q3 – Q1).

Landscape model

Understanding the relationships between material provenance, weathering and dispersion processes as well as general landform characteristics and material properties is a complex undertaking and can hinder successful greenfields exploration in areas where comprehensive datasets, field knowledge, general expertise, or time are limited. There are three different approaches to terrain classification (summarised in Gozzard (2005)) – landscape criteria (based on physical attributes; e.g., association of geology, soils etc.), genetic (based on underlying causal factors; e.g., RED scheme – residual, erosional, depositional), and parametric (quantitative, classification based on selected attributes; e.g., based on spatial feature layers). The latter is often lauded as being an approach unbiased by humans. However, while the Next Gen Analytics models are intended to have little direct human input (e.g., we do not include available geological or regolith maps), there is an inherent human bias in the parametric approach, due to assumptions relating to appropriate input data. Our input spatial feature layers were chosen based on an understanding of genetic and landscape criteria models and how these may affect soil sample characteristics in relation to metal mobility in the near surface (<30 cm depth). The main considerations for input layers were (a) indication of parent/source material (radiometric data), (b) general landscape position (DEM), (c) the depth of transported cover (MrVBF), and (d) regolith material type (Sentinel-2 derived regolith band ratios), mainly relating to clay and iron oxide content.

There are many other spatial feature layers (both remotely sensed and human-interpreted) which may be useful depending on site and commodity specific exploration needs. However, the Next Gen Analytics workflow was designed for first-pass, greenfields exploration with little knowledge of a given area. In addition, we exclusively use publicly available data, to enable application anywhere on the Australian continent with a resolution of 30 m. The workflow identifies clusters of pixels with similar feature properties without explicit consideration for the geographical location, geochemistry, or soil properties at each point, and without further human influence. It is important to note that the model output shown here (Fig. 6A) has not been adjusted to fit the ideal number of different landscape clusters, but was limited to eight clusters to demonstrate the concept in comparison to the available simplified regolith map (Fig. 6B).

The machine learning-derived landscape map is not amenable to description using concise nomenclature that would fit neatly into available regolith classification schemes. Hence, we simply refer to them here as numbered landscape clusters and their designated colour. This may present an initial challenge for human interpretation. However, it also presents an opportunity, as it allows for classification based on measurable properties of the regolith material rather than forcing a generalised landscape type into a rigid classification scheme. The clusters are not internally homogeneous (may consist of, e.g., sandplain in one area and grade into alluvial materials in other parts) but they do have recurrent characteristics that reflect the spatial feature input layers. The spatial feature layer with the greater relative variance has the most influence. This is in line with the model assumption that, where a deviation in, e.g., depth of cover, is large, this will influence the ability of a given element to migrate into surface materials. On the other hand, where a region is relatively uniform in depth of cover, but the material has vastly different regolith band ratio signals (related to clay and iron oxide content) this should be considered as to its effect on metal mobility. A brief overview of likely materials and main features observed for each cluster are noted in parentheses in the legend in Fig. 6A, and is derived in consultation with input feature layers and available surface geology and regolith maps as we did not complete on-the ground validation for this specific site.

Comparing the eight machine learned landscape clusters (Fig. 6A) to the eight regolith landforms on the publicly available map (Fig. 6B) indicates that the machine learning approach has defined clusters with similar spatial distributions to regolith landforms in the available map product. For example, landscape cluster 5 (light blue; Fig. 6A) has a similar distri-

continued on page 14

Old Soils, New Targets: ... *continued from page 13*

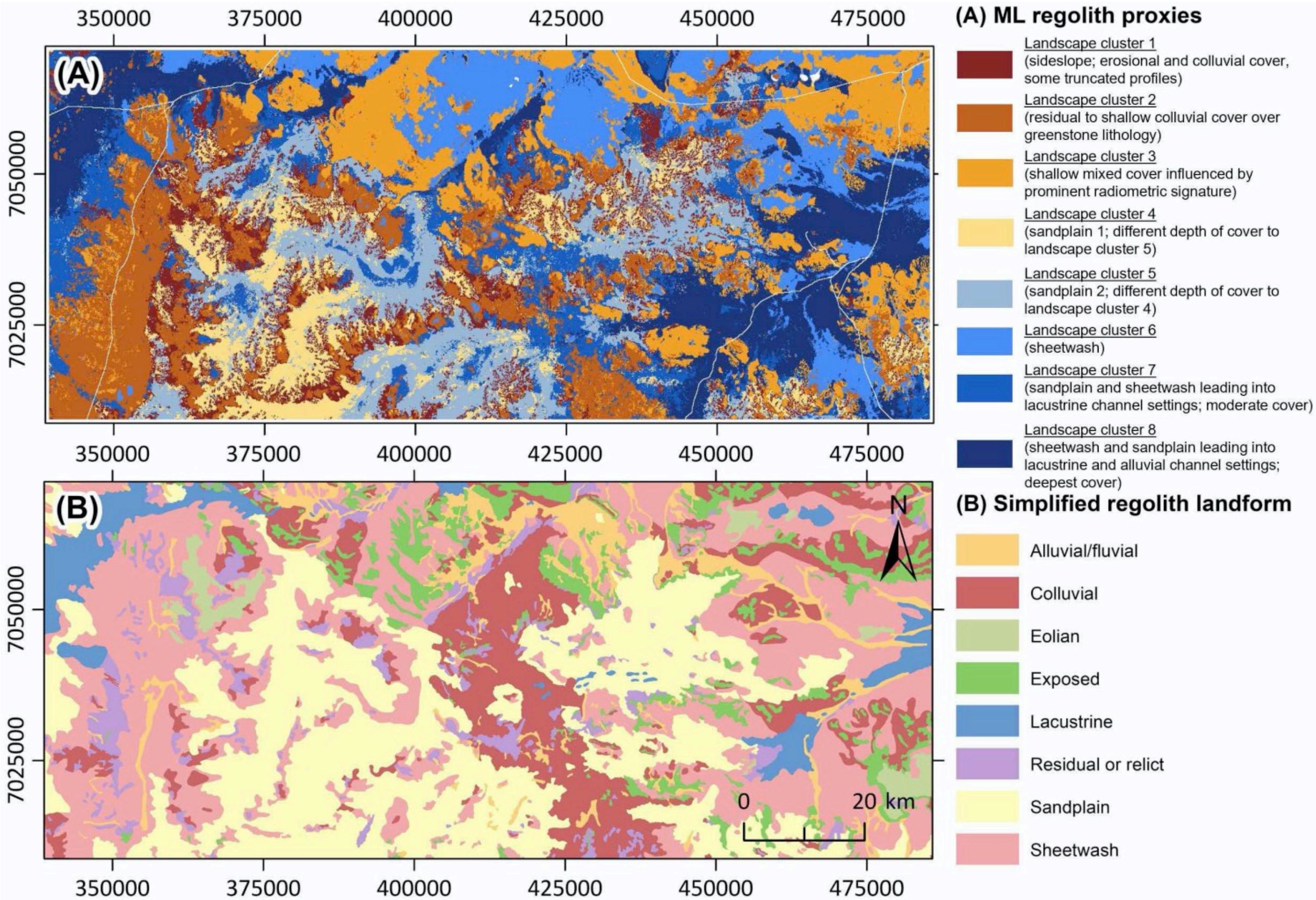


Fig. 6: Comparison of human characterisation and machine-learned characterisation of regolith over the Kingston area. (A) Simplified regolith landform map (de Souza Kovacs and Jakica 2021) derived via human interpretation. (B) Machine learned proxy-landform map based on remotely sensed spatial features. The clusters are not amenable to nomenclature that directly aligns with existing regolith classification schemes. We indicate in brackets some main features observed for each landscape cluster.

bution to what is mapped as sandplain in the available regolith landform map in the vicinity of the Yilgarn Craton granites (Fig. 6B). This is not surprising given lithological controls relating to greenstone and granitic bedrock on regolith materials are evident in surface geochemical results, which are reflected in the spatial feature layers (e.g., radiometric data) and therefore affect the machine learned landscapes. However, major differences can be observed. This is partly due to the 100 m resolution and the lack of smoothing in the machine learned approach, and partly due to the choice of input layers. For example, in the centre-left of the project area where one unit is mapped as sandplain material in the simplified regolith landform product (light yellow; Fig. 7B), the machine learning approach has assigned two clusters in this setting (light blue and light yellow in Fig. 7A). While both of these clusters likely do indicate sandplain material, the more traditional regolith landform map does not take into account the change in depth of cover (from the relatively deeper light grey to the relatively shallower grey of the MrVBF; Fig. 7D), nor the material composition related to clay and iron oxide content (relatively more blue vs. relatively more yellow colours in Fig. 7C). Topographic (DEM, Fig. 7E) and radiometric (Fig. 7F) information was consulted in the human-interpreted landform map and is well represented in both outputs. However, the change in regolith composition, in this case, does not relate to K, Th and U (radiometric data) but iron oxide and clay content (Sentinel-2 derived regolith band ratios).

In the case of the Kingston map sheet, a host of information on landform and regolith materials has been collected from remotely sensed and physical observations. However, there are many areas where no or very limited regolith landform, surface geology or regolith material maps exist, and the Next Gen Analytics approach is a cost-effective and rapid method to generate a first-pass landscape map for greenfields exploration. In addition, no physical observations are required, and the resolution is usually 30 m. Pye *et al.* (2000) also noted introduced human errors which may have skewed their geochemical statistics, such as samples that were thought to have been collected over greenstones but were instead likely to have been located over granites. The machine learning workflow aims to reduce the influence of human bias and error and provides a more consistent and objective landscape model.

continued on page 15

Old Soils, New Targets: ... *continued from page 14*

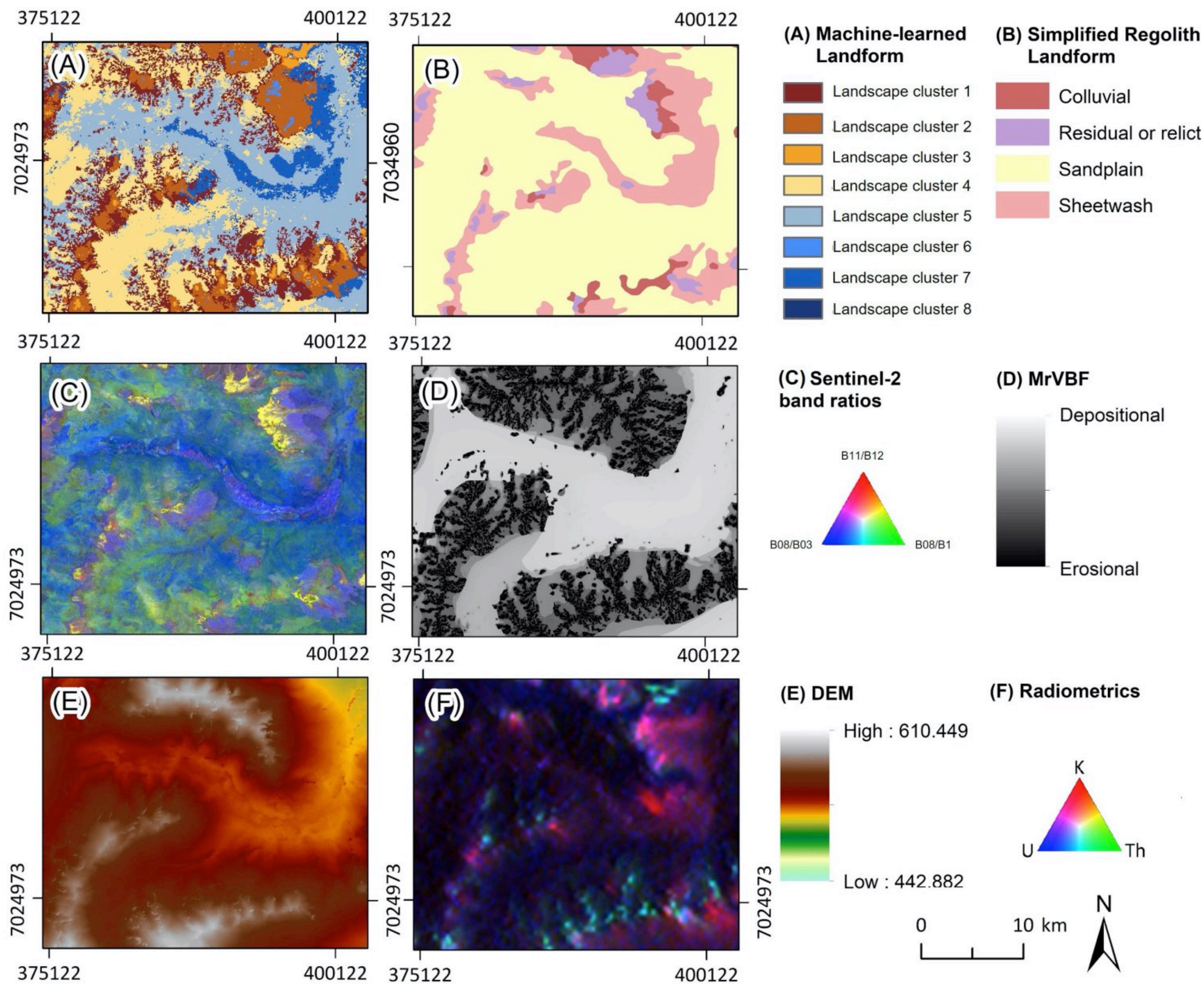


Fig. 7: Comparison of machine learned model to regolith landform map and the parameters used to generate the model. While we would consider a landscape model with more clusters as more appropriate for the Kingston area owing to its size (Fig. 5;) the usefulness of a model needs to be considered in terms of the exploration context and its application. This includes which features are likely to influence geochemical soil analytical results and the number of samples in each landscape type (for statistical outlier calculation, we consider the minimum number to be 50 samples per cluster). It is important to note when assessing geochemical data, that we group the regolith into classes of similar materials to be able to understand whether an elevated geochemical result in a surface soil is anomalous (and, therefore, potentially related to bedrock geology via dispersal mechanisms or supergene enrichment in the cover itself) or simply a variation in background material composition.

Outliers by landscape type

The goal of the Next Gen Analytics landscape modelling demonstrated above is to normalise geochemical concentrations by landscape, so samples can be compared across large areas with varying landscapes. For this purpose, statistical outliers are calculated for each analysed element based on their assigned landscape cluster (coloured boxes in Fig. 8B). By separating samples by landscape, we can assess them in separate sub-populations. For example, if we consider the distribution of Cu concentrations over the Kingston area (Fig. 8A) we might conclude that the largest values are observed near, or close to, the greenstone belts (similar to observations by Pye *et al*. 2000 on historic geochemical results). When identifying traditional outliers from the whole sample population (displayed as white triangles associated with the white box on the left-hand side in Fig. 8B) this interpretation is confirmed with one outlier located over each greenstone belt (Fig. 8C). However, this approach does not consider whether elevated metal concentrations are recorded in samples that were collected in residual or transported landscape settings. If mineralisation is present, the mobile element signature might

continued on page 16

Old Soils, New Targets: ... *continued from page 15*

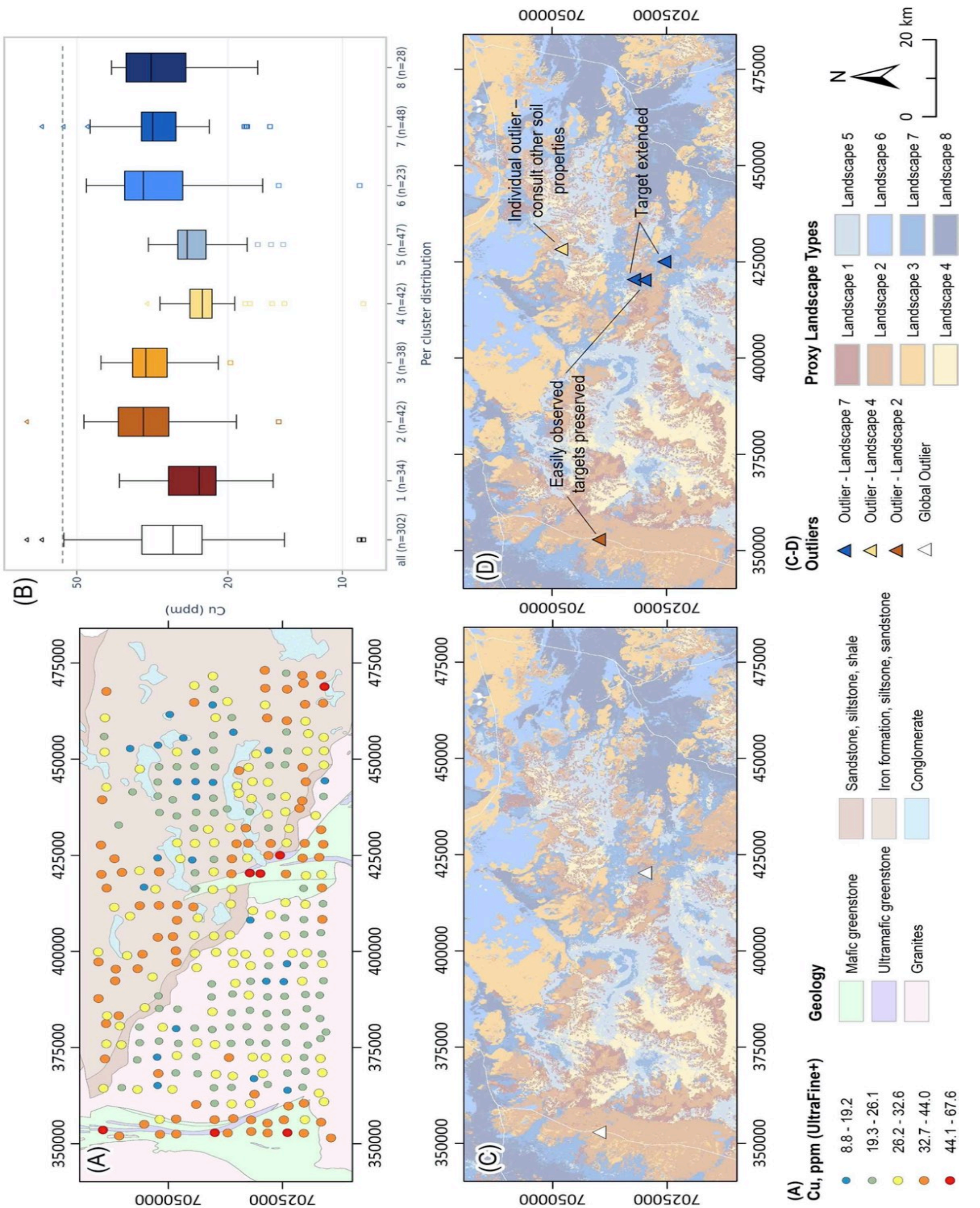


Fig. 8: Different ways of interpreting the same soil survey on the example of Cu. (A) Cu concentrations in ppm displayed based on natural breaks in the data. All data displayed was analysed via UltraFine+®. (B) Boxplots of the whole population (white box) broken down into sub-populations (coloured boxes) based on landscape clusters. (C) Location of whole population outliers. (D) Location of additional outliers coloured by sub-population. Note that the whole population outliers are still represented when viewed by sub-population.

continued on page 17

Old Soils, New Targets: ... *continued from page 16*

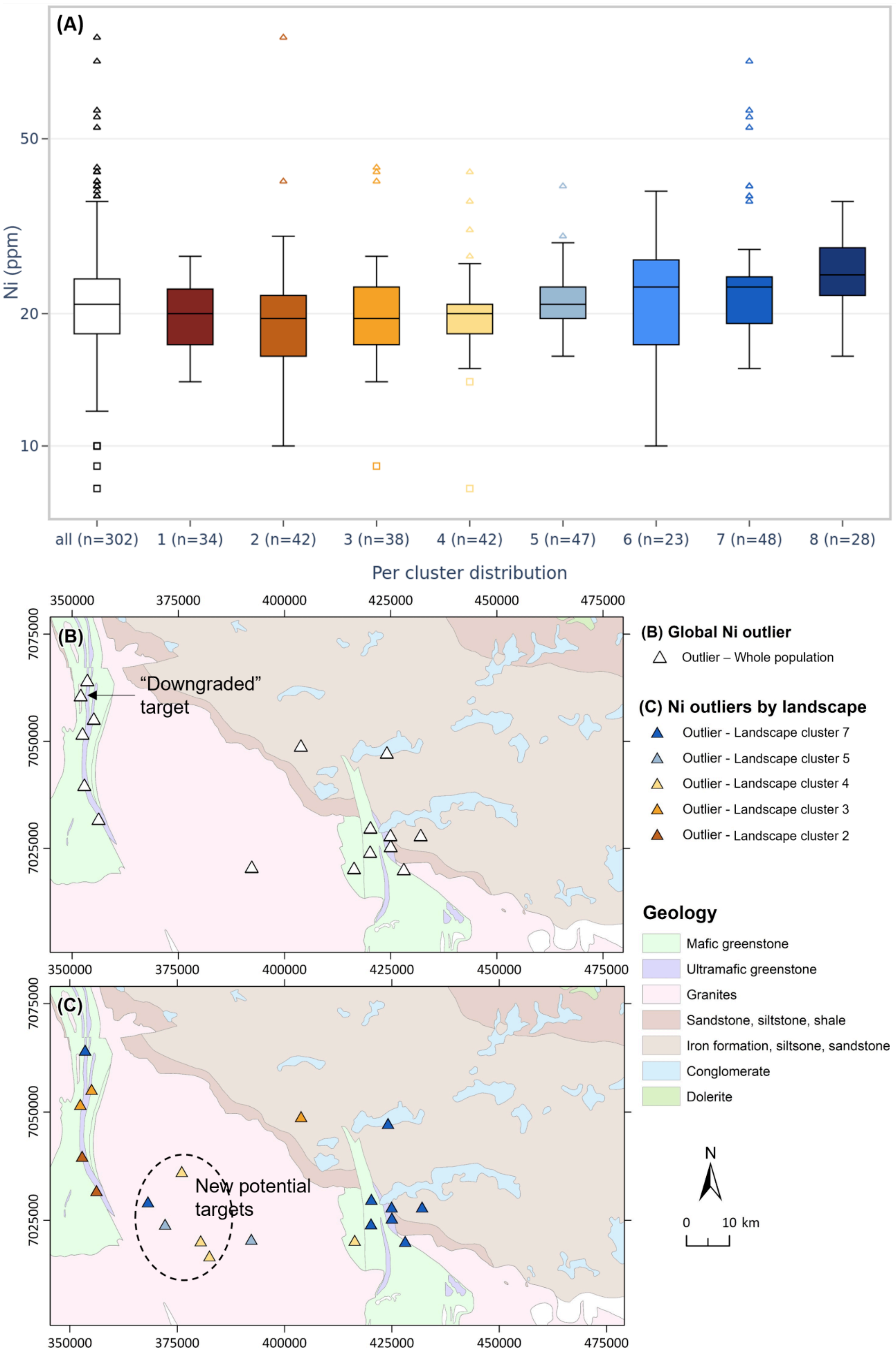


Fig. 9: Comparison of Ni outliers identified from the whole population compared to those identified in sub-populations. All data displayed was analysed via UltraFine+®. (A) Boxplots of the whole population (white box) broken down into sub-populations (coloured boxes) based on landscape clusters. (B) Location of whole population Ni outliers. (C) Location of additional Ni outliers coloured by sub-population.

continued on page 18

Old Soils, New Targets: … *continued from page 17*

have a much weaker geochemical signal in samples collected in transported (depositional) landscape settings compared to residual settings. Therefore, when the whole sample population is broken down by different landscape settings (coloured boxes in Fig. 8B), each sub-population can be assessed individually, highlighting potential anomalies within different landscape settings.

In general, the Cu outliers (triangles) below the dashed line (Fig. 8B) are often considered unremarkable (background concentration values) if evaluated as part of the whole population. However, when evaluated based on landscape context, there are two additional Cu outliers in the cluster 7 sub-population in depositional settings (sandplain and sheetwash, royal blue box in Fig. 8B and royal blue outliers in Fig. 8D). The single Cu outlier in the light brown population on the other hand, is located in residual or shallow cover and is, as expected, well represented by outliers in the overall sample population (compare white box/triangles with light brown triangle above the dashed line in Fig. 8B). Outliers in the royal blue sub-population extend one of the potential targets in the whole population and an additional potential target is identified in the yellow sub-population (Fig. 8D).

In an area where exploration has largely focussed on locations over known greenstone belts, which have generally elevated background elemental signatures (e.g., Ni, Pb and Zn; Pye *et al*. 2000) interpreting assays by landscape type can supply additional information to upgrade or downgrade traditional targets (Fig. 9). Similar to observations in the historic data, whole population outliers for Ni over the Kingston area are spatially associated with greenstone belts (white triangles in Fig. 9A and B). Most of these potential targets are confirmed when the data is assessed by sub-population providing more confidence as to their anomalous character. However, one outlier is "downgraded" (no longer an outlier in its sub-population; Fig. 9B) while other outliers are now evident in sub-populations sampled in sandplain materials (Fig. 9C).

This approach is intended as a first-pass geochemical data interpretation in greenfields settings with mixed cover. It is highly encouraged that, once an area for follow-up exploration is identified, all available landscape/landform information including field observations and human interpretation as well as geochemical and other soil property results (e.g., sizing, spectral mineralogy and pH) be reviewed. Dependent on availability and target commodity, the Next Gen Analytics approach can also be adjusted to include different spatial features. Where soil sample density permits, it is always recommended to adjust the number of clusters to fit the degree of landscape complexity.

Future developments

The UltraFine+® soil analytical method is commercially available and has to date been used by over 160 companies in Australia and has already expanded into a handful of international settings. Comparison of assay results to historic samples has shown improvements of effective detection limits as well as more repeatable and therefore more reliable results of a range of exploration relevant mobile phases in shallow cover. The Next Gen Analytics approach is currently an R&D project product, which is not yet commercially available. However, it has been tested on over 40 sites in Australia and early test work in New Zealand shows promising approaches for international settings.

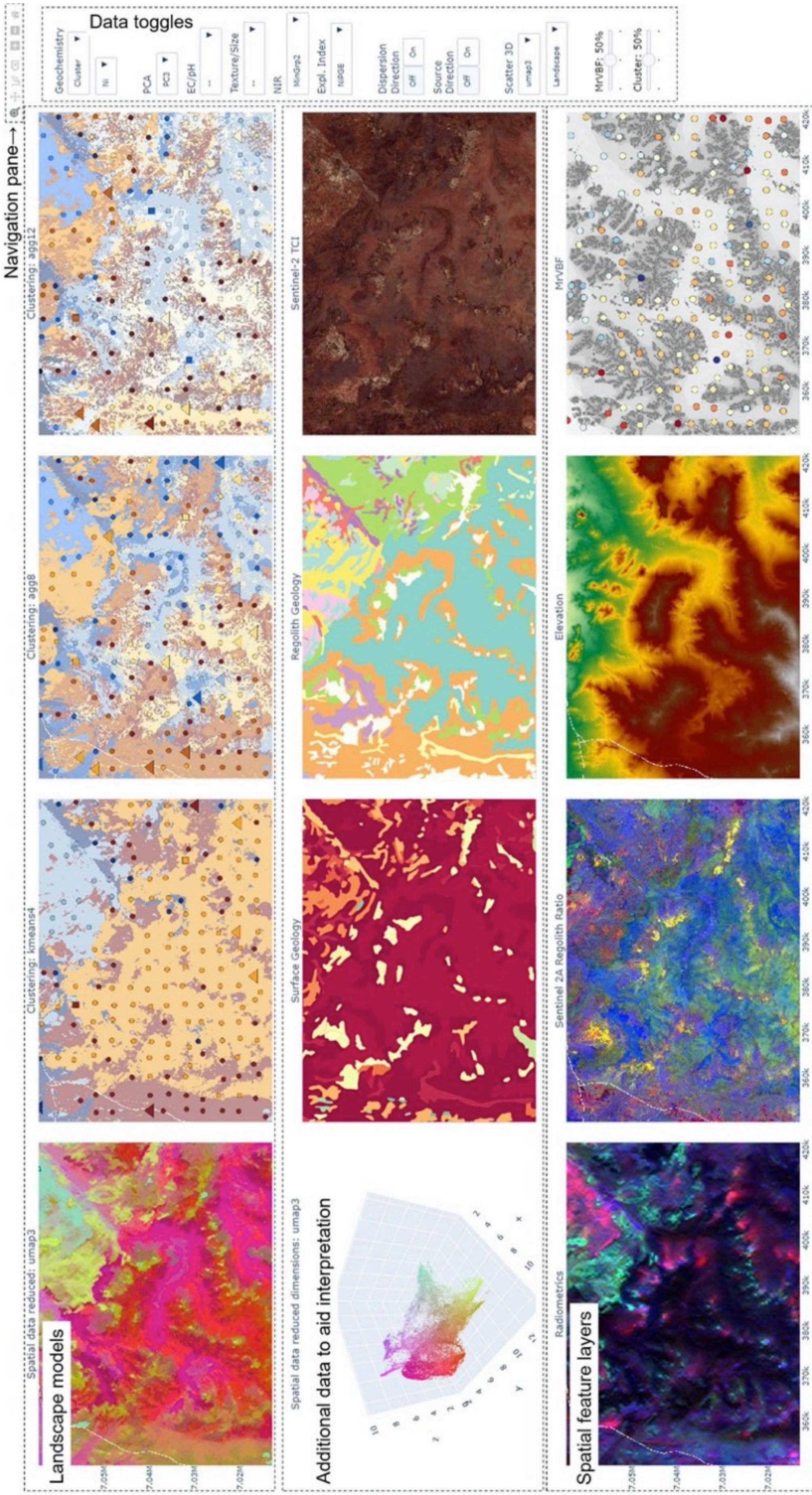
The full Next Gen Analytics workflow currently reads in exclusively UltraFine+® data. This is in part due to how the projects supporting the development have evolved, as well as its commonly greater mobile element sensitivity compared to other common methods, as fewer data points below the detection limit enables more sound statistical data interpretation. UltraFine+® also provides additional soil properties relating to particle size, pH and mineralogy which is part of the Next Gen Analytics data package. This R&D data package provides three different landscape models with 4, 8 and 12 landscape clusters for each site. This provides a standardised output that can be used for first-pass interpretation to indicate how much influence landscapes may have on statistical outlier calculations. Each data package for a given site contains CSV, GeoTIFF, PNG and shapefiles of outliers by (as well as independent of) landscape type, soil property data, exploration indices, principal component analysis of geochemical data, and DEM-derived source and dispersion directions. The data package also includes the Digital Sample Observer; an HTML-based prototype dashboard still under development that allows the explorer to easily view and interrogate all available products within the data package (Fig. 10).

Using machine learning-derived landscape models to interpret soil geochemical data by landscape type does not replace the need for in-depth regolith knowledge of an exploration area, but rather it provides a low-cost and low-impact tool in early stages of greenfields exploration to minimise time and effort while maximising outputs, ideally preventing overlooking targets and walking away from potentially prospective ground.

Future research will focus on refining the approach to address limitations such as the influence of built-up areas, densely wooded areas, and sample density. Ideally, this machine learning approach will be targeted towards a specific exploration area and commodity, and will benefit from tailored input layers, depending on specific exploration questions with a view to extending our knowledge of, and incorporating, local dispersion mechanisms. This may include a range of different or additional feature layers, such as other geophysics (magnetics, electromagnetics), and/or high-resolution company-owned data. The workflow will also be adjusted to other international settings (e.g., temperate and glacial terrains) addressing challenges in availability of regolith and landscape maps across the globe, as well as tailoring to different commodities and incorporating locally available spatial feature layers. Ideally, machine learning approaches for landscape context will be incorporated in earlier exploration stages such as during prospectivity reconnaissance and sample cam-

continued on page 19

Old Soils, New Targets: ... *continued from page18*



paign design. All of these aspects are currently being investigated, and research over the coming decades will undoubtedly increase confidence in these models, introducing them as a standard output during the exploration process.

Fig. 10: Screenshot of the Digital Sample Observer for the Kingston project area. Top row (from left to right): Reduced spatial feature layer data in 2-dimensional space, RGB coloured (the more similar the colour of pixels, the more similar the spatial feature properties for these pixels) and the resulting three landscape models with 4, 8 and 12 landscape classes. Middle row (from left to right): Reduced spatial feature layer data in 3-dimensional space followed by surface geology, regolith geology and satellite imagery for comparison to the models. Bottom row (from left to right): Spatial feature layers (radiometric data, Sentinel-2 derived regolith band ratios, DEM and MrVBF) used to derive the landscape models. Data toggles on the right-hand-side allow for viewing outliers by or independent of landscape type, as well as all other data products, in landscape context.

Acknowledgements

Thank you to the **EXPLORE** editors and reviewers, Beth McClenaghan and Paul Morris. The research to develop the UltraFine+® method and the Next Gen Analytics was supported by a broader project team including Tania Ibrahimi, Ian Lau, Bobby Pejcic, Jessica Stromberg, Neil Francis, Tenten Pinchand and Ravi Anand. The research team received financial support from many government and industry bodies, including the Minerals Research Institute of Western Australia, Geological Survey of Queensland, Geological Survey of South Australia, Geological Survey of New South Wales, Northern Territory Geological Survey, Geological Survey of Western Australia, Kalamazoo Resources, MCA Nominees, Iceni Gold, Siren Gold, Dreadnought

continued on page 20

Old Soils, New Targets: ... *continued from page 19*

Resources, De Grey Mining, Carnavale Resources, Fortescue Metals Group, Newmont, Northern Star Resources, Kairos Minerals, Emmerson Resources, Independence Group, Western Gold Resources, Capricorn Metals, Hexagon Energy Materials, Monger Gold, Strategic Energy Resources, Barton Gold, Ozz Resources, Anax Metals, Lodestar Minerals, Anglo Gold Ashanti, Antipa Minerals, Encounter Resources, First Quantum Minerals, Gold Road Resources and Southern Gold. In-kind support for the project was provided by CSIRO and LabWest.

While there is no direct financial interest in the results presented in this article, CSIRO, as the employer of the authors, owns UltraFine+® and this is licenced to LabWest as the current commercial provider. The CSIRO (not the authors) receives a financial royalty for this licence and subsequent UltraFine+® analyses.

References

Anand, R.R., Aspandiar, M.F. and Noble, R.R.P. 2016. A review of metal transfer mechanisms through transported cover with emphasis on the vadose zone within the Australian regolith. Ore Geology Reviews, 73, 394-416.

Anand, R., Lintern, M., Noble, R., Aspandiar, M., Macfarlane, C., Hough, R., Stewart, A., Wakelin, S., Townley, B. and Reid, N. 2014. Geochemical Dispersion Through Transported Cover in Regolith-Dominated Terrains - Towards an Understanding of Process. Society of Economic Geology Special Publication, 18, 97-126.

Anand, R.R., Churchward, H.M., Smith, R.E., Smith, K., Gozzard, J.R., Craig, M.A. and Munday, T.J. 1993. Classification and atlas of regolith landform mapping units, Exploration perspectives for the Yilgarn Craton, Australia. CSIRO Division of Exploration and Mining, Restricted Report 440R (unpublished).

Arne, D. and MacFarlane, B. 2014. Reproducibility of gold analyses in stream sediment samples from the White Gold District and Dawson Range, Yukon Territory, Canada. EXPLORE, 164, 1–10, https://www.appliedgeochemists.org/images/Explore/Explore_Number_164_Sept_2014.pdf

Bajc, A.F. 1998. A comparative analysis of enzyme leach and mobile metal ion selective extractions; case studies from glaciated terrain, northern Ontario. Journal of Geochemical Exploration, 61, 113–148.

Baker, P.M. 2015. Advances in reconnaissance sampling and geochemical analysis are creating new exploration opportunities in previously sampled terranes. In: Proceedings of the 27th International Applied Geochemistry Symposium 2015. Association of Applied Geochemists, Ontario, Canada, 11.

Bunting, J.A. 1980. Kingston, W.A.: Western Australia Geological Survey, 1:250 000 Geological Series Explanatory Notes, 16p.

Bunting, J.A. 1986. Geology of the eastern part of the Nabberu Basin, Western Australia. Western Australia Geological Survey, Bulletin 131, 130p.

Chao, T.T. 1984. Use of partial dissolution techniques in geochemical exploration. Journal of Geochemical Exploration, 20, 101–135.

de Souza Kovacs, N. and Jakica, S. 2021. 1:100 000 regolith geology regimes of Western Australia. Geological Survey of Western Australia, digital data layer.

Gallant, J., Dowling, T. and Austin, J. 2012. Multi-resolution Valley Bottom Flatness (MrVBF). v3. CSIRO. Data Collection. https://doi.org/10.4225/08/5701C885AB4FE

Gozzard, J.R. 2005. Part 3: Regolith-landform mapping using remotely sensed imagery. In: International Geochemical Exploration Symposium 2005 Workshop 1.3, Regolith mapping, workshop notes: Perth, Western Australia, IGES 2005, 73p.

Gray, D.J., Wildman, J.E. and Longman, G.D. 1999. Selective and partial extraction analyses of transported overburden for gold exploration in the Yilgarn Craton, Western Australia. Journal of Geochemical Exploration, 67, 51–66.

Hall, G.E.M. 1998. Analytical perspective on trace element species of interest in exploration. Journal of Geochemical Exploration, 61 (1–3), 1-19. https://doi.org/10.1016/S0375-6742(97)00046-0

Henne, A., Noble, R.R.P., Huang, F., Cole, D., Williams, M., Ibrahimi, T. and Lau, I. 2022. UltraFine+® Next Gen Analytics - Northern Territory Geological Survey - MacDonnell Ranges. CSIRO, Australia. https://doi.org/10.25919/hvzk-9x75

Jakica, S., de Souza Kovacs, N., Hogen-Esch, J. and Granado, I.M.T. 2020. 1:500 000 State regolith geology of Western Australia — compilation methodologies. Geological Survey of Western Australia, Record 2020/10, 22p.

MacGuiness, S.S. and Pye, K.S. 2000. Regolith materials, Kingston, W.A. sheet SG 51-10. In: Pye, K.J., Morris, P.A. and McGuiness, S.A. Geochemical Mapping of the Kingston 1:250 000 Sheet. Western Australia Geological Survey, 1:250 000 Regolith Geochemistry Series Explanatory Notes Plate 2. Department of Minerals and Energy, Perth, WA, Australia.

Martin, D.McB., Hocking, R.M., Riganti, A. and Tyler, I.M. 2014. 1:500 000 State interpreted bedrock geology of Western Australia, 2014, digital data layer. Geological Survey of Western Australia, <www.dmp.wa.gov.au/geoview>.

McInnes, L., Healy, J. and Melville, J. 2018. Uniform manifold approximation and projection for dimension reduction. arXiv:1802.03426

Morris, P.A. 2013. Fine fraction regolith chemistry from the East Wongatha area, Western Australia: tracing bedrock and mineralization through thick cover. Geochemistry: Exploration, Environment, Analysis, 13, 21–40 https://doi.org/10.1144/geochem2012-141

continued on page 21

Old Soils, New Targets: ... *continued from page 20*

Myers, J. S. and Hocking, R.M. 1998. Geological map of Westem Australia 1:2 500 000 (13th edition). Western Australian Geological Survey.
Noble, R., Cavaliere, M., Morris, P., Pinchand, T. and Hough, R. 2013. Determination of micro and nanoparticulate fraction gold in regolith. EXPLORE, 159, 1-13.
Noble, R., Lau, I., Anand, R. and Pinchand, T. 2018. MRIWA Report No. 462: Multi-scaled near surface exploration using ultrafine soils. Geological Survey of Western Australia, Report 190, 99 p.
Noble, R.R.P., Lau, I.C., Anand, R.R. and Pinchand, G.T. 2020a. Refining fine fraction soil extraction methods and analysis for mineral exploration. Geochemistry; Exploration, Environment, Analysis, 20, 113-128. https://doi.org/10.1144/geochem2019-008
Noble, R., Morris, P., Anand, R., Lau, I. and Pinchand, T. 2020b. Application of ultrafine fraction soil extraction and analysis for mineral exploration. Geochemistry: Exploration, Environment, Analysis, 20, 129-154. https://doi.org/10.1144/geochem2019-009
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, É. 2011. Scikit-learn: Machine learning in python. The Journal of Machine Learning Research, 12, 2825–2830.
Poudjom Djomani, Y. and Minty, B.R.S. 2019a. Radiometric Grid of Australia (Radmap) v4 2019 filtered pct potassium grid. Geoscience Australia, Canberra. http://dx.doi.org/10.26186/5dd48d628f4f6
Poudjom Djomani, Y. and Minty, B.R.S. 2019b. Radiometric Grid of Australia (Radmap) v4 2019 filtered ppm thorium. Geoscience Australia, Canberra. http://dx.doi.org/10.26186/5dd48e3eb6367
Poudjom Djomani, Y. and Minty, B.R.S. 2019c. Radiometric Grid of Australia (Radmap) v4 2019 filtered ppm uranium. Geoscience Australia, Canberra. http://dx.doi.org/10.26186/5dd48ee78c980
Pye, K.J., Morris, P.A. and McGuiness, S.A. 2000. Geochemical Mapping of the Kingston 1:250 000 Sheet. Western Australia Geological Survey, 1:250 000 Regolith Geochemistry Series Explanatory Notes. Department of Minerals and Energy, Perth, WA, Australia.
Sader, J., Benn, C., Zhivkov, N., Jeleva, T. and Anderson, R. 2018. Soil clay fraction geochemistry for surficial exploration: a case study from the Tethyan Belt. In: RFG 2018 – Resources for Future Generations Conference Proceedings, Vancouver, Canada. Society of Economic Geologists, Littleton, CO, Abstract 2469.
Scott, K.M. and van Riel, B. 1999. The Goornong South gold deposit and its implications for exploration beneath cover in Central Victoria, Australia. Journal of Geochemical Exploration, 67, 83–96, https://doi.org/10.1016/S0375-6742(99)00048-5
Wilford, J. and Roberts, D. 2021. Sentinel-2 Barest Earth imagery for soil and lithological mapping. Geoscience Australia, Canberra. http://dx.doi.org/10.11636/146125