
# REPORT NO. 462a

# M0462a Final Report

# The paradigm shift for minerals exploration using ultrafine soils and intelligent data integration tools

Results of research carried out as MRIWA Project M0462a

UltraFine+® Next Gen Analytics

at CSIRO

by

Ryan Noble, Anicia Henne, Dave Cole, Morgan Williams, Fang Huang, Ian Lau, Bobby Pejcic, Tenten Pinchand, Tania Ibrahimi, Alfredo López Alfageme, Jessica Stromberg, Neil Francis, John Hille, Tina Shelton, Selina Hutcheon

September 2024

Distributed by: MRIWA
1 Adelaide Terrace
East Perth WA 6004
to which all enquiries should be addressed

**MINISTER FOR MINES AND PETROLEUM**
**Hon David Michael MLA**

**CHIEF EXECUTIVE OFFICER, MINERALS RESEARCH INSTITUTE OF WESTERN AUSTRALIA**
**Nicole Roocke**

**CHIEF EXECUTIVE, CSIRO**
**Dr Doug Hilton AO**

**LEAD INVESTIGATOR, CSIRO**
Dr Ryan Noble

**AUTHORS:**
Ryan Noble, Anicia Henne, Dave Cole, Morgan Williams, Fang Huang, Ian Lau, Bobby Pejcic, Tenten Pinchand, Tania Ibrahimi, Alfredo López Alfageme, Jessica Stromberg, Neil Francis, John Hille, Tina Shelton, Selina Hutcheon

**REFERENCE**
The recommended reference for this publication is:
Noble, R., Henne, A., Cole, D., Williams, M., Huang, F., Lau, I., Pejcic, B., Pinchand, T., Ibrahimi, T., López Alfageme, A., Stromberg, J., Francis, N., Hille, J., Shelton, T., Hutcheon, S., 2024. M462a Final Report The paradigm shift for minerals exploration using ultrafine soils and intelligent data integration tools. Report EP2023-5262 CSIRO, Perth, 89 p.

**ISBN:** N/A
**ISSN:** mriwa

This report is available from the national Library of Australia

**PARTICIPATING ORGANISATIONS**
Minerals Research Institute of Western Australia, CSIRO, 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 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, Ozz Resources, Anax Metals, Barton Gold, New Age Exploration. Lodestar Minerals, LabWest.

**ABOUT THIS PUBLICATION**
This Report was prepared by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and commissioned by the Minerals Research Institute of Western Australia (MRIWA). Although MRIWA provided support to the project, the scientific content of the Report has been the responsibility of CSIRO.

**KEYWORDS AND TAGS**
gold, machine learning, microparticulate, regolith, landscape context, ultrafine, base metals



**Published 2024 by the Minerals Research Institute of Western Australia**
This report is published in digital format (PDF) and is available online at https://www.mriwa.wa.gov.au/research-projects/project-portfolio/

**CONTACT**
Minerals Research Institute of Western Australia
1 Adelaide Terrace
Perth WA 6000
+61 8 6180 4340
mail@mriwa.wa.gov.au
https://www.mriwa.wa.gov.au/

# Acknowledgements

This project received financial support from many government and industry bodies. Financial supporters include 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 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, Ozz Resources, Anax Metals, Barton Gold, New Age Exploration and Lodestar Minerals. In-kind support for the project was provided by CSIRO and LabWest.

Most critically, we thank the above contributors for their in-kind support in providing UltraFine+® analytical results from many, many soils across Australia. This support totalled several millions of dollars and without the large number of results the outcomes of this project would have been severely limited.

# Executive Summary

The UltraFine+® Next Gen Analytics research project developed significant advancements in greenfields mineral exploration. The outputs from this research, delivered for each project site in the UltraFine+® Next Gen Analytics data package, introduce a comprehensive set of measurable parameters and innovative data products for a more profound understanding of soil properties that have often been overlooked. This technology empowers mineral explorers to make informed decisions about where and when to direct their exploration efforts. Developed through a collaboration between CSIRO, LabWest, MRIWA, 22 industry sponsors, and five State and Territory geological surveys, the UltraFine+® Next Gen Analytics project aimed to transform the exploration landscape for precious, base, and critical metals in Australia. This was achieved by merging the cutting-edge UltraFine+® soil analytical methods with purpose-built data integration tools dubbed “Next Gen Analytics”.

Over the course of this research project, more than 80,000 soil samples were collected and analysed. A comprehensive workflow was created, encompassing both the UltraFine+® and Next Gen Analytics methodologies. This involved creating landscape models for each site and integrating soil analytical results into this landscape context for initial interpretation. All of these processes were coded into a Python workflow, allowing for the semi-automatic generation of a valuable data package. The Python code produces a range of outputs, including proxy landscape type maps, boxplots, outlier maps, principal component analysis, exploration indices, dispersion and source direction maps, and additional soil property maps. These outputs provide an initial interpretation of geochemical samples based on proxy regolith types and highlight potentially subtle anomalies of interest to mineral exploration.

While data packages were generated for each project site, not all of them can be made publicly available. Those that are not subject to confidentiality restrictions have been published as individual reports during the course of the research project. This report outlines the final research methodology and presents examples of the outputs through short case studies. The UltraFine+® workflow has predominantly been applied in Western Australia with additional case studies in Queensland, New South Wales, South Australia, Victoria and the Northern Territory. However, the research outputs have great potential for broader global applications, especially in areas with significant mineral resources concealed beneath shallow cover, and we demonstrated the approach on a case study in New Zealand.

Overall, the outlook for UltraFine+® Next Gen Analytics is promising. It has successfully ushered in a paradigm shift in precious, base, and critical metals exploration in Australia by enhancing the value of routine soil sampling in frontline exploration and is poised to influence mineral exploration practices for the foreseeable future.

The information summarised in this report has been compiled from various guides, project reports, and publications created during the research project. For further details, please refer to the project's online resources[1].

[1] https://research.csiro.au/ultrafine/resources/.

# Contents

# 1 Background

Much of Australia's remaining potential mineral wealth is masked by regolith cover that poses a challenge for future mineral exploration, especially in transported cover. The mobile element signature of interest for exploration in these materials is commonly contained in the < 2 µm "ultrafine" particle size fraction (Noble et al. 2020). This is likely due to the presence of "scavenging phases" such as clays, organic compounds and various oxides/oxyhydroxides that dominate this fine fraction (Hall 1998). The CSIRO in collaboration with LabWest developed the novel UltraFine+® method, a soil geochemistry workflow which is optimised for multi-element analysis of this ultrafine soil fraction. This improved soil geochemistry workflow compared with conventional soil analyses generates results with more contrast and increased concentrations of precious and transition metals (target elements and pathfinders), and removes the nugget effect, thereby enhancing the reproducibility and reliability of results (Noble et al. 2020). The UltraFine+® method was developed as part of the prior M0462 project.

One of the main aims of this M0462a UltraFine+® project was to add further "Next Generation Analytics" that leverage and expand on the M462 project soil methodology workflow by adding relevant spectral mineral proxies to the UltraFine+® method as standard analyses. This adds substantial additional data to the standard soil sample exploration package and enables exploration geologists to investigate the relationships between soil geochemistry and other physicochemical soil parameters and their relationship to buried mineralisation.

The second aim of the M0462a UltraFine+® project addressed a common challenge in the use of surface mineral exploration: Often soil geochemistry is interpreted with little regard for physicochemical soil parameters or landform settings, and how these relate to buried mineralisation. The UltraFine+® Next Gen Analytics research addresses this challenge by utilising machine learning approaches to produce landscape context for, and first-pass data interpretation of, these soil sample analyses by integrating soil parameters with regolith landscape models. This improves our ability to identify targets and false positives, as well as understand the spatial variance and influence of regolith types.

Assessing geochemical data in mineral exploration often focuses on understanding outliers, such as elevated copper, gold, or zinc concentrations. Often the largest concentrations are followed up, but landscape and soil types can significantly influence element concentrations. For example, high metal concentrations may be readily identifiable as outliers in a geochemical dataset where samples were collected over mineralisation in areas of shallow residual soils, while the same mineralisation has a much weaker elemental signal in samples collected over thicker depositional landscapes. With the ability to approximate landscape types from spatial data features via machine learning, the project outputs improved outlier identification within each landscape type via three main components:

a) Refining the UltraFine+® soil analytical method by lowering detection limits, adding additional elements, and adding additional soil properties (spectral mineralogy) to the standard soil analysis;
b) Developing the Next Gen Analytics workflow using machine learning to create landscape context from spatial data features for soil geochemical interpretation; and
c) Providing a comprehensive data package to project sponsors.

Following on from the successful M0462 project that developed the UltraFine+® method, 29 industry and government partners collaborated to provide the test sites, soil samples and technical knowledge to complete these components. The study sites are noted in Figure 1,

and to date 82,994 samples (not including duplicates and standards) have been collected and analysed using the UltraFine+® method.

The M0462a project expanded greatly over the three-year duration. Industry uptake and site studies were initially anticipated to include five to eight companies and a similar number of project sites with a total of 5,000 samples. The increased sponsorship and test sites (~40) responded to industry demand and resulted in project objectives aligned more strongly with industry needs for data delivery and first pass analytics, with less focus on detailed site studies. As a result, the data package component of the project outputs became the focus. The varied sites and target commodities, sample spacing, and regions also meant the team had to be flexible and adapt to changing conditions such as stream sediment samples, different target and pathfinder elements and industry interests. The method was refined early in the project and the data delivery was initially intended to include a moderate amount of machine learned regolith maps (as GeoTIFF files that are ArcGIS, MapInfo compatible) using the results of the UltraFine+ ® soil analytical elemental data in the context of proxy landscapes and the additional spectral reflectance and physico-chemical properties data. Industry expectations increased the need to deliver most of the data in formats that were ready to use and generate a first-pass interpretation tool to simplify and fast track the process to reviewing and understanding the data outputs. This report documents the general (non-IP protected) procedure and resulting UltraFine+® Next Gen Analytics data package, along with various examples of how the outputs are delivered and used to interpret soil geochemical results of the ultrafine fraction in landscape context.

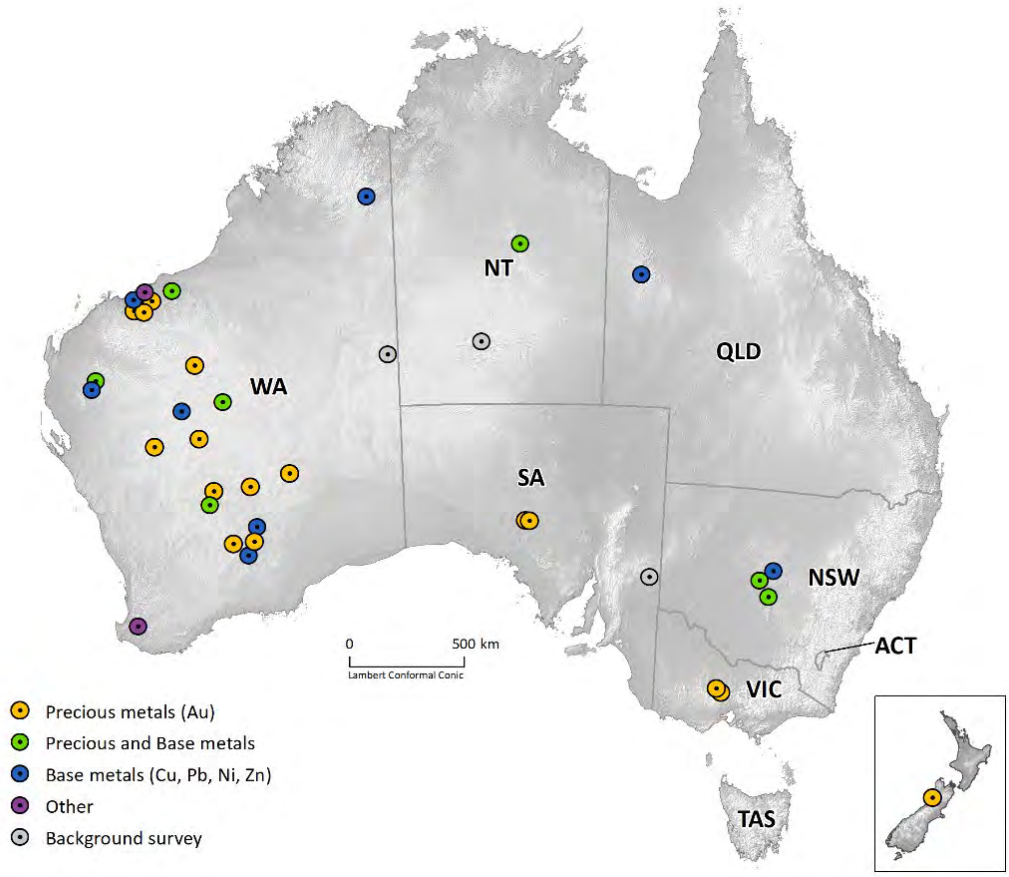



**Figure 1**: Site locations for the UltraFine M0462 and M0462a projects with 82,994 samples (not including duplicates and standards) that were part of the M0462a Next Gen Analytics project.

## 1.1 Research Outline

This research program has contributed significantly toward meeting the challenge of exploration through cover, to enhance exploration in Australia and improve our understanding of surface soil geochemistry and different commodity targets in the context of Australia's diverse regolith settings. The work was based on areas of interest for geological surveys and industry sponsors; however, the machine learning approach was adjusted as the project developed to be applicable across the Australian continent. CSIRO worked with stakeholders to deliver the following specific project outcomes:

a) Assessment of the UltraFine+®method compared to older existing data sets and other methods, examining the effects of cover and lithology/regolith types using near surface sampling techniques. This included comparison of orientation surveys and regional data with the newly developed UltraFine+® method.
b) Apply the UltraFine+® method in under-explored greenfield settings to generate initial beta test outputs around regolith proxy mapping, dispersion directions and spatial uncertainty estimations using machine learning (ML) and integrating CSIRO Background IP in regolith geoscience. The initial target outputs were data tables and GeoTIFF files.
c) Refine soil sampling and analysis strategy for industry and demonstrate the value-add of the UltraFine+® workflow.
d) Improve on the commercially available UltraFine+® workflow with the addition of Pt, Pd and Fourier Transform Infrared Spectroscopy (FTIR) analyses for organic C proxies and other broader wavelength spectral mineral proxies.

Project activities were structured around four core milestones, outlined in the summary below with further detail contained in Section 2 of this report:

### 1.1.1 Project Milestone 1: Refinement of the original UltraFine+® workflow

Analytical workflow adjustments utilised comparison of aqua regia to 4-acid extraction, ALS SuperTrace and clay separation to assess the ability to integrate geochemical datasets that used these methods with datasets derived via the UltraFine+® method. Additional comparison to partial extraction methods like Mobile Metal Ion was also conducted. This aided in the assessment of successful extraction of Pt and Pd with the UltraFine+® method, and assessment of the recovery of W, Sn and REE that are resistant to acids.

### 1.1.2 Project Milestone 2: Adding essential spectral data features and relevant FTIR mineral proxies to the UltraFine+® workflow and options for improved sample submission

An organic C estimate from vis-NIR analysis and FTIR analysis was compared. The addition of FTIR silicate and carbonate mineral proxy fields into the standard UltraFine+® method reporting was achieved as part of the workflow refinements, as was a template for sample submission for the Next Gen Analytics to enable data ingestion and move towards a "click and collect" functionality.

### 1.1.3 Project Milestone 3: Regolith landform estimates and ground truth comparison

Geospatial data related to regolith landform properties were identified and used to derive estimates of regolith type (depositional, erosional, residual as a minimum). Cover depth estimates were also assessed and eventually included as discriminators. Collation of key site studies from project partners and ground-truthing of the modelled estimates of regolith setting were carried out. A robust set of parameters to estimate regolith types from data that is sensed and not human interpreted (e.g., no regolith/geology maps used) was successfully derived.

### 1.1.4 Project Milestone 4: Machine Learning (ML) comparisons and building algorithms based on human experience in regolith geochemistry

Multiple unsupervised ML approaches were compared to generate regolith landscape models. In addition, the team produced exploration indices and scalars such as adjusting for organic content, Fe oxides and other parameters from the raw UltraFine+® data and derived Principal Component Analysis as a multielement exploration approach for each site. In total 33 indices and ratios were added as well as five principal components for a multi-element exploration approach to the outputs, in addition to soil parameters and outliers by landscape type.

Through the above stages, a robust soil analytical method with improved element detection and algorithms capable of providing landscape context for the geochemical results were developed. The project set out to deliver a proof of concept for soil surface exploration in machine-learned landscape context in Australia for future incorporation into a fully automated "click and collect" delivery of data which was initially outside the scope of the project. However, due to the additional project sites integrated into the analysis, the additional funding contributions to the project allowed for additional project staff and the project delivered a robust prototype beyond the expectations of the initial project outputs.

Much of the project findings were reported individually to sponsoring companies for specific sites that provided case studies. These do not form part of the final report due to industry confidentiality. However, each site was used to inform the method developments and machine learning workflow improvements and enabled the development and demonstrated application of the UltraFine+® Next Gen Analytics data package. A large volume of information is publicly available on the UltraFine+® website[2] and includes links to relevant publications including industry guidance notes to support method adoption. This report documents the final methodology developed over the course of the three-year research project and outlines the key outputs that are part of the UltraFine+® Next Gen Analytics data package. This report also provides key information applicable to the usability of these outputs and discusses the overall scientific contribution of the data that was collected and analysed as part of this project.

[2] https://research.csiro.au/ultrafine/

# 2 Methodology - The UltraFine+® soil analytical and Next Gen Analytics machine learning workflows

As part of this research project, over 80,000 soil samples were collected and analysed, and over 40 landscape models have been generated. These were used to generate two workflows, (1) the UltraFine+® workflow which includes all soil analytical steps conducted upon receipt of samples at the laboratory, and (2) the Next Gen Analytics workflow which includes the generation of a landscape model for each site followed by the integration of soil analytical results into landscape context for first-pass interpretation. All these steps were coded into a python workflow for semi-automatic generation of a data package. The final python code that is the core deliverable of this research project produces proxy landscape type maps, boxplots and maps for outliers by landscape type, outliers independent of landscape type, principal component analysis of soil geochemical results, exploration indices, dispersion and source directions and maps for additional soil properties. Data packages were generated for each project site and this data is confidential to the respective industry sponsors. However, the final research methodology is outlined below and examples of the outputs with short case studies to highlight the overall research results can be found in the Results and Findings section. Overarching research results are discussed in Section 3 Results and Discussion.

Most of the information collated below is reproduced from several individual short guides, project reports and publications created in several stages during the course of the research project, all of which are available online[3].

## 2.1 Sample Collection

Soil samples for this project were collected by project sponsors (exploration companies, Australian State and Territory Geological Surveys and their contractors). Only a very limited number of samples from early studies were sampled by CSIRO staff. While, therefore, no specific details of soil sampling approaches are available for most samples, all companies were provided with a detailed sampling guide. The guide recommends that sampling conducted as part of the UltraFine+ Research Project should use the following as a minimum:

### 2.1.1 Main Sample Media

At each site, a shallow soil sample should be collected from approximately 2–10 cm in depth. Depth may vary and a lower soil horizon or similar morphological feature can be a suitable target. Consistency with other soil sampling protocols in general is recommended, although the benefits of UltraFine+® is that soil morphological changes tend to be compensated for and the mass of soil required is less than other methods and requires little preparation. The sample collected should be approximately 200 g.

### 2.1.2 Sample Collection Routine

The following routine is recommended for sample collection:

- A clear space in the landscape should be selected, photographed, and documented. Typical field notes should include date, time, conditions, regolith setting, vegetation, and geology types etc.

[3] https://research.csiro.au/ultrafine/resources/

- The top 1 cm should be scraped away using a plastic trowel. The area removed should be approximately 15 cm x 15 cm.
- A further 5-10 cm should be dug using a posthole shovel or a plastic scoop. The ultrafine soil fraction is then collected from this material. In certain areas and soil types a geo-pick may make it easier to break up ground to depth and homogenise prior to sieving.
- Any coarse material from the soil >2 mm should be sieved out of the soil sample and discarded. The remaining (<2 mm size fraction) 200 g sample should be placed in a paper Geotech sample bag. Other bag types can be used but is important to have air dry samples, and breathable paper is better than plastic (for drying purposes).
- Following collection of materials, the small hole should be back filled and returned to a near flat surface

### 2.1.3 Sample Locations

Samples should be collected based on pre-planned spacing and avoid sampling areas that have clear disturbance or contamination such as animal burrows, old drilling spoil, or mine/agriculture infrastructure. If working in agricultural settings, sample below the plough depth (20 cm).

### 2.1.4 Sample Preparation

Samples should be collected when the weather is dry or dried soon after collection. This can be done in an oven at <80 °C and preferably 50 °C. Once dry, sample bags should be closed and sent to LabWest Minerals Analysis Pty Ltd for analysis.

All soil samples sent to LabWest should request the UltraFine+® method, the UltraFine+® standard, and note that they are part of the CSIRO Next Gen Analytics project. This uses a separation technique to extract the <2 µm particle size fraction and provide geochemistry, spectral mineralogy, particle size distribution and several other parameters.

## 2.2 UltraFine+® Laboratory Soil Analysis workflow (Milestones 1 and 2)

All soil samples were analysed using the UltraFine+® method (Noble et al. 2020) at LabWest Pty Ltd, Perth, Australia. The complete workflow requires <40 g of soil and includes particle size distribution analysis, pH and electrical conductivity (EC) measurements, separation of the ultrafine (<2 µm) size fraction, multi-elemental analyses, and spectral reflectance mineralogy. During this research project the original methodology was improved by adding four elements (I, Br, Pd and Pt) and lowering the detection limit of 33 elements. In addition, several spectral soil property measurements were added or refined. The final soil analytical workflow developed as part of this research project is documented below.

### 2.2.1 Bulk Soil Properties

Electrical conductivity and pH were measured on bulk sample slurries using a TPS AQUA-CP/A meter. Slurries were prepared using de-ionised water with a 1:5 w/w soil-to-water ratio.

### 2.2.2 Particle Size Analysis

Particle size analyses were conducted on the bulk sample using a Malvern Mastersizer 2,000 in suspension.

### 2.2.3 UltraFine+® Fine Fraction Separation

The ultrafine fraction (<2 µm; clay fraction) was extracted from each bulk sample via suspension in de-ionised water, addition of a dispersant and subsequent centrifugation and drying (documented in Noble et al. 2020).

### 2.2.4 UltraFine+® Extraction

The ultrafine fraction (<2 µm) of all soil samples was processed using a microwave-assisted aqua regia digestion (LabWest MAR-04) at LabWest Pty Ltd, Perth, Australia. The extractions were analysed for a suite of elements (Table 1) using ICP-OES (Perkin Elmer Optima 7300DV) and ICP-MS (Perkin Elmer Nexion 300Q). The microwave-assisted aqua regia digestion uses 0.2 - 0.4 g of soil with a 100 % mixture of 3:1 concentrated HCl:$HNO_3$. Unlike conventional extraction methods, the material is heated in a closed Teflon tube in an Anton Paar Multiwave PRO Microwave Reaction System for increased metal recovery (Noble et al. 2020).

**Table 1**: The UltraFine+® analytical stages and outputs generated. *Added to the workflow after some initial project data acquisition. †Analyses via ICP-OES.

| ANALYSES | OUTPUTS |
|---|---|
| **Microwave-assisted aqua regia on ultrafine fraction (LabWest MAR-04)** | Ag, Al†, As, Au, Ba†, Be, Bi, Br*, Ca†, Cd, Ce, Co, Cr†, Cs, Cu, Fe†, Ga, Ge, Hf, Hg, I*, In, K†, La, Li†, Mg†, Mn†, Mo, Nb, Ni, Pb, Pd*, Pt, Rb, Re, S†, Sb, Sc†, Se, Sn, Sr†, Ta, Te, Th, Ti†, Tl, U, V†, W, Y, Zn, Zr |
| **Bulk soil properties** | EC, pH |
| **Particle size distribution on bulk sample (LabWest SIZE-01)** | Size fractions <2 µm, <50 µm, <125 µm, <250 µm, <1000 µm, <2000 µm, >2000 µm; d(0.1); d(0.5); d(0.9); specific surface area (SSA) |
| **Visible-near-infrared (VNIR) on ultrafine fraction (LabWest NIR/SWIR)** | Main minerals; kaolinite crystallinity; iron oxide species; relative abundance of iron oxide, kaolinite, white mica and aluminium smectite, iron substitution in kaolinite, chlorite and dark mica, iron and magnesium smectite, mafic minerals with OH; white mica and aluminium smectite composition; palygorskite; Munsell® colour; hue; saturation; intensity |
| **Fourier transform infrared spectroscopy (FTIR) on ultrafine fraction (LabWest FTIR*)** | Clay, quartz, and carbonate abundances; total organic carbon; gibbsite index |

### 2.2.5 Visible Near-Infrared Reflectance and Short-wave Infrared Reflectance (VNIR-SWIR)

Visible near-infrared and short-wave infrared reflectance measurements were acquired on the ultrafine fraction of soil samples (< 2 µm) using a Spectral Evolution RS-3500 spectrometer (Serial Number 18980N3). The spectrometer measures electromagnetic radiation reflected off materials relative to that of a known reference material. The instrument collects spectra in the 350–2500 nm wavelength region, with a resolution of 2.8 nm at 700 nm, 8 nm at 1500 nm and 6 nm at 2100 nm. The spectral bandwidths of the RS-3500 are 1.3 nm at 700 nm, 3.5 nm at

1500 nm and 2.3 nm at 2100 nm, which are resampled to 1 nm to provide 2151 bands. A calibrated piece of sintered polytetrafluoroethylene (PTFE, also known commercially as Spectralon or Fluorilon) was used as the reflectance standard and measured before each set of soil measurements. The samples were measured with a bifurcated probe with a halogen light source. Each sample measurement consisted of 10 scans, averaged into a single measurement. Spectra were processed using The Spectral Geologist (TSG™) software[4] to extract the main features that were reported as part of the UltraFine+® output (Table 1).

### 2.2.6 Fourier Transform Infrared Spectroscopy (FTIR)

Mid-infrared spectroscopy measurements were performed using an Alpha Fourier transform infrared spectrometer (Bruker) which is equipped with a diamond-based attenuated total reflectance (ATR) sensing interface. ATR spectra were collected between 4000 to 360 $cm^{-1}$ using 16 scans at a resolution of 4 $cm^{-1}$ and a deuterated lanthanum triglycine sulfate (DLaTGS) detector. ATR measurements were conducted on the ultrafine samples by pressing about 10 mg of powder onto the diamond surface using the attached press. The surface of the diamond ATR was cleaned with ethanol and wiped dry with a laboratory cleaning tissue (Kimwipe). Background spectra were obtained in air with no sample present on the diamond surface prior to the sample measurement. All FTIR measurements were undertaken at room temperature (20 ± 3 ºC). Prior to measurements, the interferogram signal (ADC counts) was monitored and only when it passed internal QA/QC test (ADC counts / amplitude of 2800 ± 50) were spectra recorded. A reproducibility study was undertaken by recording replicate spectra and this revealed spectral intensity variations of less than 15%. The processing of spectra was performed using both OPUS software version 7.2 (Bruker) and The Spectral Geologist (TSG™) version 8.1. The ATR spectra were not corrected for changes in penetration depth/pathlength. The infrared absorption features (i.e., peak position, intensity) were determined and the band assignments are based on several references (Muller et al., 2014, Ge et al., 2014). A mineral standard (UFF_320) was used to check and monitor the analytical performance of the FTIR spectrometer with time.

## 2.3 Next Gen Analytics Machine Learning Workflow (Milestones 3 and 4)

The Next Gen Analytics workflow is a semi-automated data integration and analysis tool for combining soil analytical results with spatial feature layers. It consists of a Python library, a set of scripts for various tasks (including data import, processing and transformation, statistical analysis and machine learning, data visualisation and export) and a command-line application which provides the primary interface. The workflow configurations for individual projects were separately managed and versioned. The Python code and all associated workflow configuration remain CSIRO IP. The various stages of the workflow are described in more detail below (Figure 2).

[4] https://research.csiro.au/thespectralgeologist/

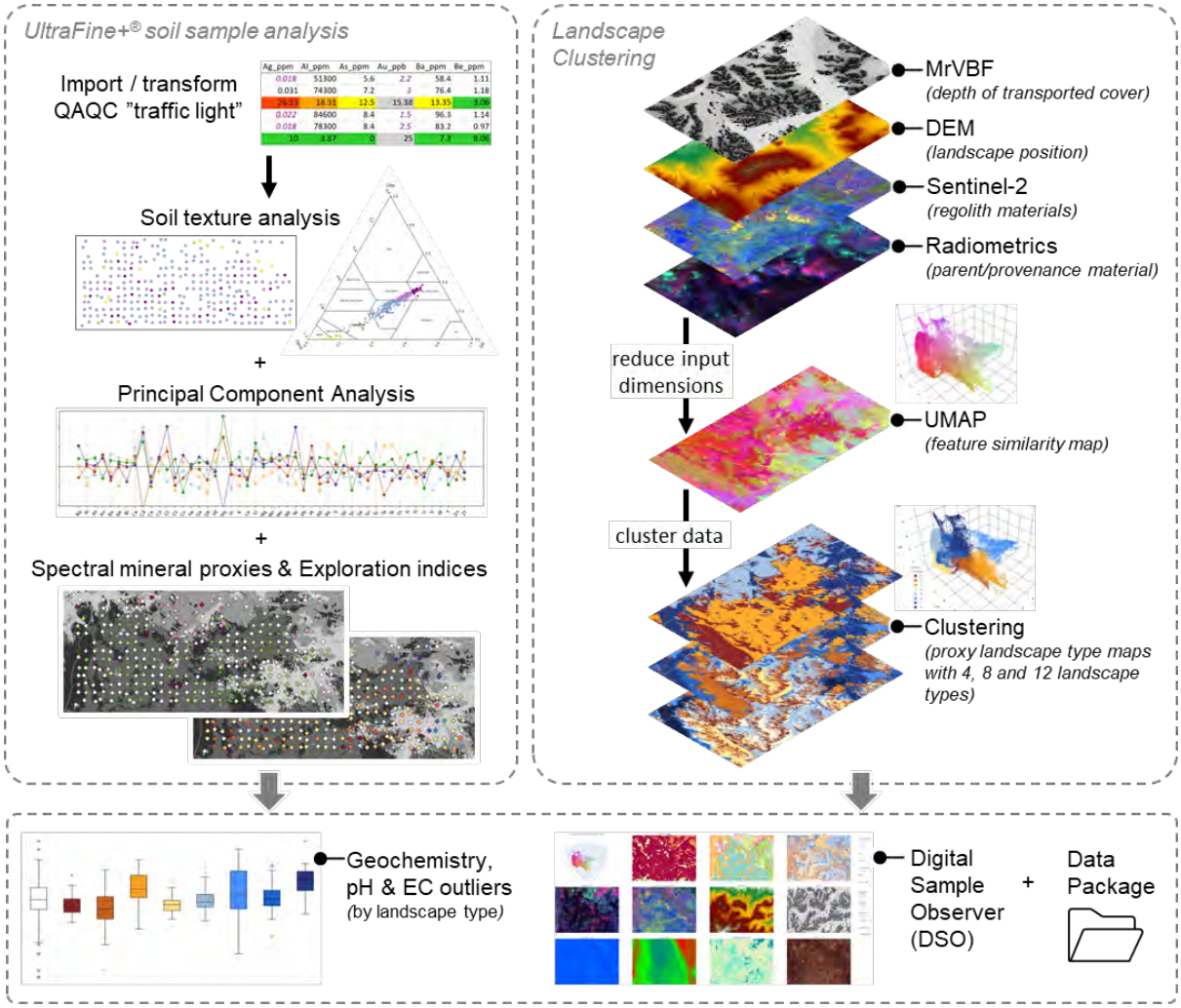



Figure 2: Simplified sketch of the major components of the Next Gen Analytics machine learning workflow deriving first-pass interpretation of the UltraFine+® soil analytical results for a given project site.

### 2.3.1 Data import and transformation

For the semi-automatic application of machine learning approaches and statistical evaluation of data, consistent data inputs without errors are paramount and often require manual handling, especially when data is collected from multiple sources and supplied in varying formats. To this end consistent input formats were generated in consultation with project partners for both the UltraFine+® data outputs and ancillary data (e.g., sample location, IDs and type). Scripts were generated to detect and remove human errors. Data transformation included the replacement of all values below detection limit in geochemical datasets with half the detection limit value for each respective element. This approach is common industry practice and has limitations (Grunsky & de Caritat, 2020; Martín-Fernández et al., 2003; Sanford, 1993), but is commonly considered fit for purpose over other simple alternatives for censored data. For outlier definition and principal component analysis, as well as regolith indices, data was also log-transformed. For spectral mineralogy datasets, the concept of detection limits does not always apply. Rather, more complex issues such as spectral interference and non-linear concentration ranges require consistent data handling for evaluation. Hence, where data was affected by either of these, a script replaces such data with a NULL.

### 2.3.2 Automated Quality Assurance/Quality Control (QA/QC)

It is important to evaluate the precision and accuracy of soil analyses prior to interpretation of results and this can be a labour-intensive process for larger datasets. As part of the Next Gen Analytics machine learning workflow a range of metrics were specifically designed to quantify and assess the quality of soil analytical data derived from the ultrafine fraction. This included the assessment of reproducibility of measurements on a specially designed reference material (QC 320) and through the assessment of measurements on duplicate samples. These were integrated into the Next Gen Analytics workflow via dedicated scripts that quantify and assess the quality of soil analytical data, and replace the error-prone, manual generation of Excel-based QA/QC reports. The outputs are conditionally formatted (colour-coded) excel sheets that provide an overview of the repeatability inherent in both the sampling (field duplicates, assessing the combined precision of sampling and analysis) and analysis (reference material, providing constraints on how accurate the provided values are likely to be). They contain individual results and statistical summaries for each project site. The underlying code handles subtleties of filtering and formatting as specified in the QA/QC configuration, as well as the generation of the formatted Excel reports. Various metrics for continuous, categorical, angular/polar and ordinal data are used for the relevant features provided as part of the expanded UltraFine+® analytical method. This system classifies the measures of repeatability of the analyses, functioning as an initial indicator of their precision via a "traffic light system". The metrics developed and how to interpret them are outlined below. More extensive guidance resources are accessible online[5] .

#### *2.3.2.1 Duplicates*

The Next Gen Analytics workflow creates a conditionally formatted Excel spreadsheet that contains four tabs to evaluate duplicate geochemistry, sizing, VNIR and FTIR analyses.

The traffic light system for duplicate pairs shows a rapid overview of the repeatability of analyses and highlights anything that warrants further investigation. For duplicate analyses, the traffic light system has four (geochemistry and sizing) or three (VNIR and FTIR) colours: If all indicators are designated as “green,” there is no cause for concern. Conversely, the designations “yellow,” “(orange),” and “red” denote a growing order of potential concern. It is imperative to acknowledge that analytical precision in proximity to the detection threshold is commonly subject to compromise. Consequently, a distinctive purple font was utilised to signify values falling below 10 × the detection threshold. Additionally, data rendered in a green font signifies values falling beneath the detection threshold, and such values are substituted with half the detection limit value (Figure 3). More information on individual duplicate pairs is accessible by clicking on the + symbol to open a drop-down menu (Figure 3). Consistent red flags for a specific duplicate pair, especially where there are also yellow and orange flags is a strong indication that duplicate pair data is mismatched and commonly is not a duplicate pair. Metadata (sample ID’s, coordinates, and field notes) are useful to consult in such cases. If specific element analyses are consistently flagged across multiple duplicate pairs, and duplicate metadata is correct, this should result in following up with the laboratory for clarification on the precision of the analyses.

[5] https://research.csiro.au/ultrafine/resources/how-to-guides/

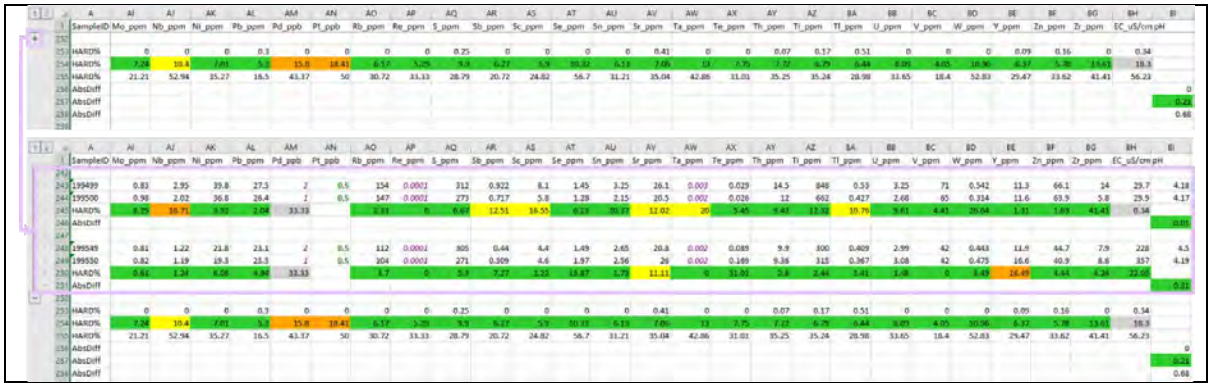


**Figure 3:** Example duplicate pair QA/QC summary output and expanded drop-down menu for individual duplicate pair geochemical analyses.

The traffic light colour system uses conditional formatting of the half absolute relative difference (Table 2) for duplicate pairs of geochemical elements and EC, and of the absolute difference (Table 2) for duplicate pairs of pH values (as it is a log scale). The specific rules for colour coding used in the outputs are outlined in

**Table 3** and take the reliability of the UltraFine+® analytical method based on the extensive sample analyses done in this project into account.

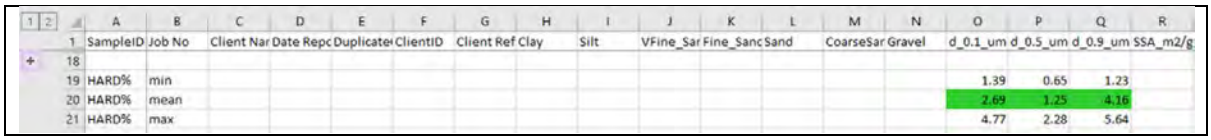
**Table 2**: Terms, parameters and calculations for QA/QC used in Next Gen Analytics outputs.

| Parameter | Formulae | Definition |
|---|---|---|
| **HARD%** | $100\frac{\lvert x_a - x_b \rvert}{\mathrm{x_a} + x_b}$ | Half relative difference between the respective value for duplicate pairs |
| **HARD% mean** | | The mean of all half relative differences between the respective values for all duplicate pairs |
| **RankMetric mean** | | Rank differences between two ordinal values, used here to compare the modal Gibbsite index (with ordinal values [NULL, 1, 2, 3]) |
| **IsEqual mode** | | Modal value of the equality check for all duplicates, indicating if the majority of duplicate pairs give equal or unequal values |
| **AbsDiff** | $\lvert x_a - x_b \rvert$ | Absolute difference between duplicate pair value |
| **AbsDiff mean** | | The mean of absolute difference of a respective value for all duplicate pairs |
| **Angle mean** | $\cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{\lVert\vec{u}\rVert\cdot\lVert\vec{v}\rVert}\right)$ | Angular difference between two vectors; here used for the hue component of colour as encoded as an angle |
| ***<parameter>* min** | | The minimum value of all respective parameter differences between the respective values for all duplicate pairs |
| ***<parameter>* max** | | The maximum value of all respective parameter differences between the respective values for all duplicate pairs |

**Table 3:** Conditional formatting rules for geochemical duplicate pairs analysed with UltraFine+®. *Al, Ca, Ce, Fe, S, Ta, Br, Sc; **Cr, Ti, Zr, W, Te, Sn, Se, La, Hf; ***Au, Pt, Pd; ****Ag, As, Ba, Be, Bi, Cd, Co, Cs, Cu, Ga, Ge, In, K, Li, Mg, Mn, Mo, Nb, Ni, Pb, Rb, Re, Sb, Sr, Th, Tl, U, V, Y, Zn; †Field will be grey if analyses are <300 µS/cm

| Problematic elements* | Resistate elements** | Precious metals*** | | Hg | | EC | pH | All other elements**** |
|---|---|---|---|---|---|---|---|---|
| | | If >10 ppb | If <10 ppb | If >0.2 ppm | If <0.2 ppm | >300† | (unit difference) | |
| <15 % | <50 % | <10 % | | <10 % | | <50 % | <0.5 | <10 % |
| >15 % | >50 % | >10 % | >10 % | >10 % | >10 % | >50 % | >0.5 | >10 % |
| >30 % | >75 % | >15 % | | >15 % | | >75 % | >0.75 | >15 % |
| >45 % | >100 % | >20 % | | >20 % | | >100 % | >1 | >20 % |

Duplicate pair data for sizing and spectral data is generated with three sizing parameters: green, orange and red. Like the geochemistry QAQC, individual duplicate pairs can be checked by clicking on the + symbol to open the drop-down menu (Figure 4). Duplicate pair QA/QC for sizing data has been developed on the 10th, 50th and 90th percentile of the particle size with the traffic light thresholds shown in Table 4 for the 3 parameters.



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | SampleID | Job No | Client Nar | Date Repc | Duplicate | ClientID | Client Ref | Clay | Silt | VFine_Sar | Fine_Sanc | Sand | CoarseSar | Gravel | d_0.1_um | d_0.5_um | d_0.9_um | SSA_m2/g |
| 18 | | | | | | | | | | | | | | | | | | |
| 19 | HARD% | min | | | | | | | | | | | | | 1.39 | 0.65 | 1.23 | |
| 20 | HARD% | mean | | | | | | | | | | | | | 2.69 | 1.25 | 4.16 | |
| 21 | HARD% | max | | | | | | | | | | | | | 4.77 | 2.28 | 5.64 | |

**Figure 4:** Example QA/QC summary output for duplicate pair sizing analyses. Individual duplicate pair comparisons can be viewed by clicking on the + sign.

**Table 4:** Conditional formatting rules for duplicate pair sizing analyses.

| D_0.1_µm | d_0.5_µm | d_0.9_µm |
|---|---|---|
| <10 % | <10 % | <10 % |
| >10 % | >10 % | >10 % |
| >15 % | >15 % | >15 % |
| >20 % | >20 % | >20 % |

Routine protocols for QA/QC of VNIR and FTIR data at the post-processing stage are not universally established. Hence, the project data was used to derive conditional formatting rules for all scalars as outlined in Table 5 for VNIR and Table 6 for FTIR. Where no data is reported for specific scalars, cells are blank. No comparison is applied to Colour, because the Munsell® Colour is derived from the three parameters Hue, Saturation and Intensity.

**Table 5:** Conditional formatting rules for VNIR duplicate pairs. *Fe-Kln, Chl, FeMgClay, OH-mafic, Paly

| MinGrp1 | MinGrp2 | FeOx | Hem_Geo | Kln_abun | Kln_cryst | WM_AS_abun | WM_AS_comp | Other scalars* | Hue | Saturation | Intensity |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Same | Same | <8% | <3nm | <7% | <0.35% | <6% | <1 nm | <5% | <2% | <2% | <7% |
| N/A | Different | 8-16% | 3-4.5nm | 7-13% | 0.35-0.7% | 6-12% | 1-1.5 nm | 5-10% | 2-4% | 2-4% | 7-14% |
| Different | N/A | >16% | >4.5nm | >13% | >0.7% | <12% | >1.5nm | >10% | >4% | >4% | >14% |

**Table 6**: Conditional formatting rules for FTIR duplicate pairs.

| Clay %, Quartz %, Carbonate %, TOC % | | Gibbsite Index |
|---|---|---|
| if values <5 % | if values >5 % | |
| <10 % | <20 % | Same |
| >10 % | >20 % | N/A |
| >20 % | >50 % | Different |

#### 2.3.2.2 *Standards*

Similar to the duplicate analyses assessment, the Next Gen Analytics workflow creates a conditionally formatted Excel spreadsheet from the CSIRO UltraFine+® reference material (QC 320) and includes three tabs for geochemistry, VNIR and FTIR analyses (note that the standards are not analysed for size distribution as the UltraFine+® standard is specifically designed to be <2 µm).

For geochemistry standards, the results are either green or red to indicate whether standard analyses are within one and two standard deviations from the mean of the reference standard, respectively (Figure 5). The mean reference standard values are derived from 312 analyses of the UltraFine+® standard reference material (QC_UFF_320). Where elements are flagged in red, the first consideration is imprecision close to the detection limit, as this might compromise results. For easy assessment a colour coding was applied to individual element analyses. Purple font indicates that a value is less than 10 times the detection limit (Figure 5) and analyses in green font indicate values below the detection limit (replaced by half the detection limit). If values are well above 10 × the detection limit value, red flags for individual elements can indicate that this element may have been consistently under- or over-reported, and this should be considered for later interpretation of geochemistry, especially if trying to merge different data sets.

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AnalysisID | Job No | Client Name | Date Repc | ClientID | Client Ref | Ag_ppm | Al_ppm | As_ppm | Au_ppb | Ba_ppm | Be_ppm | Bi_ppm |
| 3 | ALW001234-QC_UFF_320-0 | ALW001234 | Gold Rocks | ######## | QC_UFF_320 | Mt Gold | 0.024 | 77700 | 12.4 | 9.5 | 71.9 | 1.34 | 0.46 |
| 4 | ALW001234-QC_UFF_320-1 | ALW001234 | Gold Rocks | ######## | QC_UFF_320 | Mt Gold | 0.035 | 106000 | 14.2 | 9.5 | 99 | 1.42 | 0.327 |
| 5 | ALW001234-QC_UFF_320-2 | ALW001234 | Gold Rocks | ######## | QC_UFF_320 | Mt Gold | 0.023 | 74200 | 11.7 | 11.9 | 72.7 | 1.41 | 0.376 |
| 6 | | | | | | | | | | | | | |
| 24 | 1SD | | | | | | GOOD | GOOD | GOOD | GOOD | CHECK | GOOD | GOOD |
| 25 | 2SD | | | | | | OK | OK | OK | OK | OK | OK | OK |

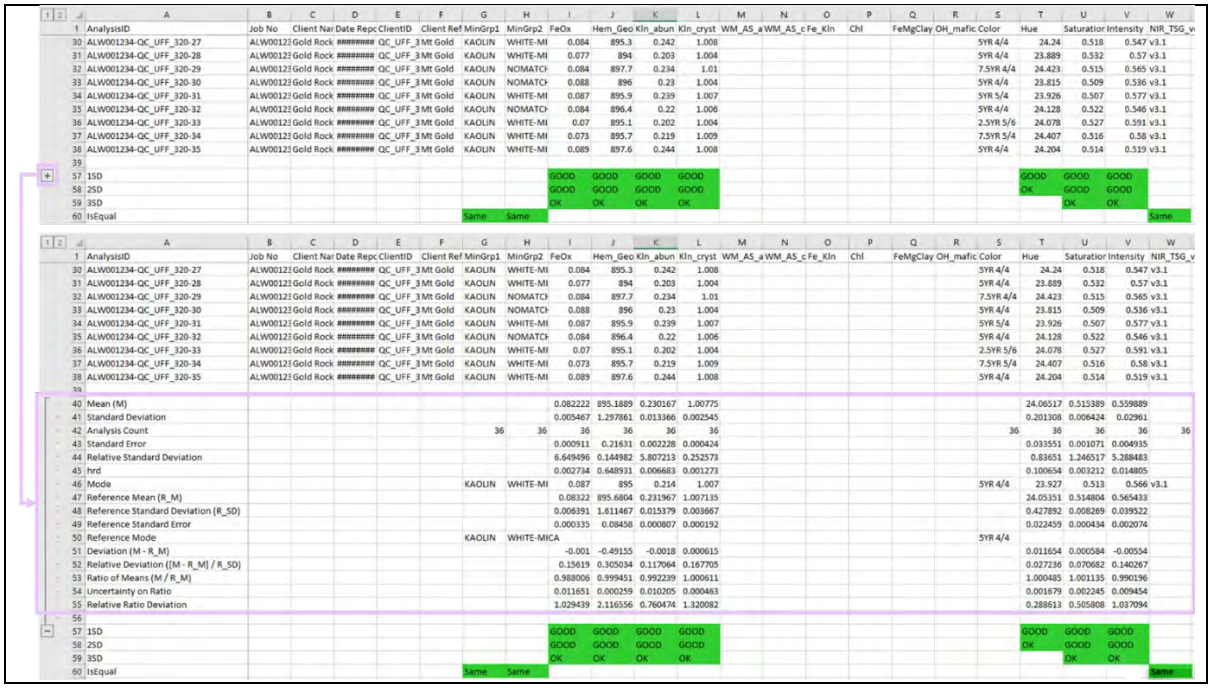
| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AnalysisID | Job No | Client Name | Date Repc | ClientID | Client Ref | Ag_ppm | Al_ppm | As_ppm | Au_ppb | Ba_ppm | Be_ppm | Bi_ppm |
| 3 | ALW001234-QC_UFF_320-0 | ALW001234 | Gold Rocks | ######## | QC_UFF_320 | Mt Gold | 0.024 | 77700 | 12.4 | 9.5 | 71.9 | 1.34 | 0.46 |
| 4 | ALW001234-QC_UFF_320-1 | ALW001234 | Gold Rocks | ######## | QC_UFF_320 | Mt Gold | 0.035 | 106000 | 14.2 | 9.5 | 99 | 1.42 | 0.327 |
| 5 | ALW001234-QC_UFF_320-2 | ALW001234 | Gold Rocks | ######## | QC_UFF_320 | Mt Gold | 0.023 | 74200 | 11.7 | 11.9 | 72.7 | 1.41 | 0.376 |
| 6 | | | | | | | | | | | | | |
| 7 | Mean (M) | | | | | | 0.027333 | 85966.67 | 12.76667 | 10.3 | 81.2 | 1.39 | 0.387667 |
| 8 | Standard Deviation | | | | | | 0.006658 | 17437.41 | 1.289703 | 1.385641 | 15.42044 | 0.043589 | 0.067263 |
| 9 | Analysis Count | | | | | | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 10 | Standard Error | | | | | | 0.003844 | 10067.49 | 0.74461 | 0.8 | 8.902996 | 0.025166 | 0.038834 |
| 11 | Relative Standard Deviation | | | | | | 24.35974 | 20.28392 | 10.10211 | 13.45282 | 18.99069 | 3.135899 | 17.35077 |
| 12 | hrd | | | | | | 0.003329 | 8718.706 | 0.644851 | 0.69282 | 7.71022 | 0.021794 | 0.033632 |
| 13 | Mode | | | | | | 0.023 | 74200 | 11.7 | 9.5 | 71.9 | 1.34 | 0.327 |
| 14 | Reference Mean (R_M) | | | | | | 0.031784 | 97115.43 | 13.45131 | 10.14805 | 91.30467 | 1.591828 | 0.389611 |
| 15 | Reference Standard Deviation (R_SD) | | | | | | 0.00486 | 11796.42 | 1.386701 | 0.818394 | 9.707469 | 0.22799 | 0.021915 |
| 16 | Reference Standard Error | | | | | | 0.000275 | 667.8407 | 0.078506 | 0.046332 | 0.549577 | 0.012907 | 0.001241 |
| 17 | Reference Mode | | | | | | | | | | | | |
| 18 | Deviation (M - R_M) | | | | | | -0.00445 | -11148.8 | -0.68464 | 0.151946 | -10.1047 | -0.20183 | -0.00194 |
| 19 | Relative Deviation ([M - R_M] / R_SD) | | | | | | 0.915819 | 0.945097 | 0.493719 | 0.185663 | 1.040917 | 0.885251 | 0.088717 |
| 20 | Ratio of Means (M / R_M) | | | | | | 0.859972 | 0.885201 | 0.949102 | 1.014973 | 0.88933 | 0.87321 | 0.99501 |
| 21 | Uncertainty on Ratio | | | | | | 0.121176 | 0.103844 | 0.055632 | 0.078969 | 0.097655 | 0.017323 | 0.099725 |
| 22 | Relative Ratio Deviation | | | | | | 1.155572 | 1.105497 | 0.914893 | 0.189605 | 1.133268 | 7.319319 | 0.050039 |
| 23 | | | | | | | | | | | | | |
| 24 | 1SD | | | | | | GOOD | GOOD | GOOD | GOOD | CHECK | GOOD | GOOD |
| 25 | 2SD | | | | | | OK | OK | OK | OK | OK | OK | OK |

**Figure 5**: Example QA/QC output and expanded drop-down menu with more detailed information for geochemical analyses of the UltraFine+® standard (QC_UFF_320).

Compared to the well-established QA/QC procedures for geochemical analyses, routine protocols for QA/QC of features/scalars derived from soil VNIR and FTIR data at the post-processing stage have not been universally established or adopted. Rather, the spectra are visually examined by an expert user and undergo QA/QC during data acquisition and processing in TSG™. However, over four hundred VNIR spectra and over 200 FTIR spectra of the QC 320 reference material were independently reviewed to enable end-user QA/QC on this data.

The automated QA/QC compares the standard analyses in the analyses batches to being within one, two or three standard deviations from the mean of the reference standard analyses for most scalars (Figure 6). This is derived from an assessment of the original VNIR and FTIR spectra and how these spectra relate to the numerical values that are reported. In addition, the *IsEqual* row is used to compare the modal values of categorical properties (non-numerical values) in the VNIR data, and the *RankMetric* row is used for the ordinal/ranking analysis output for Gibbsite reported in the FTIR data (with possible values NULL, 1, 2, 3 indicating increasing unquantified abundances), as Gibbsite should not be present in the QC 320 material.

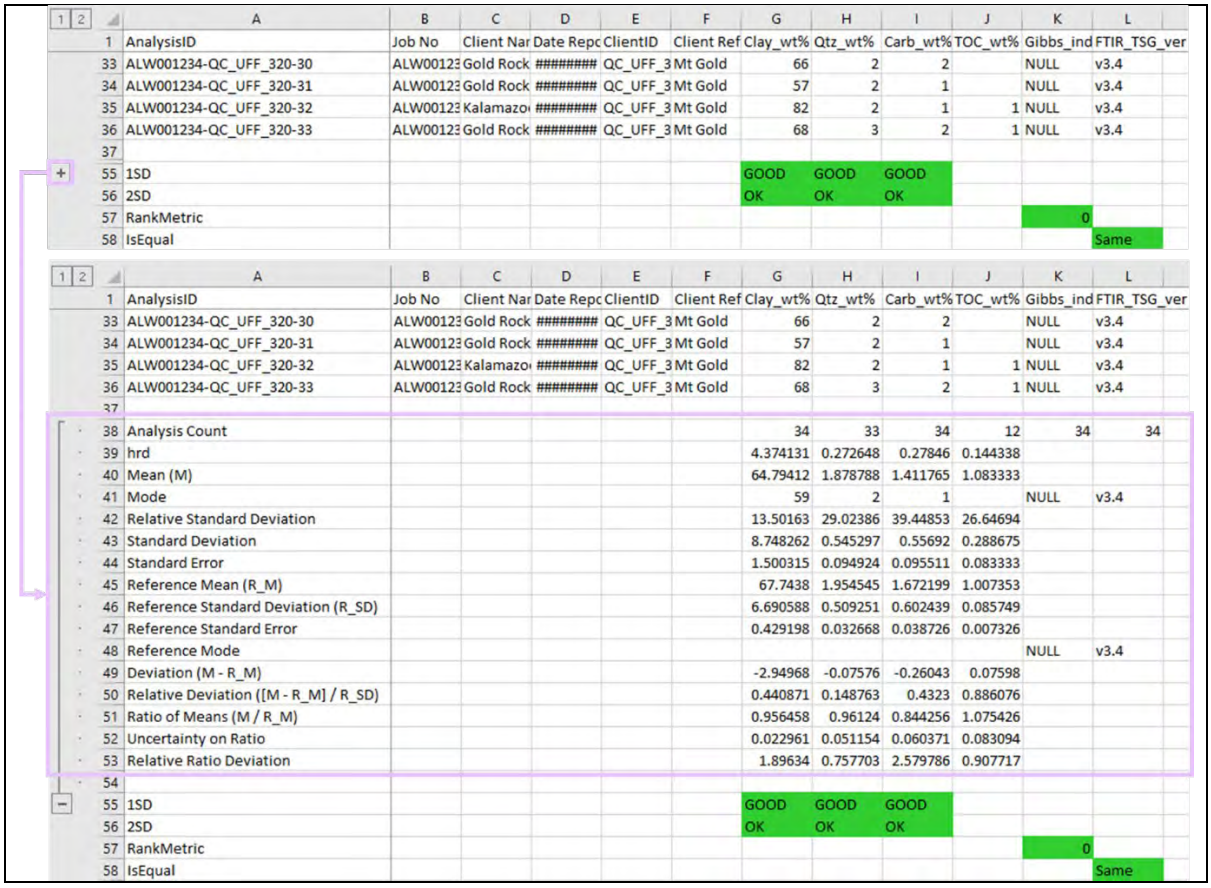
Table **7** shows the parameters and calculations that are presented or used in the QA/QC outputs derived from the M0462a data packages.



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AnalysisID | Job No | Client Nar | Date Repc | ClientID | Client Ref | MinGrp1 | MinGrp2 | FeOx | Hem_Geo | Kln_abun | Kln_cryst | WM_AS_a | WM_AS_c | Fe_Kln | Chl | FeMgClay | OH_mafic | Color | Hue | Saturatior | Intensity | NIR_TSG_v |
| 30 | ALW001234-QC_UFF_320-27 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.084 | 895.3 | 0.242 | 1.008 | | | | | | | 5YR 4/4 | 24.24 | 0.518 | 0.547 | v3.1 |
| 31 | ALW001234-QC_UFF_320-28 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.077 | 894 | 0.203 | 1.004 | | | | | | | 5YR 4/4 | 23.889 | 0.532 | 0.57 | v3.1 |
| 32 | ALW001234-QC_UFF_320-29 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | NOMATCH | 0.084 | 897.7 | 0.234 | 1.01 | | | | | | | 7.5YR 4/4 | 24.423 | 0.515 | 0.565 | v3.1 |
| 33 | ALW001234-QC_UFF_320-30 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | NOMATCH | 0.088 | 896 | 0.23 | 1.004 | | | | | | | 5YR 4/4 | 23.815 | 0.509 | 0.536 | v3.1 |
| 34 | ALW001234-QC_UFF_320-31 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.087 | 895.9 | 0.239 | 1.007 | | | | | | | 5YR 5/4 | 23.926 | 0.507 | 0.577 | v3.1 |
| 35 | ALW001234-QC_UFF_320-32 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | NOMATCH | 0.084 | 896.4 | 0.22 | 1.006 | | | | | | | 5YR 4/4 | 24.128 | 0.522 | 0.546 | v3.1 |
| 36 | ALW001234-QC_UFF_320-33 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.07 | 895.1 | 0.202 | 1.004 | | | | | | | 2.5YR 5/6 | 24.078 | 0.527 | 0.591 | v3.1 |
| 37 | ALW001234-QC_UFF_320-34 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.073 | 895.7 | 0.219 | 1.009 | | | | | | | 7.5YR 5/4 | 24.407 | 0.516 | 0.58 | v3.1 |
| 38 | ALW001234-QC_UFF_320-35 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.089 | 897.6 | 0.244 | 1.008 | | | | | | | 5YR 4/4 | 24.204 | 0.514 | 0.519 | v3.1 |
| 39 | | | | | | | | | | | | | | | | | | | | | | | |
| 57 | 1SD | | | | | | | | GOOD | GOOD | GOOD | GOOD | | | | | | | | GOOD | GOOD | GOOD | |
| 58 | 2SD | | | | | | | | GOOD | GOOD | GOOD | GOOD | | | | | | | | OK | GOOD | GOOD | |
| 59 | 3SD | | | | | | | | OK | OK | OK | OK | | | | | | | | | OK | OK | |
| 60 | IsEqual | | | | | | Same | Same | | | | | | | | | | | | | | | Same |

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AnalysisID | Job No | Client Nar | Date Repc | ClientID | Client Ref | MinGrp1 | MinGrp2 | FeOx | Hem_Geo | Kln_abun | Kln_cryst | WM_AS_a | WM_AS_c | Fe_Kln | Chl | FeMgClay | OH_mafic | Color | Hue | Saturatior | Intensity | NIR_TSG_v |
| 30 | ALW001234-QC_UFF_320-27 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.084 | 895.3 | 0.242 | 1.008 | | | | | | | 5YR 4/4 | 24.24 | 0.518 | 0.547 | v3.1 |
| 31 | ALW001234-QC_UFF_320-28 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.077 | 894 | 0.203 | 1.004 | | | | | | | 5YR 4/4 | 23.889 | 0.532 | 0.57 | v3.1 |
| 32 | ALW001234-QC_UFF_320-29 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | NOMATCH | 0.084 | 897.7 | 0.234 | 1.01 | | | | | | | 7.5YR 4/4 | 24.423 | 0.515 | 0.565 | v3.1 |
| 33 | ALW001234-QC_UFF_320-30 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | NOMATCH | 0.088 | 896 | 0.23 | 1.004 | | | | | | | 5YR 4/4 | 23.815 | 0.509 | 0.536 | v3.1 |
| 34 | ALW001234-QC_UFF_320-31 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.087 | 895.9 | 0.239 | 1.007 | | | | | | | 5YR 5/4 | 23.926 | 0.507 | 0.577 | v3.1 |
| 35 | ALW001234-QC_UFF_320-32 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | NOMATCH | 0.084 | 896.4 | 0.22 | 1.006 | | | | | | | 5YR 4/4 | 24.128 | 0.522 | 0.546 | v3.1 |
| 36 | ALW001234-QC_UFF_320-33 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.07 | 895.1 | 0.202 | 1.004 | | | | | | | 2.5YR 5/6 | 24.078 | 0.527 | 0.591 | v3.1 |
| 37 | ALW001234-QC_UFF_320-34 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.073 | 895.7 | 0.219 | 1.009 | | | | | | | 7.5YR 5/4 | 24.407 | 0.516 | 0.58 | v3.1 |
| 38 | ALW001234-QC_UFF_320-35 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | KAOLIN | WHITE-MI | 0.089 | 897.6 | 0.244 | 1.008 | | | | | | | 5YR 4/4 | 24.204 | 0.514 | 0.519 | v3.1 |
| 39 | | | | | | | | | | | | | | | | | | | | | | | |
| 40 | Mean (M) | | | | | | | | 0.082222 | 895.1889 | 0.230167 | 1.00775 | | | | | | | | 24.06517 | 0.515389 | 0.559889 | |
| 41 | Standard Deviation | | | | | | | | 0.005467 | 1.297861 | 0.013366 | 0.002545 | | | | | | | | 0.201308 | 0.006424 | 0.02961 | |
| 42 | Analysis Count | | | | | | 36 | 36 | 36 | 36 | 36 | 36 | | | | | | | 36 | 36 | 36 | 36 | 36 |
| 43 | Standard Error | | | | | | | | 0.000911 | 0.21631 | 0.002228 | 0.000424 | | | | | | | | 0.033551 | 0.001071 | 0.004935 | |
| 44 | Relative Standard Deviation | | | | | | | | 6.649496 | 0.144982 | 5.807213 | 0.252573 | | | | | | | | 0.83651 | 1.246517 | 5.288483 | |
| 45 | hrd | | | | | | | | 0.002734 | 0.648931 | 0.006683 | 0.001273 | | | | | | | | 0.100654 | 0.003212 | 0.014805 | |
| 46 | Mode | | | | | | KAOLIN | WHITE-MI | 0.087 | 895 | 0.214 | 1.007 | | | | | | | 5YR 4/4 | 23.927 | 0.513 | 0.566 | v3.1 |
| 47 | Reference Mean (R_M) | | | | | | | | 0.08322 | 895.6804 | 0.231967 | 1.007135 | | | | | | | | 24.05351 | 0.514804 | 0.565433 | |
| 48 | Reference Standard Deviation (R_SD) | | | | | | | | 0.006391 | 1.611467 | 0.015379 | 0.003667 | | | | | | | | 0.427892 | 0.008269 | 0.039522 | |
| 49 | Reference Standard Error | | | | | | | | 0.000335 | 0.08458 | 0.000807 | 0.000192 | | | | | | | | 0.022459 | 0.000434 | 0.002074 | |
| 50 | Reference Mode | | | | | | KAOLIN | WHITE-MICA | | | | | | | | | | | 5YR 4/4 | | | | |
| 51 | Deviation (M - R_M) | | | | | | | | -0.001 | -0.49155 | -0.0018 | 0.000615 | | | | | | | | 0.011654 | 0.000584 | -0.00554 | |
| 52 | Relative Deviation ([M - R_M] / R_SD) | | | | | | | | 0.15619 | 0.305034 | 0.117064 | 0.167705 | | | | | | | | 0.027236 | 0.070682 | 0.140267 | |
| 53 | Ratio of Means (M / R_M) | | | | | | | | 0.988006 | 0.999451 | 0.992239 | 1.000611 | | | | | | | | 1.000485 | 1.001135 | 0.990196 | |
| 54 | Uncertainty on Ratio | | | | | | | | 0.011651 | 0.000259 | 0.010205 | 0.000463 | | | | | | | | 0.001679 | 0.002245 | 0.009454 | |
| 55 | Relative Ratio Deviation | | | | | | | | 1.029439 | 2.116556 | 0.760474 | 1.320082 | | | | | | | | 0.288613 | 0.505808 | 1.037094 | |
| 56 | | | | | | | | | | | | | | | | | | | | | | | |
| 57 | 1SD | | | | | | | | GOOD | GOOD | GOOD | GOOD | | | | | | | | GOOD | GOOD | GOOD | |
| 58 | 2SD | | | | | | | | GOOD | GOOD | GOOD | GOOD | | | | | | | | OK | GOOD | GOOD | |
| 59 | 3SD | | | | | | | | OK | OK | OK | OK | | | | | | | | | OK | OK | |
| 60 | IsEqual | | | | | | Same | Same | | | | | | | | | | | | | | | Same |

**Figure 6**: Example QA/QC output and expanded drop-down menu with more detailed information for VNIR analyses of QC_UFF_320 standard analyses. Note that Colour is derived from Hue, Saturation and Intensity and, therefore, rules are applied to those columns but not to the Colour column.



| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AnalysisID | Job No | Client Nar | Date Repc | ClientID | Client Ref | Clay_wt% | Qtz_wt% | Carb_wt% | TOC_wt% | Gibbs_ind | FTIR_TSG_ver |
| 33 | ALW001234-QC_UFF_320-30 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | 66 | 2 | 2 | | NULL | v3.4 |
| 34 | ALW001234-QC_UFF_320-31 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | 57 | 2 | 1 | | NULL | v3.4 |
| 35 | ALW001234-QC_UFF_320-32 | ALW00123 | Kalamazo | ######## | QC_UFF_3 | Mt Gold | 82 | 2 | 1 | 1 | NULL | v3.4 |
| 36 | ALW001234-QC_UFF_320-33 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | 68 | 3 | 2 | 1 | NULL | v3.4 |
| 37 | | | | | | | | | | | | |
| 55 | 1SD | | | | | | GOOD | GOOD | GOOD | | | |
| 56 | 2SD | | | | | | OK | OK | OK | | | |
| 57 | RankMetric | | | | | | | | | | 0 | |
| 58 | IsEqual | | | | | | | | | | | Same |

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AnalysisID | Job No | Client Nar | Date Repc | ClientID | Client Ref | Clay_wt% | Qtz_wt% | Carb_wt% | TOC_wt% | Gibbs_ind | FTIR_TSG_ver |
| 33 | ALW001234-QC_UFF_320-30 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | 66 | 2 | 2 | | NULL | v3.4 |
| 34 | ALW001234-QC_UFF_320-31 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | 57 | 2 | 1 | | NULL | v3.4 |
| 35 | ALW001234-QC_UFF_320-32 | ALW00123 | Kalamazo | ######## | QC_UFF_3 | Mt Gold | 82 | 2 | 1 | 1 | NULL | v3.4 |
| 36 | ALW001234-QC_UFF_320-33 | ALW00123 | Gold Rock | ######## | QC_UFF_3 | Mt Gold | 68 | 3 | 2 | 1 | NULL | v3.4 |
| 37 | | | | | | | | | | | | |
| 38 | Analysis Count | | | | | | 34 | 33 | 34 | 12 | 34 | 34 |
| 39 | hrd | | | | | | 4.374131 | 0.272648 | 0.27846 | 0.144338 | | |
| 40 | Mean (M) | | | | | | 64.79412 | 1.878788 | 1.411765 | 1.083333 | | |
| 41 | Mode | | | | | | 59 | 2 | 1 | | NULL | v3.4 |
| 42 | Relative Standard Deviation | | | | | | 13.50163 | 29.02386 | 39.44853 | 26.64694 | | |
| 43 | Standard Deviation | | | | | | 8.748262 | 0.545297 | 0.55692 | 0.288675 | | |
| 44 | Standard Error | | | | | | 1.500315 | 0.094924 | 0.095511 | 0.083333 | | |
| 45 | Reference Mean (R_M) | | | | | | 67.7438 | 1.954545 | 1.672199 | 1.007353 | | |
| 46 | Reference Standard Deviation (R_SD) | | | | | | 6.690588 | 0.509251 | 0.602439 | 0.085749 | | |
| 47 | Reference Standard Error | | | | | | 0.429198 | 0.032668 | 0.038726 | 0.007326 | | |
| 48 | Reference Mode | | | | | | | | | | NULL | v3.4 |
| 49 | Deviation (M - R_M) | | | | | | -2.94968 | -0.07576 | -0.26043 | 0.07598 | | |
| 50 | Relative Deviation ([M - R_M] / R_SD) | | | | | | 0.440871 | 0.148763 | 0.4323 | 0.886076 | | |
| 51 | Ratio of Means (M / R_M) | | | | | | 0.956458 | 0.96124 | 0.844256 | 1.075426 | | |
| 52 | Uncertainty on Ratio | | | | | | 0.022961 | 0.051154 | 0.060371 | 0.083094 | | |
| 53 | Relative Ratio Deviation | | | | | | 1.89634 | 0.757703 | 2.579786 | 0.907717 | | |
| 54 | | | | | | | | | | | | |
| 55 | 1SD | | | | | | GOOD | GOOD | GOOD | | | |
| 56 | 2SD | | | | | | OK | OK | OK | | | |
| 57 | RankMetric | | | | | | | | | | 0 | |
| 58 | IsEqual | | | | | | | | | | | Same |

**Figure 7:** Example QA/QC output for FTIR analyses of QC_UFF_320 standard analyses with drop-down menu for more information.

Like the geochemical analyses, if everything is green, there is no cause for concern. If some scalars are flagged as red more detailed information can be assessed by clicking on the + symbol to open a drop-down of calculations (refer to the

Table **7** for an explanation of the calculated parameters). Red flags for individual scalars can indicate that this scalar may not be reliable. These values or anomalous results should be used with caution and checked against other parameters. If there are consistent red flags for multiple scalars for one, two and three standard deviations, it is recommended to identify whether this is due to a specific batch analysis. If consistent red flags across multiple batches occur for multiple scalars, following up with the laboratory for clarification on the precision of the analyses is recommended.

**Table 7**: Parameters, calculations and details used in the QA/QC data analytics for UltraFine+ Next Gen Analytics.

| | Formulae | |
|---|---|---|
| **Analysis Count** | $n$ | The total count of all available QC_UFF_320 analyses for the specific analyte |
| **Mean (M)** | $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ | The mean (or average) of all available QC_UFF_320 analyses |
| **Standard Deviation** | $s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})}$ | Measures how dispersed the data is in relation to the mean; low values indicate the analyses are clustered around the mean (are more precise); high values indicate the analyses exhibit greater variation |
| **Standard Error** | $s_{\overline{x}} = \frac{s}{\sqrt{n}}$ | The standard deviation of the sample population which measures the accuracy with which a sample represents a population |
| **Relative Standard Deviation** | $100\frac{s}{\overline{x}}$ | Percentage value describing the spread of analyses as a proportion of the Mean |
| **hrd** | $100\frac{\lvert x_a - x_b\rvert}{x_a + x_b}$ | Half absolute relative difference between two estimates of a value |
| **Mode** | | The most common value for a specific analyte; typically only used for categorical/ordinal data |
| **Reference Mean (R_M)** | $\mu_{ref} = \frac{1}{N}\sum_{i=1}^{N} x_i$ | The mean of the QC_UFF_320 standard calculated from a static set of 312 analyses for 49 elements. |
| **Reference Standard Deviation (R_SD)** | $\sigma_{ref} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})}$ | The standard deviation of the QC_UFF_320 standard calculated from a static set of 312 analyses for 49 elements. |
| **Reference Standard Error** | $\sigma_{\mu_{ref}} = \frac{s\sigma_{ref}}{\sqrt{n_{ref}}}$ | The standard deviation of the reference analysis population (the 312 QC_UFF_320 analyses) |
| **Reference Mode** | | The most common value or a specific analyte in the reference data set (the 312 QC_UFF_320 analyses); typically only used for categorical/ordinal data. |
| **Deviation (M - R_M)** | $D = \mu - \mu_{ref}$ | The difference between the mean of the reference material analyses for this site (M) relative to the mean of a static set of 312 analyses for 49 elements (R_M) |
| **Relative Deviation ([M - R_M] / R_SD)** | $\frac{D}{\sigma_{ref}}$ | The relative difference between the means (M, R_M), expressed as the Deviation (above, M-R_M) scaled by the standard deviation of the reference set (R_SD) to give an estimate of 'standard deviations from unity' |
| **Ratio of Means (M / R_M)** | $R = \frac{\mu}{\mu_{ref}}$ | The ratio between the mean of the reference material analyses for this site (M) relative to the mean of a static set of 312 analyses for 49 elements (R_M) |
| **Uncertainty on Ratio** | $\sigma_R = R\sqrt{\frac{{\sigma_{\mu_{ref}}}^2}{\mu_{ref}} + \frac{s_{\overline{x}}^2}{\overline{x}}}$ | The uncertainty on the ratio of means (M/R_M) estimated by propagated uncertainties (1σ) |
| **Relative Ratio Deviation** | $\frac{R-1}{\sigma_R}$ | An estimate of how far the means are from equality (a ratio of one, or unity) expressed in terms of the uncertainty on the ratio (σ); values greater than two likely indicate significant differences between the means |
| **IsEqual** | | Whether two categorical data values are the same (e.g., mineral groups, software versions) |
| **RankMetric** | | Rank differences between two ordinal values, used here to compare the modal Gibbsite index (with ordinal values [NULL, 1, 2, 3]) |
| **1SD** | $\left\lvert\frac{D}{s_{ref}}\right\rvert < 1$ | Indicates whether the Mean of the QC_UFF_320 analyses in this batch is within one (1) standard deviation of the Mean of the reference standard values (calculated from 312 QC_UFF_320 analyses) |
| **2SD** | $\left\lvert\frac{D}{s_{ref}}\right\rvert < 2$ | Indicates whether the Mean of the QC_UFF_320 analyses in this batch is within two (2) standard deviations of the Mean of the reference standard values (calculated from 312 QC_UFF_320 analyses) |
| **3SD** | $\left\lvert\frac{D}{s_{ref}}\right\rvert < 3$ | Indicates whether the Mean of the QC_UFF_320 analyses in this batch is within three (3) standard deviations of the Mean of the reference standard values (calculated from 312 QC_UFF_320 analyses) |

The QC 320 standard is not analysed for pH and EC (or sizing) and some of the scalars for the spectral data were also developed later in the project and are not included. This was also the case for the elements Br, I and Pd that were not part of the original UltraFine+® analysis method. While the QC 320 is not a certified CRM, means, standard deviation and tolerance and confidence intervals for this standard are available online[6] to assist industry in using the

[6] https://research.csiro.au/ultrafine/resources/other/

UltraFine+® method. It is important to note that FTIR data is only semi-quantitative and does not include tolerance intervals for this data. TSG™ version processing for spectral analyses on the QC 320 material is available for NIR TSG™ v. 3.1 and onwards, and for FTIR TSG™ v. 3.3 and onwards.

#### *2.3.2.3 Batch Uncertainty*

In addition to the evaluation of precision and accuracy of individual sample field duplicates and standards, Python coding was used to provide an overview of the repeatability of sampling and the analytical process via each batch as uncertainty maps. The statistical QA/QC metrics developed for analytical quality assessment of duplicates was applied to the duplicates of each batch/laboratory job (samples analysed in one batch, based on laboratory job ID). The averages of batch duplicate metrics were plotted on maps for each analytical parameter (Figure 8). The landscape clustering maps (see next section) were used as a base map. Batch uncertainty values of >20 % were coloured black and values between 0 % and 20 % were grey-scale, where the level of uncertainty increases with increasing grey tone. Currently, the Python code generates batch uncertainty maps for Au, Cu, Zn, Pb and Ni.

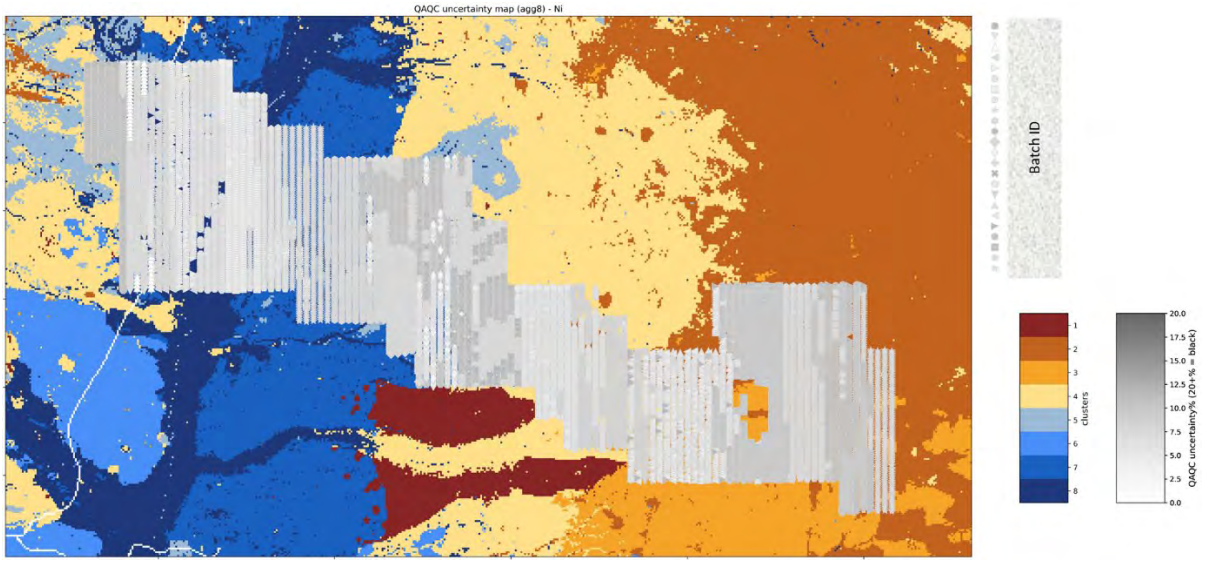



Figure 8: Example of a batch uncertainty map for Ni analyses over 24 different batches in a de-identified project area. Note that the batch uncertainty was based on duplicate samples to assess the combined precision of field sampling and laboratory analyses.

### 2.3.3 Spatial Data Integration and Clustering

#### *2.3.3.1 Spatial Data Clustering*

The UltraFine+® Next Gen Analytics workflow uses unsupervised machine learning (ML) methods to cluster publicly available spatial data to generate proxy regolith type maps (Figure 2). The spatial data feature layers used in this workflow were selected for their inferred relationship to regolith landforms and provide information on landscape position, depth of transported cover, regolith mineralogy and geological provenance. The specific features used are minimally processed or interpreted, minimising the introduction of human subjectivity associated with interpreted features such as surface geology or existing regolith landscape type maps. Spatial data feature layers included in the workflow are a Digital Elevation Model (the Copernicus GLO-30 Digital Elevation Model), Multi-resolution Valley Bottom Flatness

(MrVBF), Radiometric K (%), Th (ppm), U (ppm) and barest Earth Sentinel-2 satellite data (Table 8).

The Sentinel-2 satellites, launched via the European Union Copernicus Program's Earth observation mission, collect high-resolution multispectral imagery data of the Earth with a revisit period of 10 days[7]. The Sentinel-2 data used in this workflow is taken from the Sentinel 2A Barest Earth Analysis Ready Data, cropped to an area based on the soil sample locations. Ten out of thirteen spectral bands are collected: B02, B03, B04, B05, B06, B07, B08, B8A, B11, and B12. A three-component band ratio image is generated from the Sentinel-2 imagery (approximating the equivalent Landsat-7 ETM+ arrangement of Gozzard 2005; see USGS comparison of band equivalence[8]), with respective RGB bands calculated using R= B11/B12, G=B08/B12 and B=B08/B03.

**Table 8**: Spatial data layers used for landscape clustering of each project area.

| LAYER | LANDSCAPE INFORMATION | DATA SOURCE |
|---|---|---|
| **DEM** | Landscape Position | Copernicus GLO-30 Digital Elevation Model was accessed on 11/05/2022 from https://registry.opendata.aws/copernicus-dem. |
| **MrVBF** | Depth of transported cover | Gallant, J., Dowling, T., Austin, J., 2012. Multi-resolution Valley Bottom Flatness (MrVBF). v3. CSIRO. Data Collection. https://doi.org/10.4225/08/5701C885AB4FE |
| **Radiometrics K pct** | Parent material | Poudjom Djomani, Y., 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 |
| **Radiometrics Th ppm** | Parent material | Poudjom Djomani, Y., 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 |
| **Radiometrics U ppm** | Parent material | Poudjom Djomani, Y., 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 |
| **Sentinel-2** | Surface material and dispersion | **Digital Earth Australia** https://explorer.sandbox.dea.ga.gov.au/products/s2_barest_earth |

Prior to dimensionality reduction and clustering, all data layers were re-gridded to the highest input layer resolution (typically on the order of 20 m) and scaled according to zero median and unit inter-quartile range. The dimensionality reduction algorithm UMAP (Uniform Manifold Projection and Approximation; McInnes et al. 2018) was applied to embed the spatial data pixels (or, grid-cells) into a three-dimensional latent space (Figure 9) for more efficient clustering and to provide a framework for visualisation. The UMAP algorithm was used as it captures non-linearities within the data and balances capturing the global dataset structure with preserving local structure. The method does not explicitly include any location information, spatial relationships, or spatial features (e.g., textures) as only the per-pixel values of each input layer were considered.

[7] See Digital Earth Australia data product documentation at https://knowledge.dea.ga.gov.au/notebooks/DEA_datasets/DEA_Sentinel2_Surface_Reflectance/
[8] https://www.usgs.gov/media/images/comparison-landsat-7-and-8-bands-sentinel-2

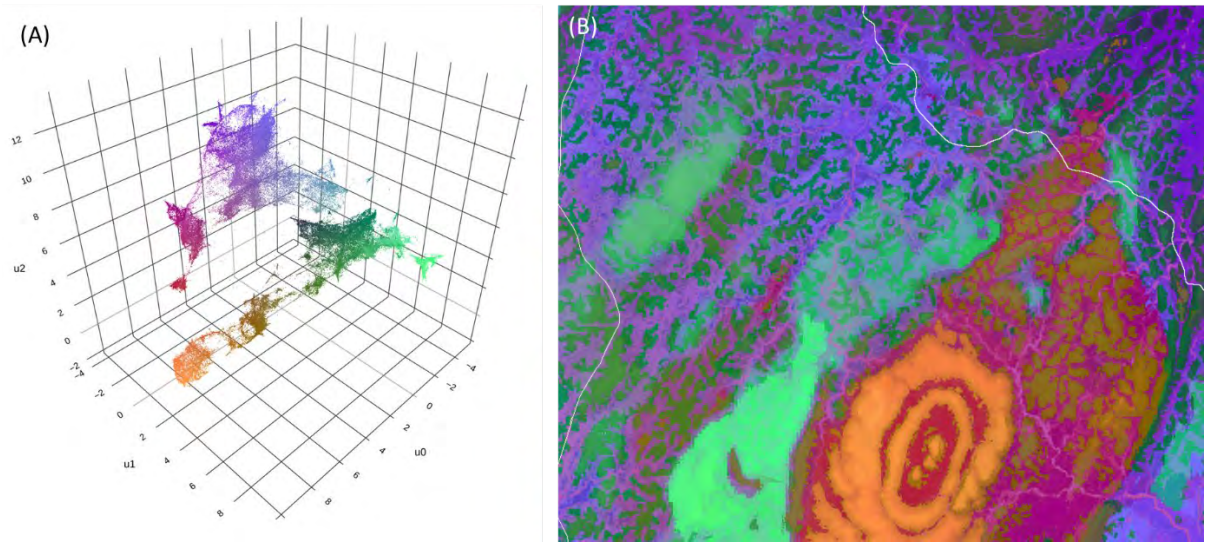



Figure 9: (A) Scatter diagram of pixel values for one of the project sites embedded in a 3-dimensional latent space using the dimensionality reduction algorithm UMAP. Points are coloured using an RGB of the axes (u0, u1, u2) values. (B) Pixel values projected into 2-dimensional space for spatial context over the de-identified, approximately 290 km$^2$ large project area. Note that roads (white lines) have been masked out.

The Next Gen Analytics workflow uses the *k* -means and agglomerative hierarchical clustering algorithms (scikit-learn implementation in Python; Pedregosa, 2011) to cluster UMAP-embedded pixels with similar features. Both clustering techniques require a-priori selection of a specified number of clusters and assign each pixel to one of these clusters. Due to the intended application of the clustered proxy landscape types to geochemical sample interpretation and the added complication that these landscape models are intended to be used over a variety of scales (1 km$^2$ to 50,000 km$^2$), the number of clusters used in the workflow has been standardised for practical purposes. Four clusters are used for *k*-means (for project areas <20 km$^2$), eight clusters are used for agglomerative hierarchical clustering (typically for project areas >20 km$^2$), and 12 clusters are used for agglomerative hierarchical clustering for large project areas with dense sampling. Herein, they are referred to as 'kmeans4', 'agg8' and 'agg12'.

*k* -means is an iterative clustering algorithm that initialises randomly generated cluster centres, and hence involves a degree of random variability (MacQueen, 1967) which in the case of the Next Gen Analytics workflows is explicitly controlled.

Agglomerative hierarchical clustering is a hierarchical clustering algorithm that recursively merges pairs of clusters, minimising a specified distance metric (and is in this sense, deterministic). Within the Next Gen Analytics workflows, the scikit-learn default linkage criterion is used (Ward distance), which minimises the variance within newly merged clusters. Training an agglomerative clustering model is computationally intensive and, once trained, a model is not able to assign new samples to established clusters. To assign cluster labels to pixels at a larger scale, a multi-step process was used:

a) fit an agglomerative clustering model using a random subset of 20,000 pixels;
b) use this same subset and their cluster labels (from step 1) to train a random forest classifier; and
c) use this classifier to predict cluster labels for all pixels across the area of interest.

The random forest classifier has been observed to be an effective proxy for the agglomerative clustering model, with out-of-bag (OOB) prediction accuracy typically above 99%. For areas

smaller than 20,000 pixels (such as the Jericho Project area in Queensland; Henne et al. 2023a), the agglomerative clustering model is instead used directly, however, for larger areas, the process described above was used.

For all clustering models, a colour is assigned to each cluster based on rankings of mean cluster MrVBF (proxy for depth of transported cover) values. These clusters can be visualised spatially to generate maps which are used to evaluate the proxy regolith type that each cluster corresponds to (Figure 10). These maps are exported in multiple formats, including GeoTIFF and PNG (see Appendix A; see Section 1.6).

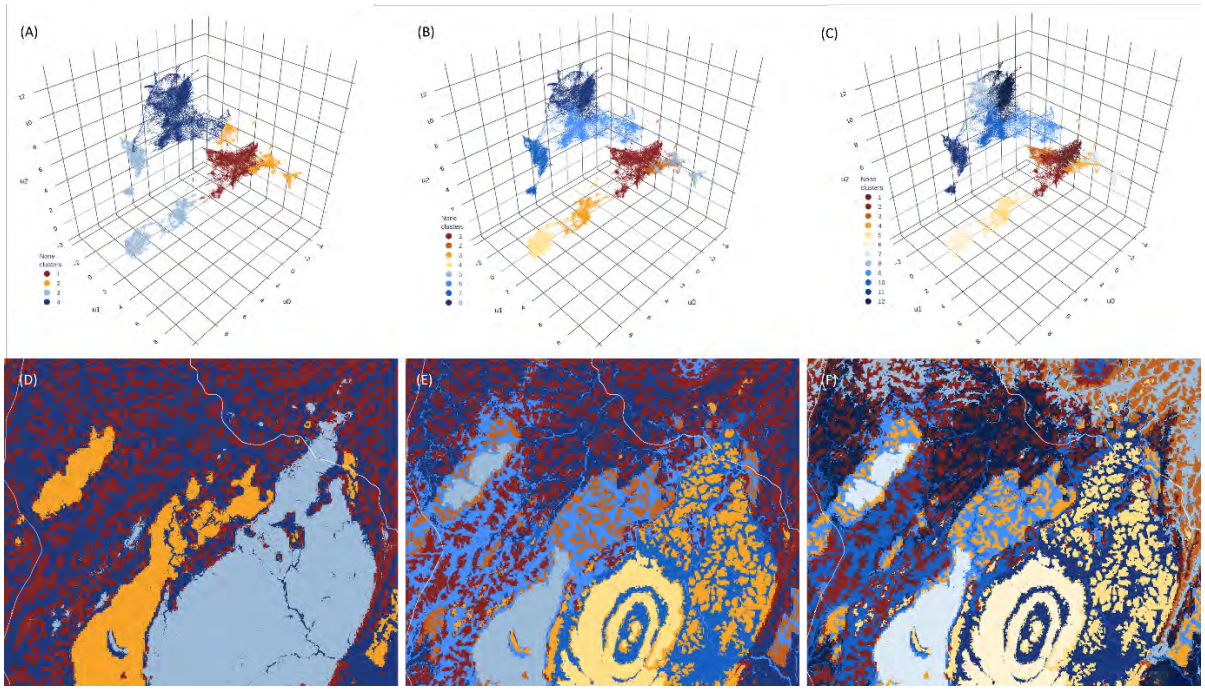



Figure 10: (A – C) Pixels of a large project site (approximately 290 km$^2$) clustered into 4, 8 and 12 clusters in the 3-dimensional latent space displayed by cluster colour to visualise data separation. (D-F) Pixel values grouped by cluster colour into 4, 8 and 12 "landscape" clusters projected into 2-dimensional space for spatial context over the de-identified project area. Note that roads (white lines) have been masked out.

### 2.3.4 Geochemical Outliers by Landscape Type

Outliers for the 52 elements reported by the UltraFine+® method were calculated on log-transformed data and are defined as values that are greater than 1.5 times the interquartile range (IQR) beyond the first or third quantile. The main goal of deriving machine-learned landscape maps was to provide context for geochemical soil surveys. Hence, geochemical data were also grouped by their corresponding landscape cluster into 4, 8 and 12 groups respectively, and outliers were calculated for each of these clusters. The workflow produces boxplots for each element by landscape type (i.e., cluster), as well as in spatial context plotted over a landscape cluster map for each element (see section 1.6.5). In addition, shapefiles with outliers grouped by element and landscape type are produced. The data is also available as a CSV file (see Appendix A).

### 2.3.5 Principal Component Analysis

Principal Component Analysis (PCA) is performed on quantile-normalised centred-log-ratio transformed geochemical data for each soil sample. A threshold of 10 % is applied for missing data points (due to non-analysis) to determine if an element is included in the PCA analysis.

If less than 10 % of data points of a given element are missing, the element is included in the PCA analysis, but no principal components are calculated for samples with missing data. The workflow reduces each data point (soil sample) from *n* dimensions (number of analysed elements) into five principal components (PC0 to PC4). The explained proportion of variance of the principal components decreases from PC0 to PC4. For each principal component, the loading of each element is visualised on a spider diagram to illustrate the general geochemical affinity (see 1.7.2). The weights of each principal component for each sample are also visualised spatially (see Section 1.7.2).

### 2.3.6 Regolith ratios and indices

Multi-element indices and regolith ratios based on common examples used in Australian exploration settings were calculated from soil sample geochemical data (see Appendix A). These include indices for the purposes of normalisation (e.g., Au/OrgC) as well as indicators and proxies of mineralisation (e.g., CHI6). Multi-element compositional indices are derived from centred-log-ratio (CLR) transformed geochemical data for each soil sample (addressing proportional dependence related to closure). Regolith ratios were calculated from measured elemental abundances (without further transformation), sizing data and/or spectral analyses of ultrafine soil samples.

### 2.3.7 Other Soil Properties

#### *2.3.7.1 Soil Sizing*

Particle size analysis was conducted on field samples (<2 mm; dried soil) with the results reported in percentage of sand, silt, and clay. These values are plotted in a ternary plot commonly used in soil science to provide a general textural class (Soil Science Division Staff 2017). Plots were generated using pyrolite (Williams et al. 2020) and symbolised by colour according to broad soil textural classes (e.g., sandy loam, silty clay). These classes are presented on the soil textural triangle and in spatial context allowing users to see key changes in the landscape soil morphology (see Appendix A).

### 2.3.8 Dispersion and source direction

Dispersion/source direction and topographic slope are extracted at each point from Australia-wide slope (Gallant and Austin 2012a) and aspect (Gallant and Austin 2012b) datasets, which were derived from a smoothed version of the Shuttle Radar Topography Mission Digital (SRTM) Elevation Model (DEM) processed to reduce noise. Dispersion direction indicates the direction in which the land surface slope faces (i.e., the aspect) and is generated on a grid over a given project area. The grid spacing for the dispersion direction is automatically adjusted based on the survey area size. Source direction is the opposite of this (i.e., upslope direction) and is generated for each sample point. Outputs are available as shapefiles with associated layer files that display both source and dispersion direction as scaled arrows (see Appendix A). The size of the arrow is proportional to the slope degree.

### 2.3.9 Data Package and html-viewer

The semi-automated Next Gen Analytics workflow generates a data package that contains all the outputs described above. These include:

1) The UltraFine+® analytical results (geochemistry, pH, EC, VNIR, FTIR and sizing) integrated with sample locations and transformed (CLR) data (csv and shapefiles)

2) “Traffic Light” QA/QC for all QC 320 standards and duplicates analysed (csv files)
3) Landscape models with 4, 8 and 12 clusters (TIFF, PNG and HTML files; including UMAP outputs) and identification of landscape type for each soil sample (csv and shapefiles)
4) Identification of geochemical outliers by landscape type for 52 elements (PNGs of boxplots and maps as well as csv and shapefiles for all three landscape models)
5) Identification of outliers from the whole data population for 52 elements (PNGs of boxplots and maps as well as csv and shapefiles)
6) Principal component analysis of 52 elements (PNGs, csv and shapefiles)
7) Regolith ratios and exploration indices (PNG, CSV and shapefiles)
8) Visualisation of other soil properties (maps and shapefiles of VNIR, FTIR, pH, EC, sizing) and first-pass interpretation of data (soil texture as PNG and shapefiles; pH and EC boxplots)
9) Source direction of each soil sample and dispersion direction within the survey area (shape files)
10) The DSO (Digital Sample Observer), a HTML-based web viewer

The Digital Sample Observer is an interactive HTML-based web viewer that shows the respective landscape model and corresponding UMAP (3-dimensional representation of the spatial attributes of the project area) as well as the input layers (MrVBF, radiometrics, DEM and regolith ratios) for each model area along with a true colour satellite image. The available surface geology and regolith geology maps are also included for comparison. Drop down tabs on the right side of the DSO contain individual sample results that can be projected onto the landscape map. For geochemistry the options are to present all values or outliers only. Spectral results, PCA, indices and other soil properties are presented as the whole data (no outlier options).

As part of the project deliverables for each project site, a *readme.txt* file within the Data Package that is provided contains a recommendation on which model (kmeans4, agg8 or agg12) is the most appropriate for the project site soil survey and other important considerations specific to that project site/data package.

# 3 Results and Discussion

Much of the detailed project site specific findings were reported individually to industry sponsors and these do not form part of the final report due to industry confidentiality. However, analyses of each site informed the method developments and machine learning workflow improvements and enabled the development of the UltraFine+® Next Gen Analytics workflow and data package. A large volume of information was made publicly available on the UltraFine+® website[9] to enable timely and applied knowledge transfer, and much of this information is reflected here. This report documents the methods and the key outputs that form the UltraFine+® Next Gen Analytics data package, and key information to support usability of these outputs is documented below.

## 3.1 Analytical Refinements of the UltraFine+® Method (Milestones 1 and 2)

As part of this research, the UltraFine+® soil analytical method, developed during the prior M0462 project, was extended and refined to achieve key improvements. These include the successful addition of platinum group elements, lowering of detection limits for Au and other key pathfinder elements, and the addition of mineralogy proxies from spectral data. The robustness of the approach was tested via reference soil samples as well as through comparison studies to other commercially available methods. The extended UltraFine+® soil analytical method (including soil geochemistry, sizing, pH, EC and spectral mineralogy proxies) is now commercially available.

### 3.1.1 Reproducibility of ultrafine Reference Material Analyses (QC 320 and QC 420)

Given the novelty of the UltraFine+® method and the lack of appropriate reference materials, two reference materials were specifically designed for the analysis of the clay fraction. Reference material QC 320 was manufactured by the CSIRO based on blending the ultrafine fraction of soils from 10 locations in Western Australia that represent a variety of typical exploration settings and targets (documented in Noble et al. 2018) and the more recent reference material QC 422 was manufactured by LabWest. While QC 320 is not a certified CRM, means, standard deviation and tolerance and confidence intervals for this standard are available online[10].

A total of 420 batches of ultrafine reference materials were subjected to analysis in order to assess the reproducibility of these reference materials and to identify any long-term variations linked to changes in the analytical procedure, which encompassed enhancements to detection limits, and the inherent heterogeneity of the reference materials. During the timeframe relevant to this dataset, a comprehensive total of 1563 analyses was conducted on QC 320, and an additional 400 analyses were performed on QC 422, which is intended to replace QC 320 once its supply is depleted. The means and relative uncertainties were systematically quantified and compared. Discernible trends, apparent at the level of individual batches, were also observed across the complete set of measurements pertaining to the reference materials. Figure 11 offers additional insights, including an examination of the variability of halogens, in conjunction with certain elements linked to resistant phases, such as Hf, Zr, and the higher

[9] https://research.csiro.au/ultrafine/
[10] https://research.csiro.au/ultrafine/resources/other/

relative uncertainties associated with rare and noble metals, as expected due to their lower concentrations. The reference material analyses were additionally scrutinised for any discernible drift relative to the sequential batch identifiers. Over the defined time period, a degree of drift was identified for several elements. This phenomenon is visually illustrated for a subset of elements in Figure 12, wherein least square regression analyses were employed to quantify the magnitude of drift. With regards to QC 320, 26 out of the 65 measured elements exhibited statistically significant drift exceeding 10 % over this period, while QC 422 exhibited significant drift exceeding 10 % for 21 out of the 65 measured elements (it should be noted that this corresponds to a smaller number of analyses conducted over a shorter duration). Overall, this will influence the QC assessment slightly, making the assessment indicate the precision slightly poorer than it is, but it does not affect the exploration sample results or interpretation of the data.

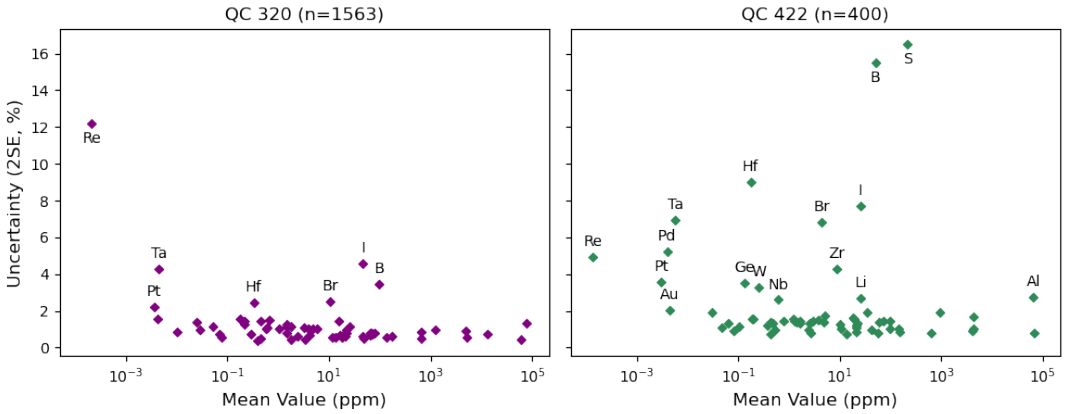



Figure 11: Comparison of mean values and uncertainties for elements across QC 320 (left) and QC 422 (right), where elements with higher than 2% relative uncertainty in the mean value are annotated. Note that the dataset of QC 422 analyses is significantly smaller but shows a higher number of elements with higher-than-expected uncertainty in the mean for a given abundance (particularly for B, S). Some elements display relatively high uncertainty across both reference materials, including the halogens (I, Br), elements associated with resistate phases (e.g., Hf, Zr, Ta); rare and noble metals typically exhibit relatively high uncertainties, as expected from their low overall abundances and proximities to detection limits. Reproduced from Henne et al. 2023b.

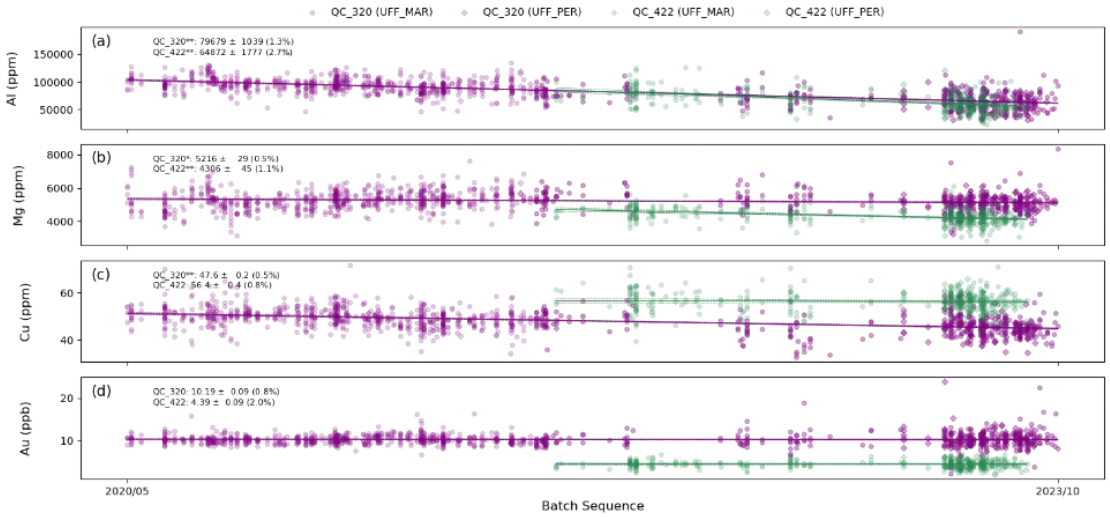



Figure 12: Linear regression of QC 320 and QC 422 analyses over the time period relevant to this dataset for a subset of elements (Al, Mg, Cu, Au). Regression has been performed against sequential batch identifiers, as dates for batches were not consistently available; no outliers were excluded from the regression, and both analytical schemes (UFF MAR, UFF PER) were included in each regression. Confidence intervals are shown either side of the regression with dashed lines (95% confidence on estimate of the mean value; more clearly visible for QC 422).

Means and uncertainties (± 2SE) are given for each reference material, with an asterisk used to denote statistically significant drift (where the regression coefficient is beyond 2SE from zero), and two asterisks used to denote both significant drift and differences in the regression of the mean between the first and last batches of greater than five percent. Reproduced from Henne et al. 2023b.

### 3.1.2 Detection limit improvements and additional soil geochemical properties

A total of four elements, Pt, Pd, Br and I, were added to the standard suite of analytes. In addition, the analytical detection limits of numerous key pathfinder elements such as W, Sn and REE that are resistant to acids were improved (33 elements; Table 9). Reference materials specifically developed and characterised for the UltraFine+® method (QC320 and QC422) were used in repeatability studies to determine the robustness of the new proxies and new chemical extraction requirements (Figure 11).

The UltraFine+® method was not developed for measuring REE. However, it has been used by a number of companies and an option for the full REE suite was recently added by LabWest in response to industry demand. This addition and testing was outside of the project scope, and only initial test work for sample sites was completed, with a full REE suite available for approximately 100 repeats for QC 422. This has provided initial validation around the analytical method and the ability to analyse a wider range of elements than initially proposed. This is discussed in the following section comparing UltraFine+® to aqua regia and 4-acid extractions.

Table 9: Comparison of detection limits (DL) before commencement of the research project and upon completion of the research project (blue font). Full REE not included.

| Element | Unit | DL April 2020 | DL July 2022 |
|---|---|---|---|
| **Ag** | ppm | **0.01** | 0.003 |
| **Al** | ppm | 10 | 10 |
| **As** | ppm | 0.5 | 0.5 |
| **Au** | ppb | 0.5 | 0.5 |
| **Ba** | ppm | 0.2 | 0.2 |
| **Be** | ppm | **0.2** | **0.01** |
| **Bi** | ppm | **0.1** | 0.002 |
| **Br** | ppm | **N/A** | 1 |
| **Ca** | ppm | 10 | 10 |
| **Cd** | ppm | **0.01** | 0.004 |
| **Ce** | ppm | 0.05 | 0.05 |
| **Co** | ppm | **0.2** | 0.01 |
| **Cr** | ppm | 2 | 2 |
| **Cs** | ppm | **0.1** | 0.03 |
| **Cu** | ppm | 0.2 | 0.1 |
| **Fe** | ppm | 100 | 50 |
| **Ga** | ppm | 0.05 | 0.05 |
| **Ge** | ppm | 0.05 | 0.05 |
| **Hf** | ppm | 0.02 | 0.002 |

| Element | Unit | DL April 2020 | DL July 2022 |
|---|---|---|---|
| **Hg** | ppm | **0.01** | 0.001 |
| **I** | ppm | **N/A** | 1 |
| **In** | ppm | **0.01** | 0.001 |
| **K** | ppm | 10 | 10 |
| **La** | ppm | 0.05 | 0.05 |
| **Li** | ppm | 0.5 | **0.05** |
| **Mg** | ppm | 10 | 10 |
| **Mn** | ppm | **2** | 0.5 |
| **Mo** | ppm | **0.1** | 0.03 |
| **Nb** | ppm | 0.01 | 0.01 |
| **Ni** | ppm | **2** | 0.2 |
| **Pb** | ppm | **0.2** | 0.05 |
| **Pd** | ppb | N/A | 1 |
| **Pt** | ppb | **N/A** | 1 |
| **Rb** | ppm | 0.1 | 0.1 |
| **Re** | ppm | 0.0001 | 0.0001 |
| **S** | ppm | **50** | 5 |
| **Sb** | ppm | 0.1 | **0.001** |
| **Sc** | ppm | **1** | 0.2 |
| **Se** | ppm | 0.05 | 0.05 |
| **Sn** | ppm | **0.2** | 0.02 |
| **Sr** | ppm | 0.1 | 0.1 |
| **Ta** | ppm | **0.005** | 0.001 |
| **Te** | ppm | 0.001 | 0.001 |
| **Th** | ppm | 0.02 | 0.02 |
| **Ti** | ppm | **10** | 2 |
| **Tl** | ppm | **0.1** | 0.003 |
| **U** | ppm | **0.02** | 0.003 |
| **V** | ppm | **2** | 1 |
| **W** | ppm | **0.01** | 0.001 |
| **Y** | ppm | 0.05 | **0.005** |
| **Zn** | ppm | 0.2 | 0.2 |
| **Zr** | ppm | **1** | 0.1 |

### 3.1.3 Comparison studies of UltraFine+® with other analytical methods

The initial M0462 project and follow-on commercialisation of the UltraFine+® soil analytical method provided comparisons to industry standard extractions and identified the required sample weight for reproducible analyses. Additional improvements of the method were carried out during this subsequent M0462a project to meet the objectives associated with Milestone 1 (i.e., further lowering detection limits, adding Pt and Pd to the standard suite of element analyses and refining recovery of a range of resistate elements). Emphasis was placed on comparing UltraFine+® microwave-assisted, closed-vessel aqua regia to 4-acid extraction,

ALS clay separation, and Mobile Metal Ion (MMI) which was, at the time of the project onset, one of the most recent “new technologies” in partial extraction geochemistry for mineral exploration.

The ability to integrate available historic data is reliant on understanding the degree of correlation between the different methods and individual element responses. The integration of multiple data types with UltraFine+® results was outside of the project scope and not pursued further. However, the results described below suggest that integrating an UltraFine+® soil survey with soil survey results acquired with other techniques is not recommended, due to the increased recovery of many elements from the ultrafine fraction. The following discussion summarises the main findings of the various comparison studies.

#### *3.1.3.1 Comparison of Mobile Metal Ion (MMI®) and UltraFine+® analytical results*

Soil samples (n=146) from across numerous study sites, including the 26 composite representative sites used in the method development published in Noble et al. (2020) were analysed by the MMI® proprietary method used by SGS laboratories and compared to the UltraFine+® results of the same samples. A good correlation was evident for Au and a positive correlation for Cu and Ni (Figure 13). Silver and Pb were positively associated albeit with a weaker correlation. The concentrations of the metals measured for the same soils using UltraFine+® were orders of magnitude greater for most elements of interest. Gold concentrations extracted via the UltraFine+® method were 4 × greater, whereas Cu concentrations were almost 30 × greater. The enhanced but similar response for Au is positive since the methods are both aimed at targeting mobile metals, hence a positive association was expected. Direct comparisons with MMI® results were somewhat limited as numerous metals were not reported in the initial MMI data that were measured in the updated UltraFine+® method. However, key pathfinder and target element recovery such as Ag, As, Bi, Pd, Pt, Sb, Sn were poor using MMI® with most of these analytes below the method detection limit. Overall, the UltraFine+® method outperformed MMI® in the recovery of most metals from the soils tested (mostly WA soils).

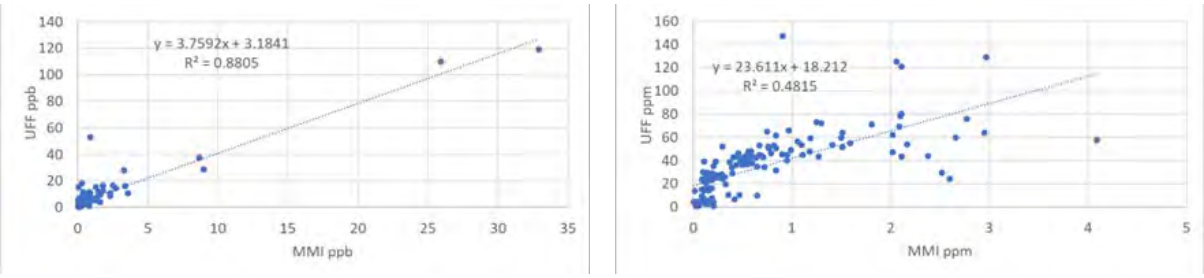



Figure 13. Comparison between MMI and UltraFine for Au (left) and Cu (right).

#### *3.1.3.2 Comparison of ALS Clay separation and UltraFine+® analytical results*

Another method for fine fraction geochemistry is offered by ALS in its Vancouver Laboratory. This method separates the clay fraction (<2 µm) and analyses for elements using aqua regia, very similar to the UltraFine+® method. However, ALS uses an open block (lower pressure and temperature) digestion compared to the pressure vessels used in the microwave method for UltraFine+® at LabWest. It was therefore expected that both methods would be very strongly correlated. Indeed, 21 Western Australian representative composite soil samples were compared and produced a nearly perfect correlation for Au ($R^2$ = 0.99; Figure 14). The correlation was less robust for concentrations below 5 ppb as shown in the inset of Figure 14. Copper was also very strongly correlated ($R^2$ = 0.95; Figure 14). However, the UltraFine+®

method extracted slightly greater concentrations which could be attributed to the more aggressive, closed-block digestion. Overall, there was no appreciable difference in the relative abundances recovered for the elements tested. The limitation of the ALS approach is the greater cost and that other parameters, such as pH and spectral mineralogy, are not part of the analysis.

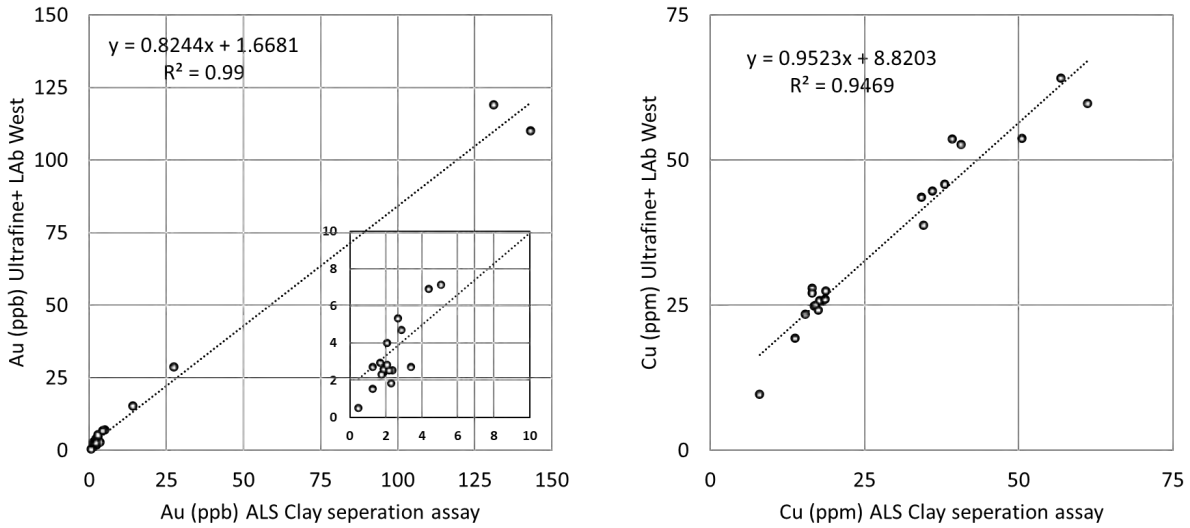



Figure 14. UltraFine+® analysis compared to ALS clay separation analysis for Au (left; inset shows the Au correlation at lower concentrations) and Cu (right).

### *3.1.3.3 Aqua regia v 4-acid v UFF*

A large dataset of 1,885 samples that were subject to a traditional aqua regia digestion (open block digestion at ALS), 4-acid digestion (near-total digestion at ALS) and UltraFine+® (at LabWest) were compared. All assays were conducted on the same samples. However each method uses a different size fraction. UltraFine+® separates the <2 µm fraction from the bulk or <2 mm soil sample fraction, whereas the ALS analysis uses a dry-sieved <75 µm fraction for the standard aqua regia and 4-acid analysis.

The UltraFine+® results for most trace, target and pathfinder metals were greater than the standard aqua regia (blue versus green in Figure 15). Gold is shown on a log scale with higher gold concentrations recovered via the UltraFine+® method. It should be noted that Au results are not reported in the 4-acid extraction. For the other reported elements, the 4-acid extraction (pink distribution in Figure 15) recovered greater concentrations than the aqua regia (green distribution in Figure 15), but lower concentrations than the UltraFine+®. These results were expected due to the increased surface area/exchange capacity and host materials contained in the ultrafine fraction recovered via the UltraFine+® method compared to the larger size fraction (up to 75 µm) recovered via the other methods which was expected to dilute mobile element concentrations due to the presence of, e.g., quartz in the coarser particle sized materials. These findings were confirmed in other smaller study sets (Noble et al. 2020; Henne et al. 2022; Henne et al. 2023).

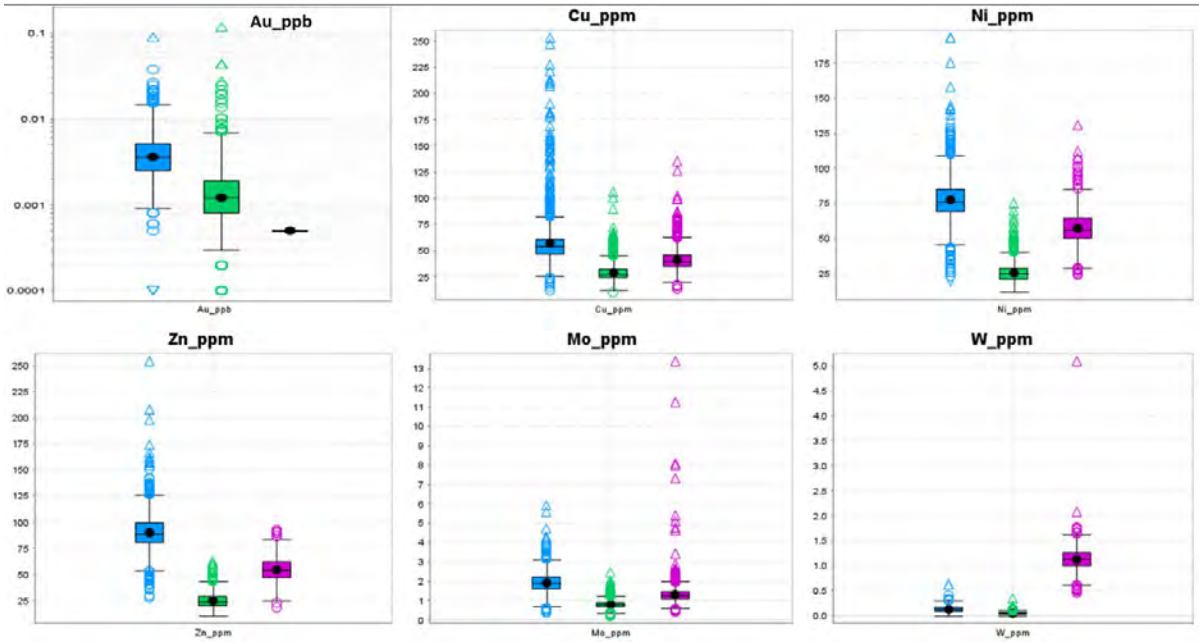



Figure 15. Comparison of extractions from samples subject to UltraFine+® aqua regia assay (blue), standard aqua regia (green) and 4-acid "total" digestion (pink) for Au, Cu, Ni, Zn, Pb and W; n = 1885. Note the log scale on the Au plot. Boxplots include values below the detection limit (replaced by half the respective detection limit).

While base and precious metal recovery were greater with the UltraFine+® method, the 4-acid was expected to better recover those elements that are more resistant to weathering and acid digestion. These resistate elements include Ge, Ti, W and Zr and, as expected, recovery was greater in the 4-acid extracted results. Silica is not part of the UltraFine+®assay. Concentrations of elements Mo, Cr and Sn that can also be resistant to weathering were recovered well via the UltraFine+® method. These are important findings for explorers looking for polymetallic ores, or those who use Cr as a lithological indicator element. However, of the Mo results, 11 samples would not have been measured as elevated using the UltraFine+® results compared to those measured with the 4-acid digestion. This is equivalent to 0.5% of the data (Figure 16). Hence, while the UltraFine+® method is effective for exploration purposes, results indicate that, if Mo was the sole target metal, the 4-acid approach is likely preferable.

A similar consideration was applied to REE at the start of the project, hence only limited REE analyses are available (Ce, La, Sc and Y) as the method was not believed to be effective in capturing the more resistate element signatures. With a rapid increase in demand for REE, LabWest added a range of additional REE and the recovery of REE via the UltraFine+® method was assessed in preliminary tests. Two data sets were compared, one containing samples from the North Yilgarn (soils discussed for Mo previously) with the caveat that only a few REE representative elements were available (Ce and La; Figure 17). This area was not sampled for REE target generation although it is likely that soils respond similarly. The second data set (n=15) contains soil analyses from an exploration survey targeting ionic REE-bearing clays. This data set reported the full suite of REE measured (Figure 18). Importantly, the results showed REE patterns were similar for both data sets. The UltraFine+® method measured more than 100 % of the REE signature that was extracted using the 4-acid (near total) extraction. This indicates that there is no advantage in applying the stronger extraction method for REE exploration using near surface samples. This is likely due to the closed, pressured vessels used during the UltraFine+® method, which appear to produce a near-complete digestion. This difference is evident in the larger dataset comparing Ce, where the

traditional aqua regia extraction only recovered approximately 66 % of the total recovered via 4-acid extraction. UltraFine+® recovered on average 106% from the same soil samples. It is important to note that the <2 µm fraction is not the exact same sample as the <75 µm milled samples used for the 4-acid digestion and hence this likely contributes to the increased recovery > 100 %. Correlation analysis between the results shows $R^2$= 0.91 and very similar concentrations (y=1.02x; Figure 19).

For the ionic clay samples, a high degree of similarity was expected as the REEs are hosted in the clays specifically targeted by UltraFine+®, but this was not the expectation in other soils. Extended testing of soils containing coarse, resistate REE minerals hosting REE in soils will be completed in the future, and it is anticipated that the 4-acid extraction may perform better than the UltraFine+® method in this setting and also in (down hole) saprolite samples.

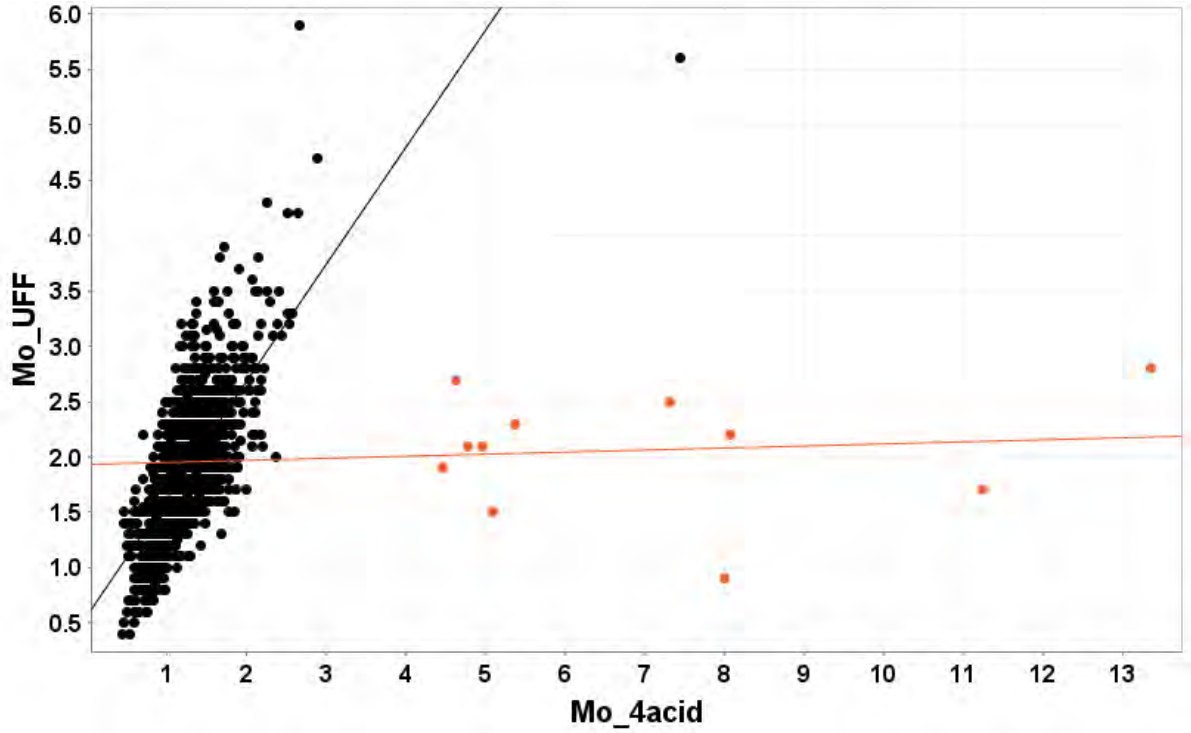



Figure 16. UltraFine+ (closed-vessel aqua regia) compared to 4-acid extracted Mo from ~1885 soils collected in the Yilgarn Craton. The 11 orange sample points were elevated in the 4-acid digestion and not observed in the UltraFine+ extraction.

Other comparison studies showing the comparisons to hydroxylamine hydrochloride, weak aqua regia and water (Noble et al. 2020) for ~30 representative soils from WA were conducted as part of the previous M0462 project and presented results that support the above observations.

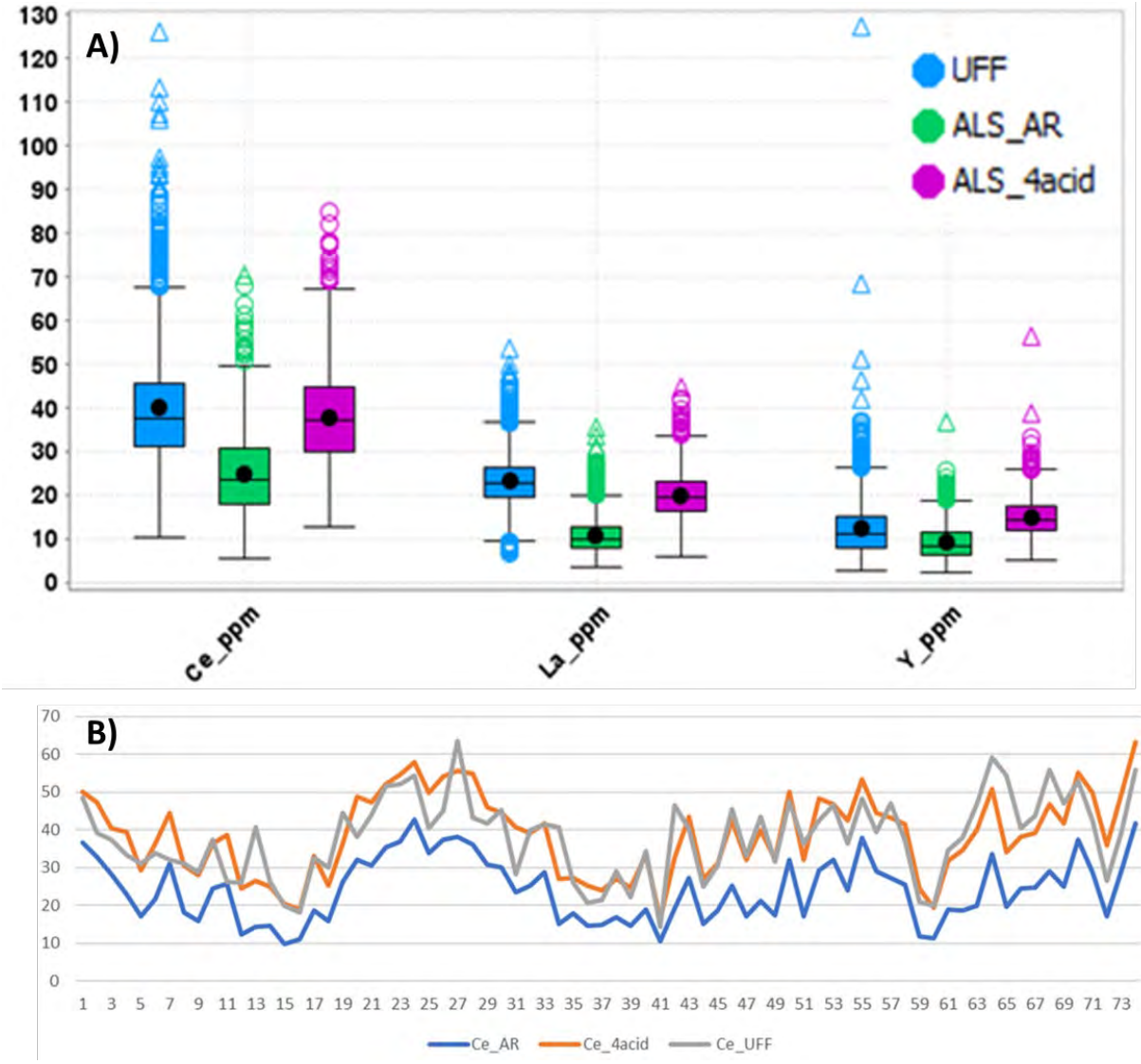



Figure 17. A) Comparison of 1885 soils from the NE Yilgarn using traditional aqua regia and 4-acid digestion on the 75 µm size fraction compared to UltraFine+® (UFF) on the same samples for Ce, La and Y (ppm). B) Ce concentrations of the first 75 sample pairs showing the similarity of relative abundance patterns.

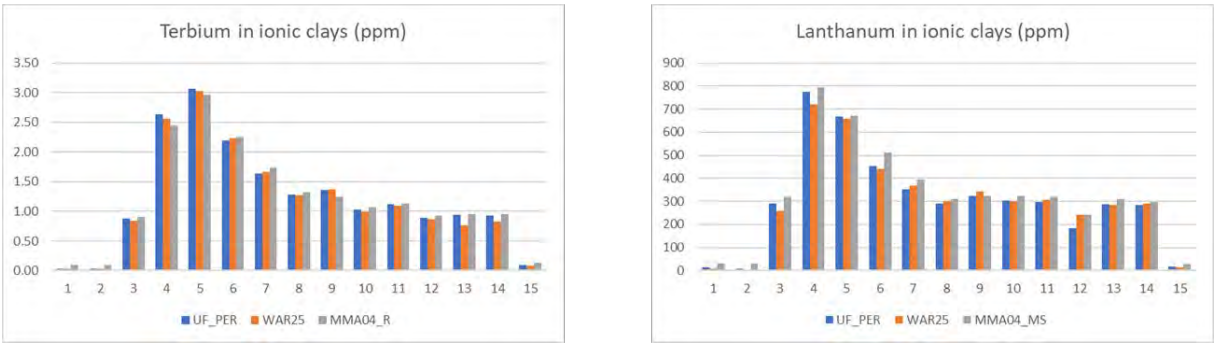



Figure 18. Comparing UltraFine+®, traditional aqua regia and 4-acid digestion for Tb (HREE example) and La (LREE) from 15 pairs of ionic clay soil samples.

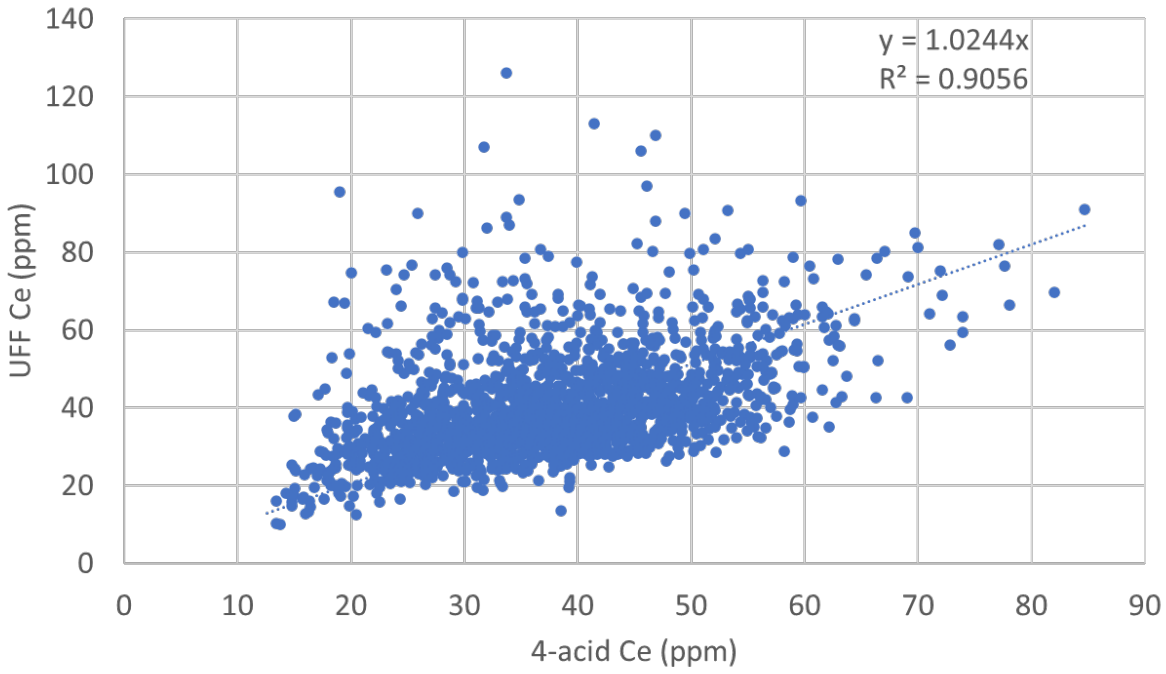



Figure 19. Cerium concentration of soils from UltraFine+® extraction (<2 µm) compared to 4-acid extraction (<75 µm) using the same samples (n=1885).

### 3.1.4 Additional comparison studies

In addition to the extraction method comparison studies above, two further comparison studies were necessitated by the submission of soil samples from overseas, and submission of samples that were pulped or milled prior to submission.

#### *3.1.4.1 Heat treatment for quarantine*

Soil samples imported for analysis from overseas are required by law to undergo strict Australian quarantine procedures to prevent the introduction of harmful biological materials. This presented challenges in the project for selected sites due to limited access to approved quarantine premises, the logistics and expense of releasing samples from quarantine and concerns about the effect of approved procedures to release samples from quarantine. Gamma irradiation was initially considered the most suitable approach to minimise any potential alteration of materials by other methods (e.g., heating or addition of strong acids). However, initial trials indicated that logistics, access and costs were prohibitive, and LabWest opted to acquire a quarantine licence for sample receival to enable rapid, commercially viable sample processing. For samples to be released from quarantine they are required to be heat treated (to >160°C for 2 hours). While there was minimal concern that the heat treatment would alter the chemistry of the samples, clay interlayers can be altered at higher temperatures. The removal of loosely bound water and nanoscale-confined interlayer water occurs when heating kaolinite from ambient temperature to 150 °C (Gadikota et al. 2017). This does not greatly alter the interlayer distance or the structure in non-swelling clays (e.g., kaolinite), but it can affect swelling clays such as Na-montmorillonite and Ca-montmorillonite. More significant transformations of clays occur at greater temperatures such as illite alteration at >300°C and kaolinite/halloysite alteration between 400 °C to 700 °C (Jackson 1964). Therefore, it was possible that dehydration of clays and the interlayers above 105 °C might modify the available metals and alter the spectral mineralogy. For this reason, a simple heat

treatment test was conducted where 103 samples were split, with half undergoing heat treatment and half used as controls. Both batches were analysed with the UltraFine+® workflow to measure elemental abundance, soil physico-chemical properties and spectral parameters. Results indicated that there were no significant differences for most trace metals such as Cu, Mo, Ni, Pb, Sb, Zn and that the spectral responses in NIR and FTIR also showed no appreciable difference between treated and untreated samples (Figure 20).

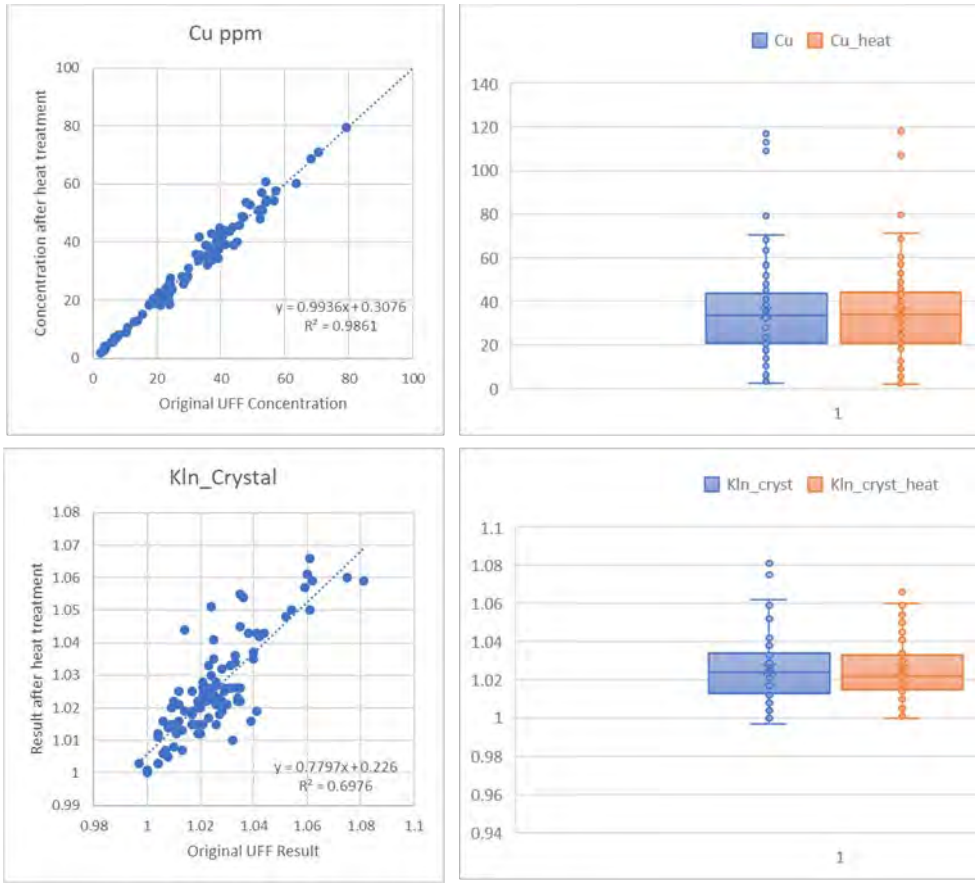



Figure 20. Comparison of ultrafine soil samples that underwent heat treatment and control samples. Cu results (top row) show excellent correlation similar to VNIR kaolinite crystallinity index (bottom row).

The measurement of pH was found to be reduced (more acidic) by less than half a unit on average in the heat-treated samples, although there was a strong correlation between samples. Given that the entire batch of a given survey is subject to heat treatment (if imported) it is deemed that interpretation will not be impacted by this slight shift (Figure 21). Relative Au concentrations showed noticeable variation only near detection limit concentrations and are not expected to influence interpretation. Overall, the test results demonstrated that using heat treatment to release international soil samples from quarantine did not affect the usability of the UltraFine+® method.

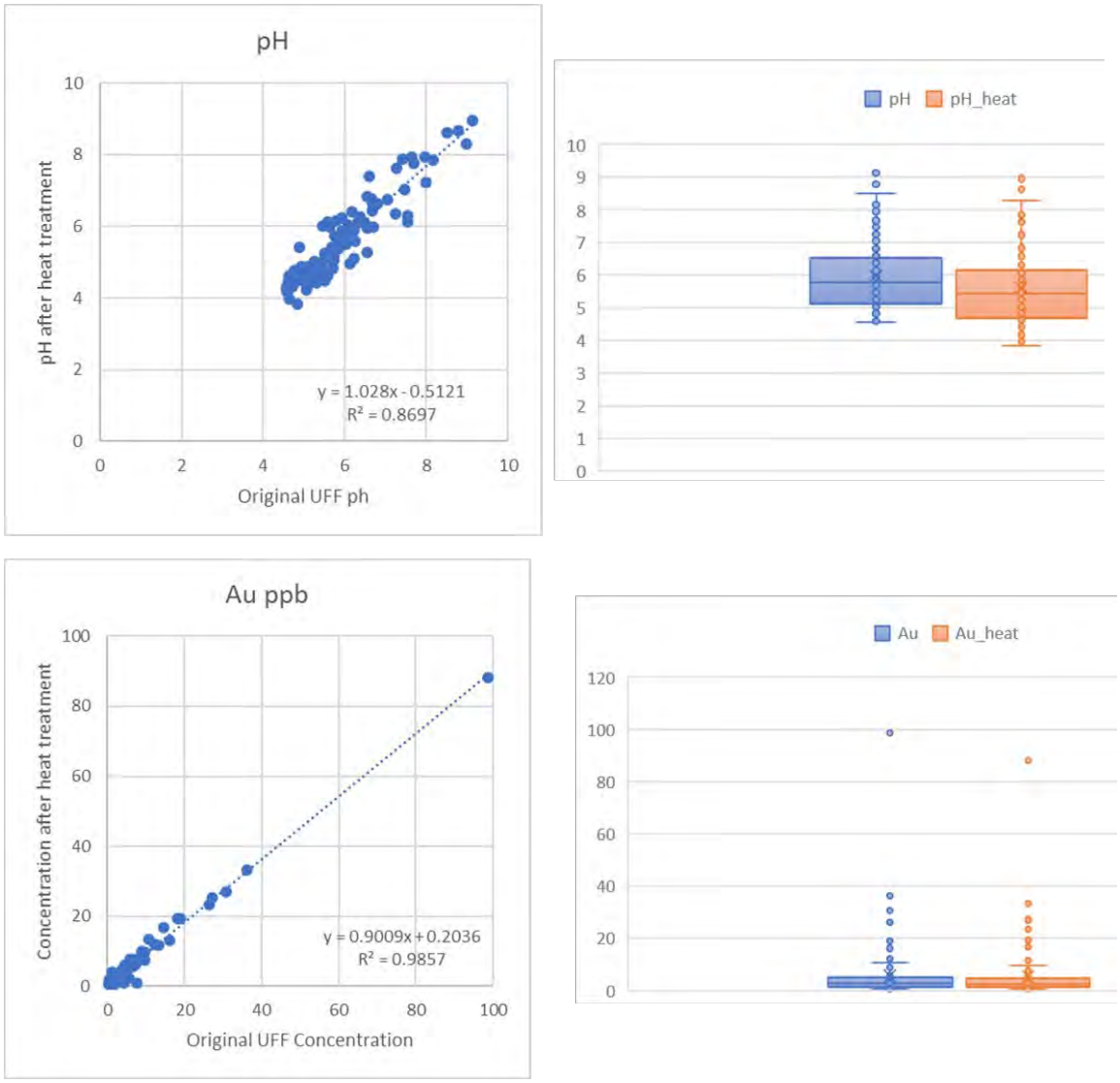



Figure 21. Comparison of ultrafine soil samples that underwent heat treatment and control samples for pH measurements (top) and gold (bottom). The results have excellent correlation but were found to be statistically significantly different.

### *3.1.4.2 Analysis of milled samples compared to bulk samples via UltraFine+®*

Historically, many companies have collected bulk soil samples that were milled to reduce the particle size for standard analytical methods. Since the UltraFine+® does not require this step, the concern arose that archived samples that require re-processing (rather than re-sampling) via the UltraFine+® method may not be suitable for comparative analysis. Early research testing therefore investigated whether UltraFine+® results may be altered when soil samples are pulped or milled.

Pulped/milled (nominal <50 µm) and non-pulped/milled samples were initially compared for 30 samples, with some minor variance observed in the major elements Fe, Al and Ca. Results showed no significant differences for Au (Figure 22) as well as trace metals such as Cu, Zn, and Pb. The initial outlier noted in the overall population was the same sample in both test series although the relative concentrations differed with 11.7 and 8.6 ppb Au recovered, via milled and non-milled, respectively. This result was reassuring for analysing historic samples to identify previously undetected anomalous samples and indicates the samples do not host

nuggety gold. Additional test results (scatter plot; Figure 22) showed a positive correlation between the milled samples and those that were not subject to milling.

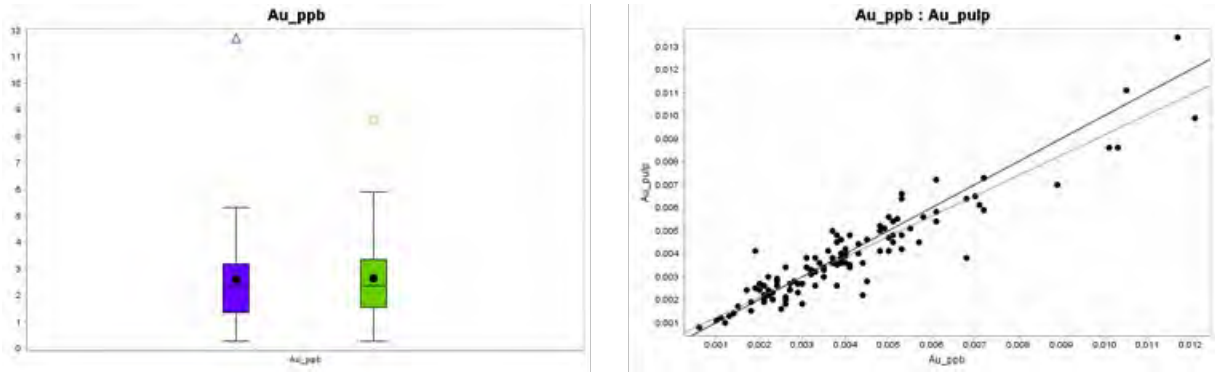



Figure 22. Milled (purple) versus non milled (green) soil samples subject to UltraFine+® analysis. The distribution of the initial Au data on the left and the scatter plot of the later test samples on the right. The perfect fit (1:1) line is shown in the darker trendline with the lighter trendline showing the line of best fit (linear regression).

### 3.1.5 Addition and refinement of other soil properties in the UltraFine+® workflow

One of the main aims of the M0462a project was to develop more mineralogical proxies acquired via spectroscopy. To this end, the existing visible and near or shortwave infrared (VNIR) scalars were refined and additional proxies using FTIR were developed. Reflectance spectral soil properties provide important information about soil samples that relate to mineralogy and landscape settings and were added to the UltraFine+® method to better interpret soil geochemical results.

#### *3.1.5.1 Visible and near or shortwave infrared spectroscopy (UV-Vis/NIR/SWIR)*

A total of 17 spectral soil parameters in the visible to shortwave infrared (350-2500 nm) reflectance range were refined or added to the standard UltraFine+® method.

The initial scalars developed during earlier CSIRO projects focussed predominantly on alteration minerals in fresh rocks, drill core, chips and pulps. The naming system of these initial scalars were abbreviated descriptions of the wavelengths and parameters. As part of this project the scalars were reassessed and re-designed to apply the commonly used algorithms to Ultrafine+® soil materials based on spectral measurements of the samples. Scalars that targeted features that were extremely rare and were limited to very few occurrences in the test dataset were excluded. The naming system of these scalars was improved to be more suited to mineralogical naming conventions as well as making them easier to understand for non-spectral geologists, and an explanation document was created. In addition, a quality control scalar was added using spectral features and properties extracted by TSG to flag potentially noisy or unusual spectra to aid consistent quality control within the laboratory. The scalars listed in Table 10 are the final parameters developed during the project that are now routinely reported as part of the commercially available UltraFine+® soil analytical workflow. Importantly, the application and limitations of these scalars in a surface exploration context were investigated and are recorded in column 3.

**Table 10**: Spectral outputs (parameters) from the Next Gen Analytics workflow for the various relative abundances of spectrally active minerals, their mineral chemistry and an indication and limitation of these parameters.

| Parameter | Definition | Applications & Limitations |
|---|---|---|
| **MinGrp1** | Main mineral-group detected in the shortwave infrared region | Report of the dominant mineral-group present, which may be deemed to have contributed >51% to the spectral unmixing algorithm. Uses mineral-groups rather than individual minerals (i.e., white mica rather than muscovite). Only reports mineral-groups which are spectrally active in the 1300-2500 nm wavelength region, thus does not report iron oxide minerals.<br>See section below this table for more information. |
| **MinGrp2** | Secondary mineral-group detected in the shortwave infrared region | May make up <49% of the predicted mineral-group in the sample and reported the same as MinGrp1. |
| **FeOx** | Relative abundance of iron oxide based on the strength of the iron oxide absorption around 900 nm | Indicates the presence of iron oxides and may indicate that the original material was iron-rich. Typically, as this is a surface technique it can disproportionally represent coatings on grains, but with UltraFine samples the material is typically fine grained. Ranges are <0.01 very low; 0.01 - 0.02 = low; 0.02 - 0.2 = moderate; >0.2 = high amounts of iron oxide. |
| **Hem_Goe** | Iron oxide species (measurement of the wavelength in nanometres of the iron oxide absorption, which relates to the hematite or goethite proportion) | Hematite absorption is ~890 nm whereas goethite is ~920 nm. Thus, lower numbers represent more hematitic materials, whereas higher values mean the material is more goethitic. Can indicate the type of weathering that has occurred and redox conditions. In some regions can be used to differentiate palaeoenvironments during deposition. Can also include less common iron oxide and oxyhydroxide minerals (such as acid rock drainage minerals: jarosite, schwertmannite, ferrihydrite etc. and lepidocrocite). Magnetite and maghemite are not spectrally active in this wavelength region. |
| **Kln_abun** | Relative abundance of kaolinite | Main absorption features coincide with white mica and aluminium smectite and can be influenced where both minerals exist. Ranges are 0.1 - 0.15 = low; 0.2 - 0.25 = moderate; >0.25 = high amounts of kaolinite. |
| **Kln_cryst** | Kaolinite crystallinity | Useful for regolith mapping of soils (poorly crystalline = transported vs. highly crystalline = *in-situ*). Ranges are <1.004 = disordered/mixture with smectite; 1.025 - 1.004 = moderate; >1.04 = high. Note very low numbers <0.97 could be dickite. |
| **WM_AS_abun** | Relative abundance of white mica and/or aluminium smectite | This result is excluded if the material contains a kaolin absorption feature at 2160 nm >1.005. Ranges are <0.05 = low; 0.07 - 0.14 = moderate; >0.14 = high amounts of white mica/Al-smectite. |
| **WM_AS_comp** | White mica and/or aluminium smectite composition (amount of Al-substitution based on wavelength in nanometres) | This result is excluded if the material contains a kaolin absorption feature at 2160 >1.005. Typically indicates less weathered material, or soils from parent rocks. May indicate shallow depth of cover/subcrop or colluvial transport of material from outcrop upslope. Changes in the wavelength (chemistry) of white micas has been linked to alteration and has been used as a vector for mineralisation.<br>Lower values are Al-rich (<2200 nm to - 2208 nm = muscovite), whereas higher values are Al-poor white mica (>2209 nm = phengite) or Al-smectite (~2209 nm = montmorillonite, >2215 nm = palygorskite). |
| **Fe_Kln** | Relative abundance of iron substitution in kaolinite | Iron substitution in kaolinite has been found in highly weathered materials over mafic materials and can indicate the parental materials were originally iron-rich. Ranges are 0.01-0.056 = low; 0.016 - 0.03 = moderate; >0.03 = high. This result is excluded if the material contains chlorite. |
| **Chl** | Relative abundance of chlorite and dark mica (biotite and phlogopite) | Requires a high proportion of chlorite or biotite to be present, where white mica and kaolinite are absent. This parameter can also falsely include palygorskite (compare to Paly parameter below). Biotite is spectrally dark and can be |

| | | |
|---|---|---|
| | | difficult to detect. Uncommon in highly weathered soils but may occur close to exposures of fresh rocks or very shallow cover. Ranges are 0.01 - 0.015 = low; 0.015 - 0.025 = moderate; >0.025 = high. |
| **FeMgClay** | Relative abundance of iron and magnesium smectite | Indicates the presence of Fe and Mg smectite minerals, such as nontronite and saponite, which are associated with the weathering of mafic materials. Ranges are 0.01 -0.014 = low; 0.014 - 0.022 = moderate; >0.022 = high. |
| **OH_mafic** | Mafic minerals with OH (i.e., amphiboles) relative abundance as well as Fe/Mg smectites | Indicates the presence of significant amounts of OH-bearing mafic minerals and carbonates, where there is a lack of kaolin and white mica. Ranges are 0.06 – 0.08 = low; 0.08 - 0.12 = moderate; >0.12 = high, |
| **Paly** | Palygorskite | Palygorskite (also known as attapugite) is a Mg Al clay mineral often associated with alkaline conditions and inland lakes. It is highly absorbent and may be associated with elevated metal concentrations. Ranges are 0.04 – 0.07 = low; 0.07 - 0.17 = moderate; >0.18 = high. |
| **Colour** | Munsell Colour; results from the spectral data in standard "hue value/chroma" Munsell notation, typically used for describing soil and regolith materials colour | Colour can be related to the mineralogy and organic content. |
| **Hue** | A numerical (0 - 360) representation of the colour wheel of Red-Yellow-Green-Cyan-Blue-Magenta(-Red); hues around 0 are red, whereas yellows are ~40 - 60 | Gives an indication of the colour of the material. |
| **Saturation** | Refers to how washed out or pure the hue is, with low values appearing greyer; values range from 0 - 1 | |
| **Intensity** | Also known as Value on the HSI or HSV scale of colour or Luminance and refers to how light or dark a hue is; values range from 0 - 1. | Dark material may contain organic carbon or dark minerals, whereas lighter material may contain highly weathered materials or carbonates. |

The Mineral Groups 1 and 2 report the main mineral phases in the ultrafine fraction of a sample, based on a spectral reference library from TSG (The Spectral Geologist® software). Some of the minerals reported have been grouped into broader classes as noted in the definition of Table 11 where kaolinite, dickite, halloysite and nacrite may be individually identifiable, but for more general application were deemed to be captured as “kaolinite”. Not all minerals have diagnostic features in the wave lengths measured and thus are not identifiable. For example, quartz is not directly detectable in the SWIR region but can be diagnostically detected in the thermal infrared wavelength using FTIR (Fourier transform infrared) spectroscopy (see next section). The Mineral Group 1 represents the major SWIR-active mineral component of the reflectance spectra (> 50 %), whereas Mineral Group 2 is the secondary component and can constitute up to 49% of the match to the spectral library. Carbonates are detectable by SWIR in high abundances but were almost never observed in the soil samples that were part of this research project. This may partly be due to the size fraction separation and/or that carbonate materials are often coated by clays and Fe-oxides. Carbonates are reported in the FTIR data allowing cross reference and confirmation of key changes observed in the spectral output delivered in the UltraFine+® data.

**Table 11**: Minerals identified from the SWIR in the Mineral Group 1 and 2 outputs and notes of the applications and limitations of these identified mineral signatures in the ultrafine fractions.

| Mineral | Definition (example minerals) | Applications & Limitations |
|---|---|---|
| **Kaolin** | Kaolinite, halloysite, dickite, nacrite minerals | Kaolinite is a common mineral in weathered terrains, the crystallinity can indicate transported vs. *in situ* material. Halloysite has a very similar spectral appearance to a mixture of kaolinite and montmorillonite. |
| **Smectite** | Montmorillonite, nontronite and magnesium clays (saponite) | Commonly found in alluvial settings (montmorillonite) and weathering close to mafic rocks (nontronite/saponite), as well as in weathering of volcanic rocks. |
| **White-Mica** | Muscovite, phengite, paragonite, illitic varieties of white mica | Commonly found associated with shallow soils near or adjacent to granitic material. May also be associated with feldspar, which is not detected by SWIR. |
| **Amphibole** | Hornblende, tremolite, actinolite, riebeckite | Uncommon, but may be associated with shallow soils near or adjacent to mafic material. |
| **Chlorite** | Clinochlore, chamosite | Associated with shallow soils near or adjacent to chlorite bearing rocks, particularly where they are fresh. |
| **Dark-Mica** | Biotite and phlogopite | Dark micas tend to weather rapidly once exposed and are uncommon. |
| **Other_ALOH** | Group of other Al-bearing minerals | Pyrophyllite, prehnite, diaspore and gibbsite. Of these, gibbsite is the most likely to be seen in soils, particularly in highly weathered terrains (gibbsite can be cross checked against the FTIR data). |
| **Other_MgOH** | Group of other Mg-bearing minerals | Indicator for talc and hornblende. Not commonly seen as a major component of soils. |
| **Carbonate** | Calcite, dolomite, magnesite | May be present where calcrete occurs, but SWIR is less reliable than the thermal infrared wavelength regions used by the FTIR for detecting carbonate. Carbonates can be easily coated by other materials which are active in the SWIR, making them less detectable. Due to the carbonate region overlapping with other minerals, sometimes white mica, chlorite and FeMg clays can be misinterpreted as carbonate (cross check with FTIR carbonate data). |
| **Sulfate** | Gypsum, alunite, jarosite | Gypsum can be found near salt lakes and saline soils; the other minerals are uncommon outside of acid rock drainage environments. |
| **NOTAROK** | Anything that is not a mineral, this could be plastic, wood, etc. | |
| **OTHERMINS** | Other rare minerals for UltraFine soil | Minerals that occur only very rarely in UltraFine soils. One example could be serpentine that may be a product of weathering of ultramafic rock. |
| **NOMATCH** | No mineral-groups in the library match the sample | NULLs reported in the Mineral Groups 1 and 2 are translated to NOMATCH as these are minerals that have no match in the current library. NOMATCH will not appear in raw data from LabWest and will only be present in machine learning outputs. |

Null values were common and indicate the sample may lack a significant enough diagnostic absorption feature for that parameter, which would indicate that the mineral pertaining to that parameter is not present or cannot be spectrally detected (possibly due to the mineral being coated with another mineral). Much of the soils in Australia contain minerals associated with weathered materials, but some areas may contain minerals more associated with primary rock minerals (e.g., lithosols in the Australian Soil Classification System), particularly where samples are close to exposures of outcropping rock. To address this in the project the workflow replaces NULLs in Mineral Groups 1 and 2 with NOMATCH, to indicate that a mineral may be present, but cannot be reliably identified from the shortwave infrared spectral library.

For example, a sample with high amounts of Fe-oxides and quartz will produce NULLs for the Mineral Groups.

#### *3.1.5.2 Fourier transform infrared spectroscopy (FTIR)*

In addition to the spectral parameters acquired in the VNIR region, scalars were developed for a range of chemical properties using the mid-infrared spectrum, and these were tested for inclusion in the UltraFine+® workflow. Given that the mid-infrared spectral region comprises many fundamental molecular/bond vibrations that are associated with different substances that are present in the soil, effort was devoted towards understanding a few of the essential mineral and organic properties. ATR-FTIR spectroscopy was used to provide semi-quantitative information on five parameters, and these are displayed in Table 12. The final UltraFine+® workflow includes FTIR spectroscopy data on the clay, carbonate, quartz and organic carbon content. In addition, the absence or an approximate estimate of the amount of gibbsite present is also included. Generally, differences that are greater than ±1 % (for quartz, carbonate and organic carbon contents) and ±10 % (for clay) are considered significant. Most of the IR bands that appear in the mid-infrared region between 4000 to 400 $cm^{-1}$ are predominantly due to mineral/inorganic materials along with some minor features arising from organic substances. Multiple absorption features and peaks commonly occurred at 3700-3100 $cm^{-1}$, 1800-1400 $cm^{-1}$, 1200-800 $cm^{-1}$, and 800-400 $cm^{-1}$. The peak intensities varied significantly from one ultrafine soil sample to the next and this is indicative of variations in the relative amounts of the different mineral and non-mineral phases. It is important to note that certain features stood out in the infrared spectra. Generally, peaks occurred at 3750-3500 $cm^{-1}$, 1150-850 $cm^{-1}$ and 700-400 $cm^{-1}$, which are due to clay minerals. As for silicate-based minerals such as quartz, intense absorption peaks were observed around 1080 $cm^{-1}$, 780 $cm^{-1}$, 695 $cm^{-1}$ and 460 $cm^{-1}$. However, carbonates displayed characteristic bands at 1450-1400 $cm^{-1}$, 870 $cm^{-1}$ and 700-740 $cm^{-1}$, whereas organic matter exhibited weak peaks but the most prominent were at 3000-2800 $cm^{-1}$ and 1750-1450 $cm^{-1}$. The mineral (clay, quartz, carbonate and gibbsite) and TOC (total organic carbon) quantification was investigated in the ultrafine soil using internally developed calibration standards and the results were compared against various laboratory techniques (i.e., X-ray diffraction, inductively coupled plasma spectrometry, carbon analyser). It is important to note that regression analysis was performed using a wide range of samples and generally high correlation coefficients ($R^2$ >0.7) were achieved between ATR-FTIR spectroscopy and the standard techniques.

**Table 12. FTIR spectral outputs (parameters) from the Next Gen Analytics workflow for the various relative abundances of spectrally active minerals, their definition and an indication and limitation of these parameters**:

| Parameter | Definition | Applications & Limitations |
|---|---|---|
| **Clay_wt%** | Estimate of total clay content | It is sensitive to certain clay groups (i.e., kaolinite, montmorillonite/smectite, illite, muscovite) but it does not distinguish between different clay species. The FTIR results on samples that have clay contents of < 20 % are less reliable and are limited to samples with low (<20 %) quantities of gibbsite (a NULL or 1 result in the Gibbs_Index). The quantification error can vary by up to ± 10 wt. % (depending on the clay content). |
| **Qtz_wt%** | Estimate of quartz content | It is sensitive to quartz and is limited to samples with low amounts of feldspar (<10 %). Quartz contents of <1 % and above 10 % are less reliable. The quantification error can vary by up to ± 1 wt. % (depending on the quartz content). |
| **Carb_wt%** | Estimate of total carbonate content | It is sensitive to a range of carbonates, but it does not distinguish between different carbonate species. Carbonate contents of <1 % and above 40 % are unreliable. If large quantities (>20 %) of carbonate are measured, further processing can separate some carbonate types, but this is very uncommon in UltraFine+ soil samples and is not part of the standard output. The quantification error can vary by up to ± 4 wt. % (depending on the carbonate content). |
| **TOC_wt%** | Estimate of total organic carbon content | It is sensitive to a range of organic matter, but it does not distinguish between different types of organic matter. The TOC values <1 % or >15 % are less reliable and it is limited to samples with a low carbonate content. Typically, samples with <15 % carbonate will not influence the result but can be cross checked with Carbwt% result. The quantification error can vary by up to ± 2 wt. % (depending on the TOC content). |
| **Gibbs_indx** | A classification that provides an approximate amount of gibbsite abundance | Limited to samples with a low metal hydroxide content. Typically, samples with >10 % metal hydroxide may influence the result and can be cross checked with the VNIR data. |

Similarly to the NIR spectral data, Null results are reported and are due to either (a) the amount of various components (i.e., clay, carbonate, quartz, organic carbon or gibbsite) being below the detection limit, (b) spectral interferences/variations which occur due to the nature of the sample that result in less accurate data, and/or (c) the abundance values may be outside the linear concentration range. The gibbsite index is a proxy for gibbsite abundance developed in the Next Gen Analytics project as a very broad classifier (Table 13). This is required in the workflow as the exact abundance (in wt. %) was not established with known gibbsite standards.

**Table 13**: The FTIR gibbsite index parameters used in the Next Gen Analytics outputs.

| Gibbsite Index | Definition | Actual values |
|---|---|---|
| **NULL** | Excluded data | values <0.019 |
| **1** | Low | values between 0.02 to 0.19 |
| **2** | Medium | values between 0.20 to 0.59 |
| **3** | High | values >0.6 |

#### *3.1.5.3 pH and sizing*

Particle size analysis, pH and EC measurements were developed during the initial M0462 project and did not require further development. However, during this subsequent M0462a project, the results of duplicate soil analyses submitted by project sponsors were investigated to design metrics which could be applied by the end-user to quantify and assess the quality of soil analytical data. These were added to the machine-learned QA/QC procedure described above (see section 1.4.2;

**Table 3**).

## 3.2 Next Gen Analytics Landscape Models (Milestones 3 and 4)

Understanding the relationships between soil geochemistry and general landform characteristics is a complex undertaking and can hinder successful greenfields exploration in areas where comprehensive datasets, field knowledge, general expertise, or time are limited. The second major objective of the M0462a project was therefore to develop an unsupervised machine learning workflow to generate proxies for landscape types from spatial feature layers to provide context for surface geochemical data interpretation (see section 1.4.3 for details on the methods applied). Next Gen Analytics models were developed with the intention of having little direct human input (e.g., they do not include available geological or regolith maps). The resulting proxy landscape maps were generated for all project sites and used to group geochemical data according to the landscape cluster corresponding to each soil sample location. In the following discussion, examples are provided of how to assess these models against spatial input layers, available regolith and surface maps, as well as how the landscape models and associated outlier calculations can be used to interpret soil surveys in the context of the respective machine-learned landscape setting.

### 3.2.1 Assessment of landscape models

The standard output of the workflow provides three different landscape maps for each site (with 4, 8 and 12 landscape clusters), selected and developed to accommodate the wide range of different project sites. The appropriate number of clusters to choose for a given project area depends on area size, the resolution of the input layers, the complexity of the respective landscape, and the needs of the end-user.

In the first instance, the appropriate number of landscape types in a given project area can be ascertained using the 3-dimensional representation of pixels in a project site (top row in Figure **23**). The closer the data points are to each other in the derived clustering, the more similar they are. If most points are close together, this indicates that there is not much landscape variability between those points. Points that form distinct clusters, especially offset from other clouds of data points, are more distinct landscape types – the more distinct clusters that are evident, the more different landscape types are likely to exist in a modelled area. Figure 23 demonstrates the use of multiple clustering algorithms (kmeans4, agg8, and agg12) comparing the three standard outputs for one site spatially in 2D and in the cluster boxes in 3D. In all three models of this project area, the main landscape types (sub-crop/residual in dark brown, and varying types of cover) are differentiated and remain the same. However, with an increase in landscape clusters, additional landscape feature details (e.g., colluvial/side-slope materials, differentiation of transported cover by e.g., composition (based on sentinel-2 ratios and radiometric data) and depth of cover (based on MrVBF) become apparent.

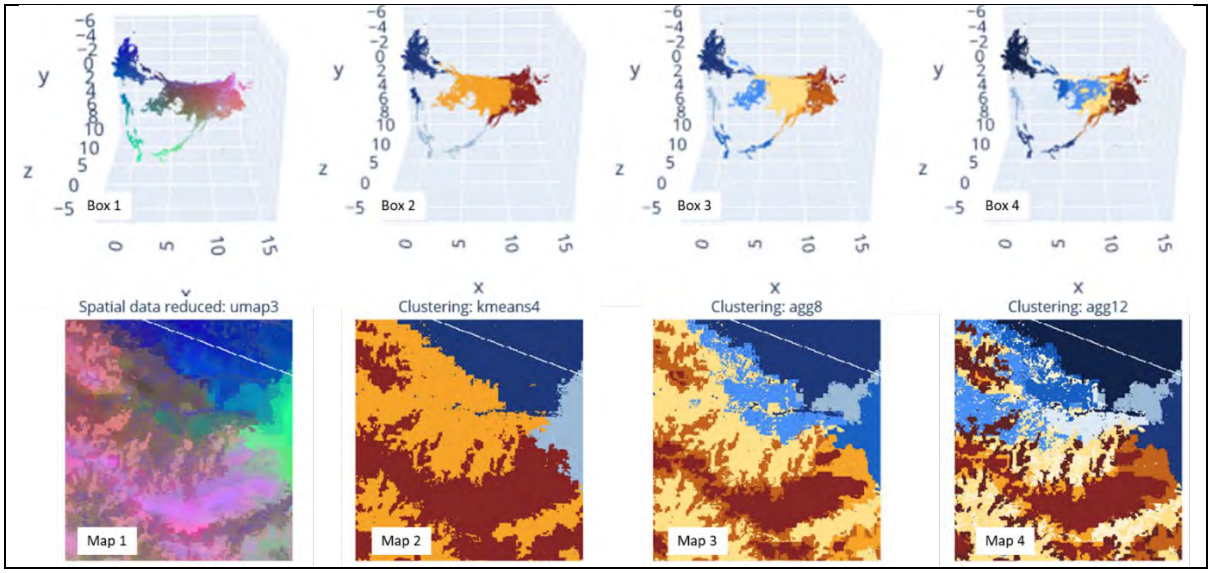



Figure **23**: Top row: Example of cluster boxes showing how the samples are associated in 3D space using an RGB colour scheme (Box 1) and coloured by landscape type (Box 2 with four landscape types, Box 3 with eight landscape types, Box 4 with twelve landscape types). Bottom row: Points from the cluster boxes projected in real-world spatial context creating proxy regolith maps with an RGB colour scheme applied (Map1) and with four (Map 2), eight (Map 3) and twelve (Map 4) landscape types over the same project area. The white diagonal line is a masked-out road and is not part of the clustering schemes.

Since the UltraFine+® Next Gen Analytics workflow was developed to generate landscape context for geochemical samples (see section 1.6.5), the number of samples across a given area must also be considered to avoid small clusters with few or no samples (see discussion below). The resulting number of landscape clusters are, therefore, a balanced approach to represent the major landscape types while also enabling meaningful interpretation of geochemical data.

Frequently, a landscape model featuring a reduced number of landscapes is favoured. This may fall below the ideally recommended number of landscape classes of the site and is underpinned by the primary objective of the Next Gen Analytics outputs which aims to impart contextualisation to geochemical data. Consequently, a pivotal consideration revolves around the density of samples encompassing the project site, with the intention of circumventing the emergence of minute clusters housing a limited number (<50) of samples or even devoid of any samples altogether. Such scenarios could result in rendering the statistical analyses for a project site devoid of meaningful significance. The selection of an appropriate number of landscape clusters is poised between the representation of significant landscape types and the facilitation of substantive interpretation of the geochemical data (i.e., having enough landscapes and enough samples in each landscape to confidently identify anomalous samples). It is noteworthy that the utilisation of a more exhaustive landscape model remains viable for strategic purposes in planning forthcoming or supplementary exploration endeavours. Hence, all three landscape models were added to the Next Gen Analytics Data Package and the research team provided each project site with a specific recommendation of which model to use.

### 3.2.2 Comparison of Landscape Models to input layers

While the unsupervised landscape clustering approach separates an area's regolith into domains based on the information provided by each input layer at every pixel, it does not label these clusters. Hence, the model does not assign a particular regolith type to a cluster and

only two sites were physically ground-truth tested. Therefore, the landscape clusters, or landscape types, are considered proxies for the physical regolith types. However, each landscape model produced for this project has been validated by comparison against the model input layers and satellite imagery. In general, the spatial representation of the modelled landscape clusters corresponded well with broad landscape settings such as outcrops or residual hill areas, active fluvial channels, river terraces, deltas and large linear dune fields for all sites. Examples from two sites in Queensland (Henne et al. 2023a; Figure 24) and in the Northern Territory (Henne et al. 2022a; Figure 25 and Figure 26) are shown below. These are easily identified and interpreted and often compare well where detailed regolith maps are available.

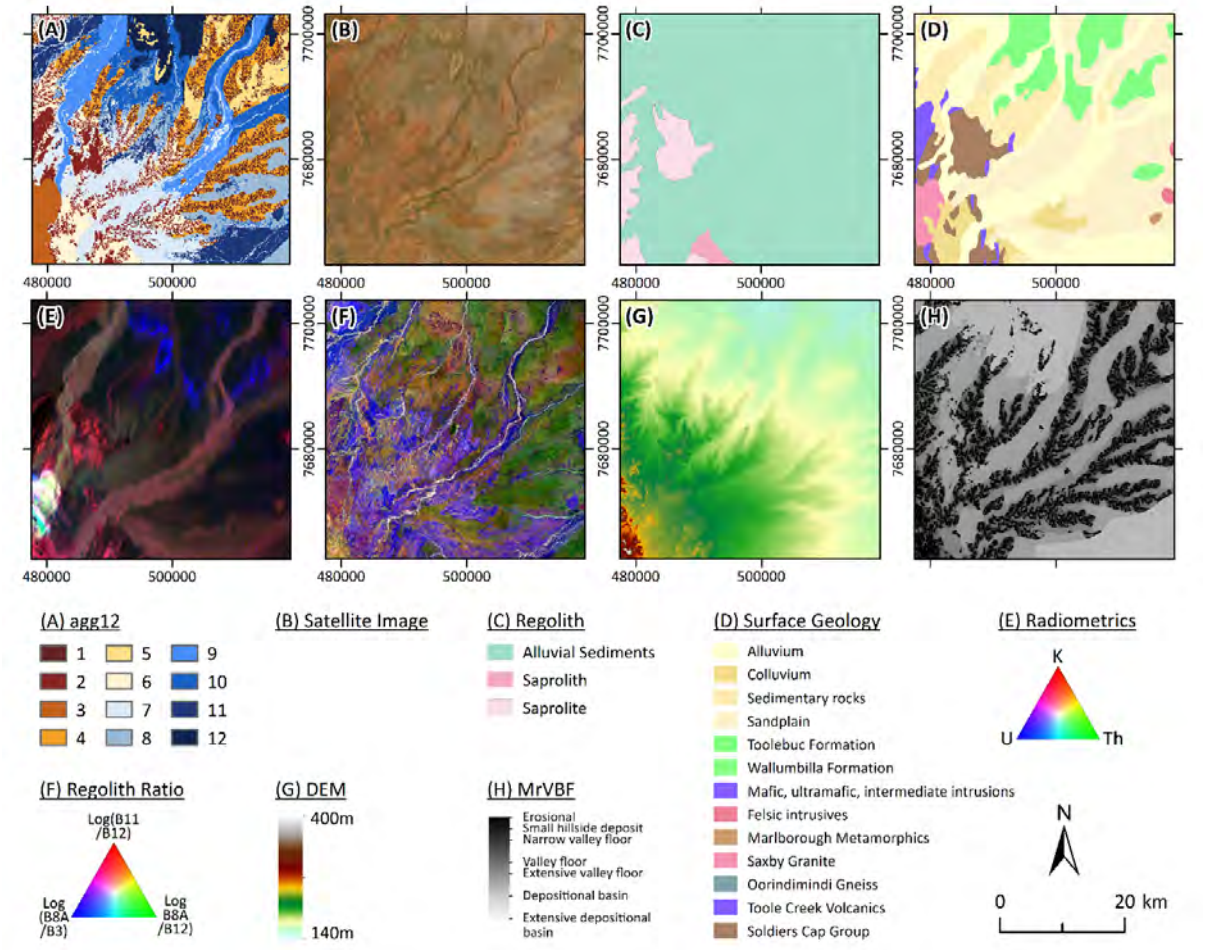



Figure 24: Example of a proxy regolith landscape cluster map generated in the Mt Isa region in QLD, input layers and comparison to traditional geological maps over the general Jericho region. (A) Landscape clusters derived via agglomerative clustering with twelve clusters. (B) True-colour satellite image. (C) Regolith geology. (D) Surface geology. (E) Radiometric data (indication of parent materials). (F) projection of spectral band regolith ratio (indication of regolith material). (G) Digital Elevation model (indication of landscape position). (H) MrVBF (indication of depth of transported cover). This figure was reproduced from Henne et al. 2023 with permission.

Some features, however, may be less distinct, e.g., erosional mid slopes trending into foot slopes, toe slopes, and pediplains/sheetwash plains may not display distinct property changes that the model can cluster with sharp boundaries and may result in different landscape proxies partly corresponding to similar settings (e.g. NE section of Figure 25 is a broad toeslope/pediplain trending into aeolian dunes). It is important to note that the clusters are heterogeneous, similar to traditional regolith map units and may be mixtures near transitions

such as residual soils trending into colluvial soils (erosional side slopes). In addition, the inclusion of the relative depth of cover (MrVBF) in the models, while adding valuable additional information, may be unfamiliar to the traditional regolith geologist when assessing maps. The MrVBF also determines the colour classes of the modelled landscape clusters and have been assigned based on the mean value of the MrVBF for each cluster. The lowest mean MrVBF values are assigned dark red-brown, and often indicate outcrop and residual upland settings. The highest mean MrVBF values are assigned dark blue, and often indicate areas of deep transported cover. However, the scale is relative (based on the lowest and highest mean MrVBF values in the project area). For example, the presence of elevated lowlands (e.g., depressions in higher elevation regions of a survey) or outcrops in otherwise deep cover (e.g., silcrete outcrops on salt lakes) can influence the order of landscape colours. The project outputs typically show red tones as outcrop/sub-crop/residual cover grading into orange/yellows/light blues for side-slopes, channel deposits, shallow cover and the darker the blue the deeper the cover. It is important to recognise that each site is processed individually, and landscape clusters/colours are not comparable to other sites.

The following shows several examples of the overall positive alignment of landscape clusters with landscape features and notes some of the constraints. Where available, regolith and surface geology maps were also used for comparison and validation. However, in many cases, these maps were at a lower resolution than the landscape cluster maps (e.g., compare Figure 24a and c).

In the example from the MacDonnell Ranges in the Northern Territory (Henne et al. 2022) the dark red-brown in Figure 25A are elevated ranges of outcrop/subcrop. The light yellow and mid blue classes were separated using the satellite image (Figure 25B) that shows a pronounced outcrop, the DEM (Figure 25E) and MrVBF (Figure 25F) indicate that there is a difference in landscape position and depth of cover between dark red-brown (outcrop) and orange, hence the orange cluster is a side slope position. Further downslope, the light yellow and blue colours show a change in regolith materials (regolith ratio map; Figure 25H) and a change in the radiometric signal (Figure 25G) as well as the depth of cover (Figure 25F). Both light yellow and mid blue clusters are mixed shallow cover, the main distinction being the difference in depth of cover that is important for interpretation. The light grey-blue colour is identified as a landscape feature akin to fluvial channels. The blue clusters remaining are mainly distinguished by the depth of cover.

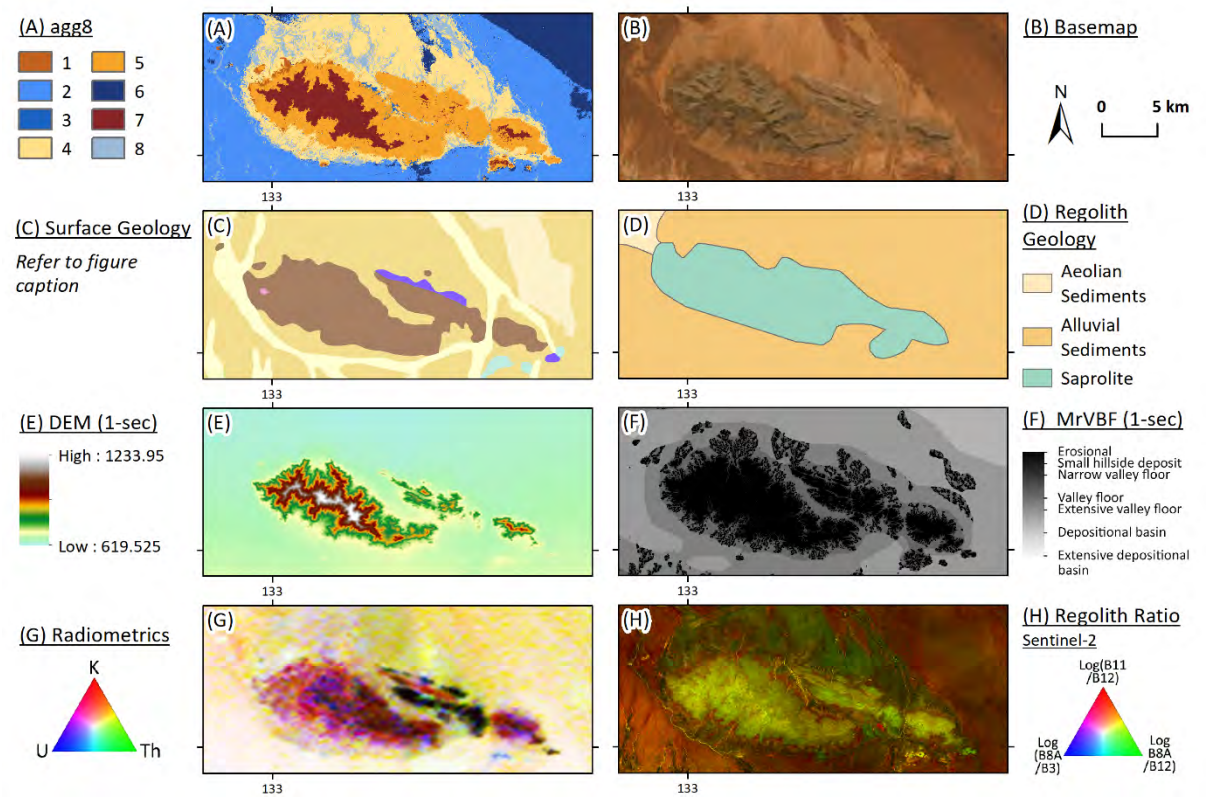



Figure 25. Assessment of landscape clusters indicative of sideslope material produced by agglomerative clustering (agg8) in (A) compared to satellite imagery in (B) and traditional map outputs (surface geology in (C) and regolith geology in (D)). Landscape clusters were produced from the input layers DEM (E), MrVBF (F), radiometric data (G) and regolith ratio (H).

Another interesting example is the mid blue cluster wedged between side slope material (orange) shedding from outcrops/subcrops (dark red-brown) to the north and south (Figure 26A). This is a relatively lower landscape position (Figure 26E) with deeper cover (Figure 26F) and a distinct colour in the regolith ratio (Figure 26H; indicating a change in regolith materials). The satellite imagery shows a large sand dune. This was further confirmed by looking for the same features in other parts of the area, hence the mid blue cluster is denoted as aeolian cover.

A key difference between the Next Gen Analytics regolith proxies and a typical regolith landform type surface map is that the M0462a project was designed for surface geochemical exploration and not just surface landscape classification. For surface exploration it is essential to evaluate anomalies distinctly based on the thickness of cover and the MrVBF provides this difference. However, this means that the machine learning derived maps for some areas look different to what a traditional landscape map might show. This is particularly pronounced for alluvial channels that switch colour classes on the ML derived map. This is because the overall landscape cover thickness has changed with the MrVBF. Hence, although the samples may be in a similar geomorphological setting (fluvial channel) the depth of cover below has changed. In the context of surface geochemistry this enables better discrimination and consideration of the geochemical data, since the elemental signatures in deeper cover are likely to be weaker or perhaps unlikely to be detected with surface geochemical techniques.

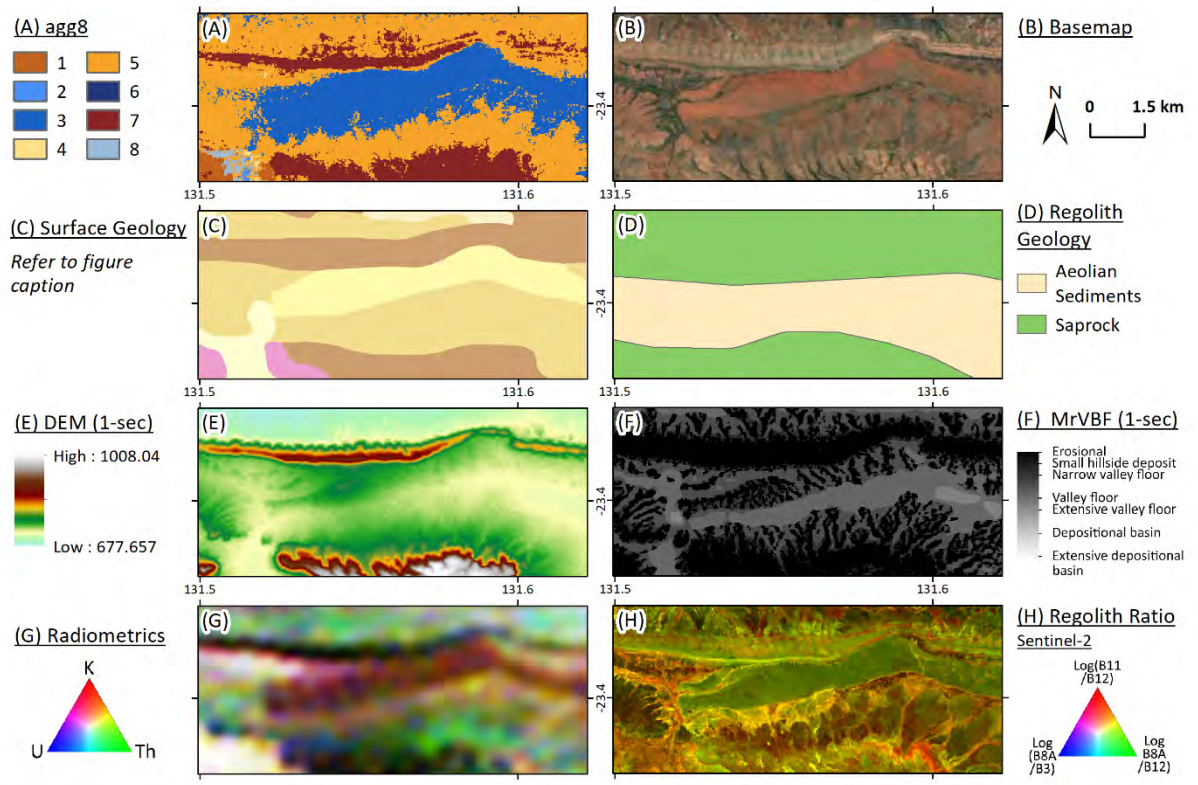



Figure 26. Assessment of landscape clusters as proxy for aeolian cover produced by agglomerative clustering (agg8) in (A) on the example of a sand dune, compared to satellite imagery in (B) and traditional map outputs (surface geology in (C) and regolith geology in (D)). Landscape clusters were produced from the input layers DEM (E), MrVBF (F), radiometric data (G) and regolith ratio (H). Clustering the data into proxy landscape classes inherently puts a hard boundary on the data even though, in many cases, there is a gradual change and the exact location of the boundaries is somewhat arbitrary. This is no different to other geological maps, although less biased than human interpretation, and samples near cluster boundaries should be considered when using the outputs.

### 3.2.3 Interpreting ("naming") machine learning derived proxy regolith landscape maps and comparison to existing map products

The ML-derived landscape map is not amenable to description using concise nomenclature that would fit neatly into available regolith classification schemes. This may be an initial challenge for human interpretation. However, 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). 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 Sentinel-2 band ratio signals (related to clay and iron oxide content) this should be considered as to its effect on metal mobility. This is demonstrated in the below example from the NW Yilgarn Craton in Western Australia (for more details on this model refer to Henne et al. 2023b). This site was chosen as an ideal comparison of the machine learning derived landscape model with a detailed, high-quality regolith map and we chose eight landscape clusters for this model to compare to the eight regolith types available in the simplified regolith landform map (de Souza Kovacs and Jakica 2021) of this area. A simple overview of likely materials and main features

observed for each cluster were derived via comparison with input layers, satellite imagery and the available regolith and surface geology maps and are noted in parentheses in the legend in Figure 27A.

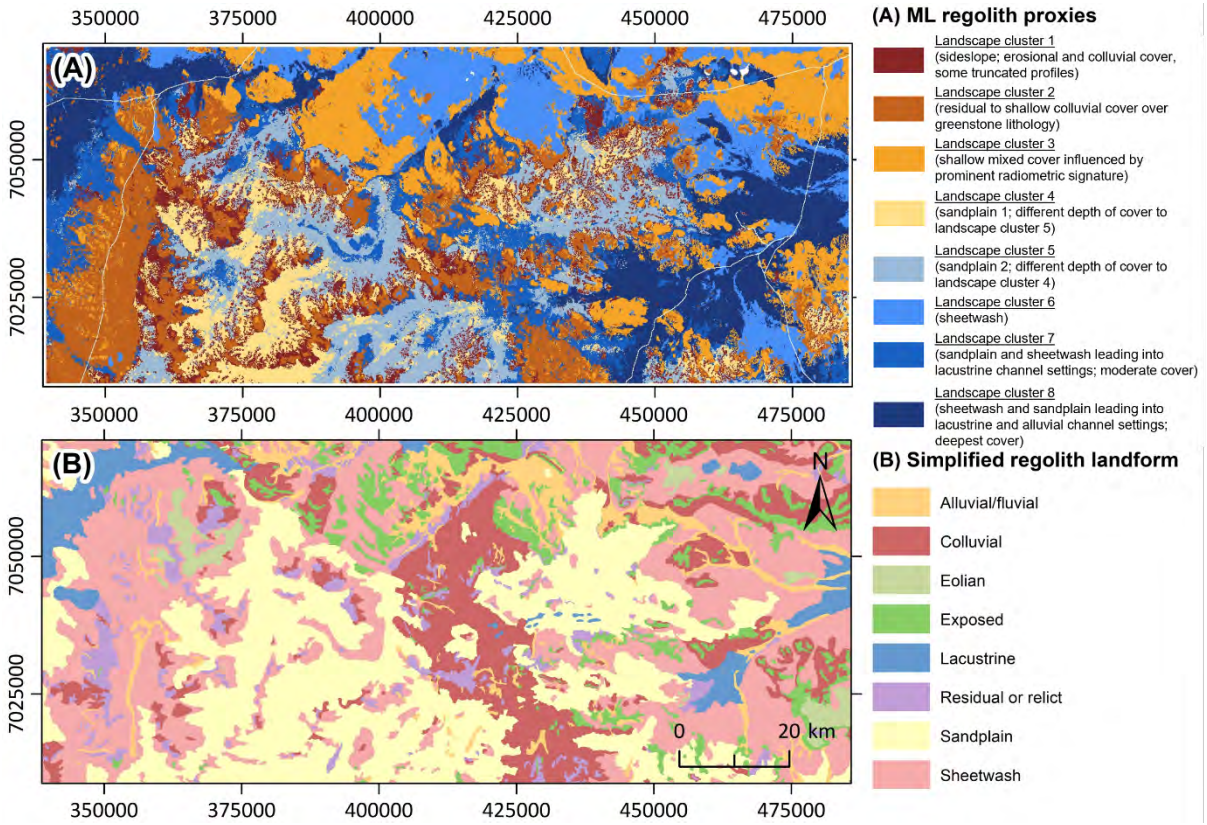



Figure 27. Comparison of human characterisation and machine-learned characterisation of regolith over an area in the NW Yilgarn Craton. (A) Machine learned proxy-landform map based on remotely sensed spatial features. (B) Simplified regolith landform map (de Souza Kovacs and Jakica 2021) derived via human interpretation. The ML 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. (Reproduced with permission from Henne et al. 2023)

Comparing the eight machine learned landscape clusters over the Kingston project site (Figure 27A) to the eight regolith landforms on the publicly available map (Figure 27B) indicates that the ML approach has defined clusters with similar spatial distributions to regolith landforms. For example, landscape cluster 5 (light blue; Figure 27A) has a similar distribution to what is mapped as sandplain in the available regolith landform map in the vicinity of the north-western Yilgarn Craton granites (Figure 27B). This is expected 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, some major differences can be observed. This is partly due to the coarser resolution of this example (100 m pixels), the lack of smoothing in the machine learned approach, and the choice of input layers. For example, in the centre-left of the project area of Figure 27 the regolith is mapped as sandplain material in the simplified regolith landform product (light yellow Figure 28B), whereas the machine learning approach has assigned two clusters in this setting (light blue and light yellow in Figure 28A). 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; Figure 28D), nor the material composition related to clay and iron oxide content (relatively more blue vs. relatively more

yellow colours in Figure 28C). Topographic (DEM, Figure 28E) and radiometric (Figure 28F) information was likely used for 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 likely to iron oxide and clay content (Sentinel-2 derived regolith band ratios).

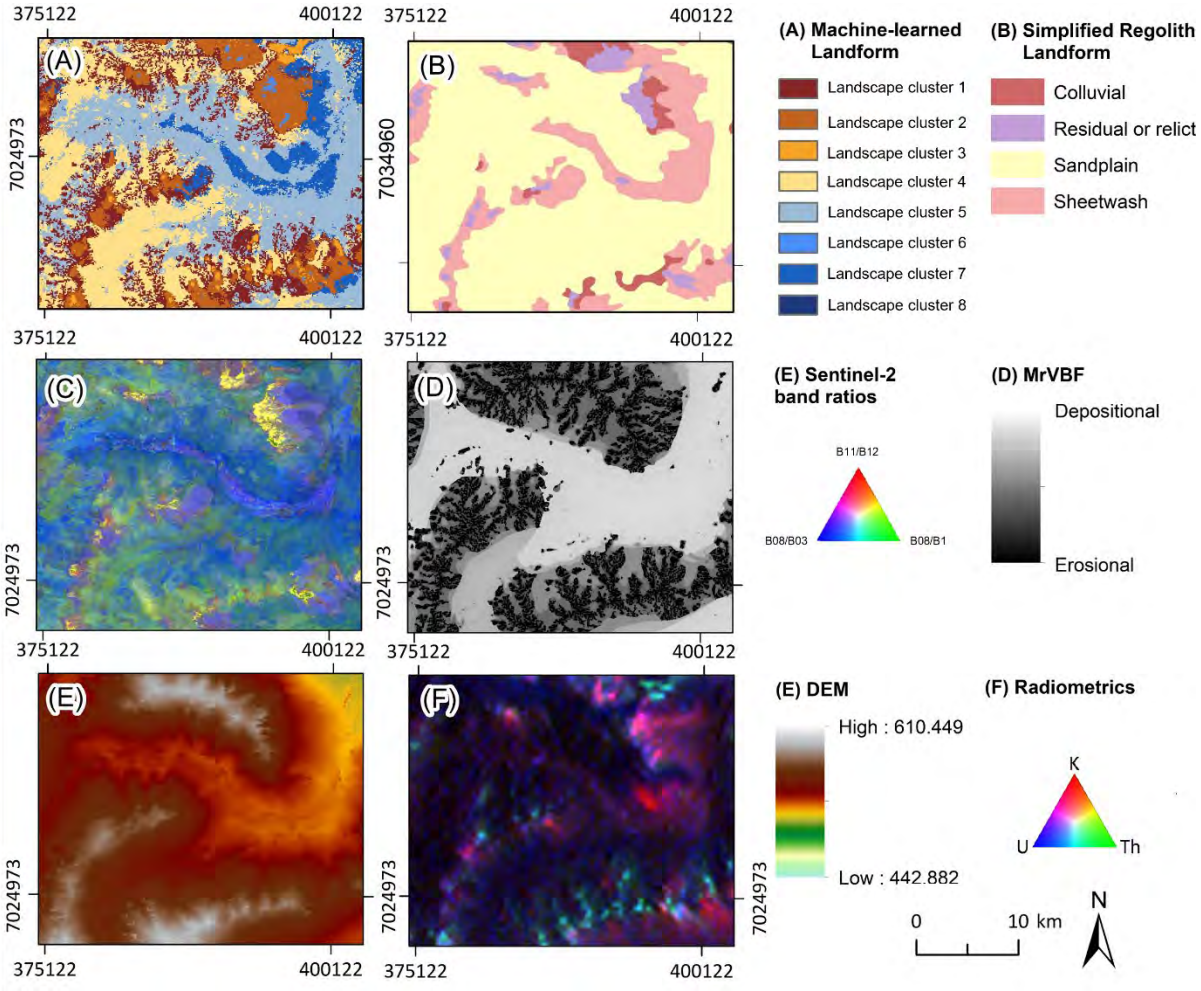



Figure 28. Comparison of machine learned model to regolith landform map and the parameters used to generate the model.

For a ML derived landscape model of such a large area in the NW Yilgarn Craton (Figure 27) more clusters are often appropriate, but 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). For example, more in-depth information in the publicly available regolith map product takes the underlying bedrock geology into account. While this is a valuable addition to interpret soil geochemical results in in situ regolith settings, if there is no connection of the regolith material to bedrock geology, such an interpretation could lead to misleading results. For example, the sandplain unit (light yellow; Figure 27B and Figure 28B) in the simplified regolith landform map covers sandplain material over both the Yilgarn Craton and the Earaheedy Basin (residual and aeolian sand, dominated by undulating sandplain and aeolian dunes; McGuiness and Pye, 2000). However, the more detailed regolith map (Figure 29B) differentiates this material based on the underlying bedrock. By using this output to interpret

soil geochemical results, an assumption is made that there is a genetic relationship between sandplains over the Yilgarn Plateau Physiographic Province (YPP) and the Pilbara Physiographic Province (PIP; Figure 29B). Radiometric data is often referred to as an indication of parent material from underlying bedrock. However, radiometric data is derived from the top 30 cm of the Earth's surface (IAEA 2003). In areas of transported cover, this feature layer is therefore related to surface regolith material composition. In the example, radiometric data (Figure 29D) indicates materials of similar composition, implying that parent material (not bedrock, but rather material providence) is relatively similar between these two units. The model instead has drawn out differences in depth of cover (distance to bedrock; Figure 29C) and divided the sandplain materials over the granites into two units (yellow and light blue; Figure 28; Figure 29A). Sentinel-2 derived regolith band ratios (Figure 29E) support similar patterns. It is important when assessing geochemical data, that the regolith is grouped 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.

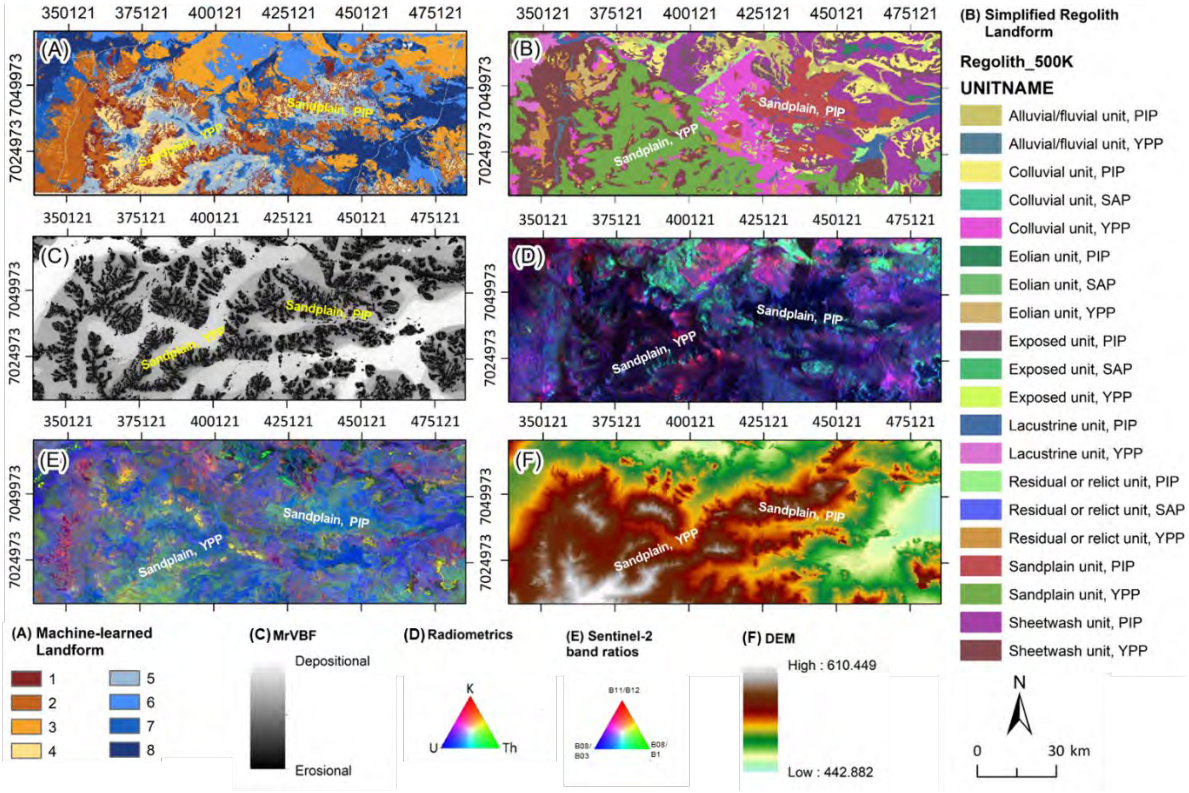



Figure 29: Comparison of machine-learned model to regolith unit names (including information such as underlying bedrock geology) model.

In the case study in the NE Yilgarn, a host of information on landform and regolith materials has been collected from remotely sensed and physical observations (Figure 30). 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 commonly 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 likely to actually

have been located over granites. The ML workflow aims to reduce the influence of human bias and error and provides a more consistent and objective landscape model.

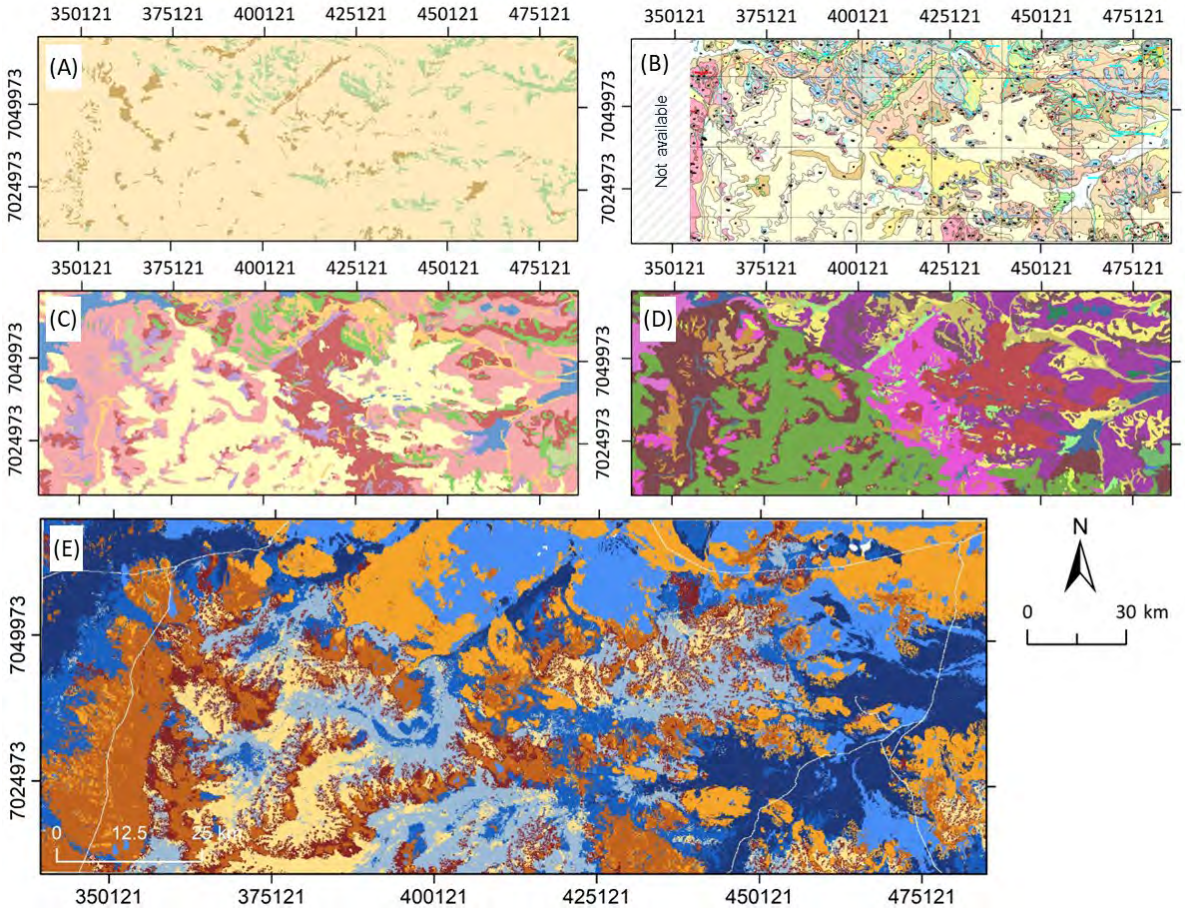



Figure 30: Various human-derived models and a spatial feature layer-based model. (A) Looking for outcrops (RED scheme; 1:100 000 regolith geology regimes of Western Australia; © State of Western Australia (Department of Mines, Industry Regulation and Safety, 2022). (B) Physical landscape characteristics, including geochemistry (C) Identifying regolith landforms (Extended RED scheme; Pye et al. 2000) (D) Extended RED including underlying bedrock geology (1:500 00 State Regolith Geology; © State of Western Australia (Department of Mines, Industry Regulation and Safety, 2022). (E) Machine-learned regolith proxy units for greenfields exploration. See Figure 27 and Figure 29 for legends.

### 3.2.4 Relationship of machine learned landscape clusters to soil properties

Two sites in Western Australia and New South Wales, respectively, were chosen to validate the mapped regolith materials of the machine learning models against real world regolith settings in a qualitative manner by walking the ground, recording observations on elevation, ground cover, vegetation type, and soil colour. The Western Australian site was more thoroughly investigated with the goal to establish the relationships between machine learned landscape clusters, input layers and soil properties (e.g., Figure 31). This research was part of a separate student internship research project using the Ultrafine workflow at a non-industry site and will therefore be published separately. However, a short summary of findings is provided below.

An area of 2000 km$^2$ was modelled as per the workflow described above and the output with eight landscape clusters was chosen. A machine learning approach (conditional Latin Hypercube Sampling; Minasny et al. 2006) was used to distribute 40 sampling sites (and 8

duplicate sites) across the eight landscape clusters with the aim of best capturing the variety of soil properties within each modelled unit (landscape cluster). Field observations and physical soil samples (soil cores of 30 cm length) were collected. Cores were split into 3 sections each (0 - 5 cm, 5 - 15 cm and 15 - 30cm), each split was homogenised and analysed using pXRF for major element chemistry, handheld ASD and FTIR for spectral mineralogy and other physicochemical properties (including pH, electrical conductivity, and size fraction analysis above and below 2 mm).

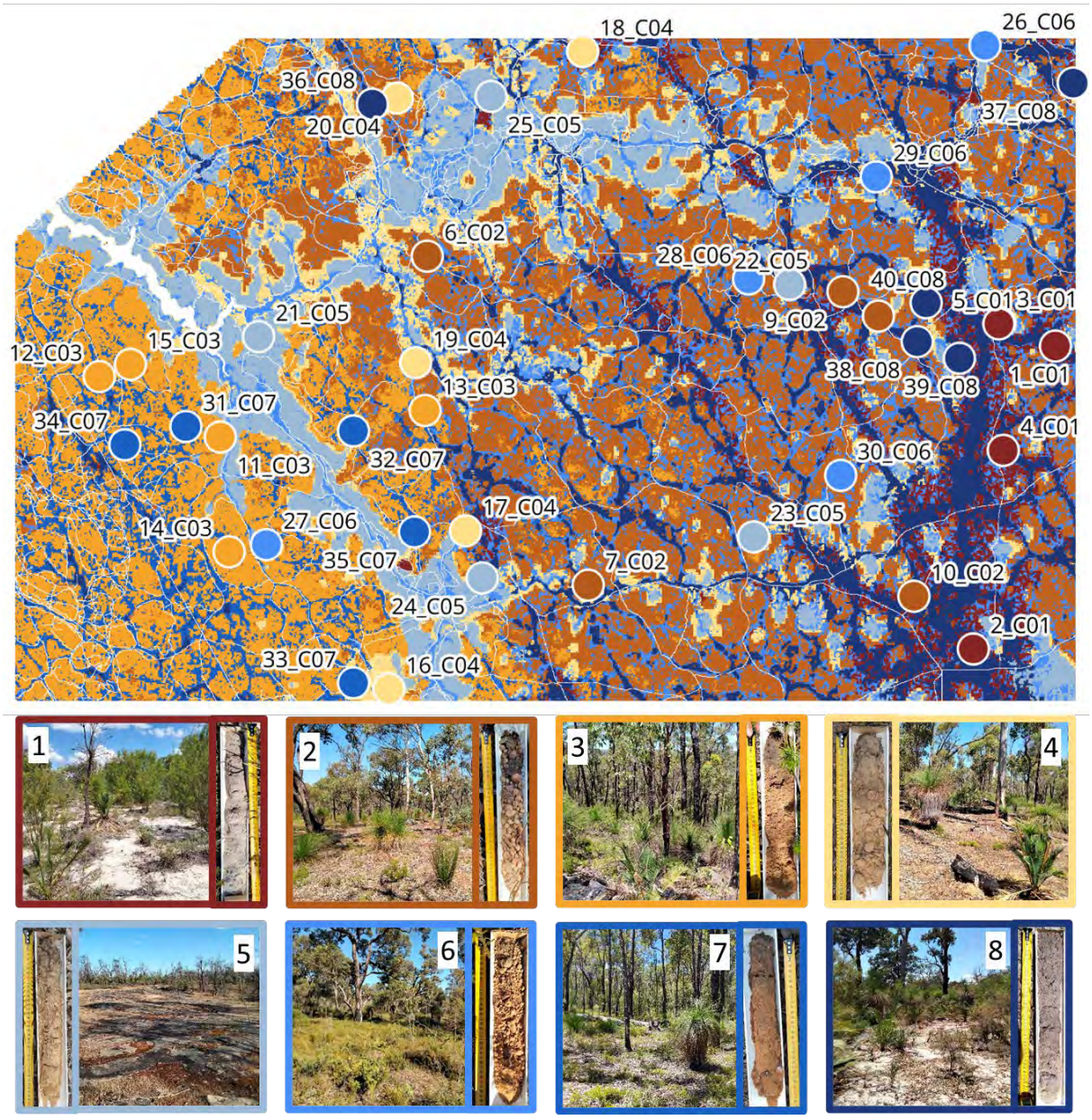



Figure 31: Distribution of soil cores in the study area (map) and representative photograph of landscapes and soil cores in each modelled landscape cluster. The labels (1 to 8) and colours refer to the eight modelled clusters and correspond to those in Figures 23 and 24.

In general, broad correlations between soil property changes in the physical samples and machine learned landscape clusters as well as the spatial feature layers used to create these models were observed (Figure 32 and Figure 33). Examples include similar trends in the distribution of relative Fe-oxide abundance in physical soil samples of each landscape cluster (measured via handheld ASD; Figure 32B) with the distribution of Sentinel-2 band ratio 8/3

(indicative of Fe-oxide composition; Figure 32A), as well as the relative K abundance (measured via pXRF on soil samples; Figure 32D) and the radiometric K percent distribution (from radiometric data; Figure 32C) for each landscape cluster.

Most landscape clusters showed distinct soil properties that differentiate landscape clusters from each other, including soil colour, soil particle sizing, presence or absence of pisoliths, soil carbon content (Figure 31) and major element concentrations. An exception were landscape clusters 1 and 8 which were both characterised by sandy soils with little organic carbon, similar particle size distribution and similar major geochemical composition. Sentinel-2 regolith ratios, radiometric data (Figure 32C) as well as landscape position (DEM, Figure 33B) also indicated that these landscape clusters are very similar in input layer properties. However, the relative depth of cover indicated by the MrVBF shows an appreciable difference in mean MrVBF values of pixels within these landscape classes (Figure 33). This provides essential additional information in a surface exploration context and differentiates the machine learning derived landscape maps from traditional regolith maps.

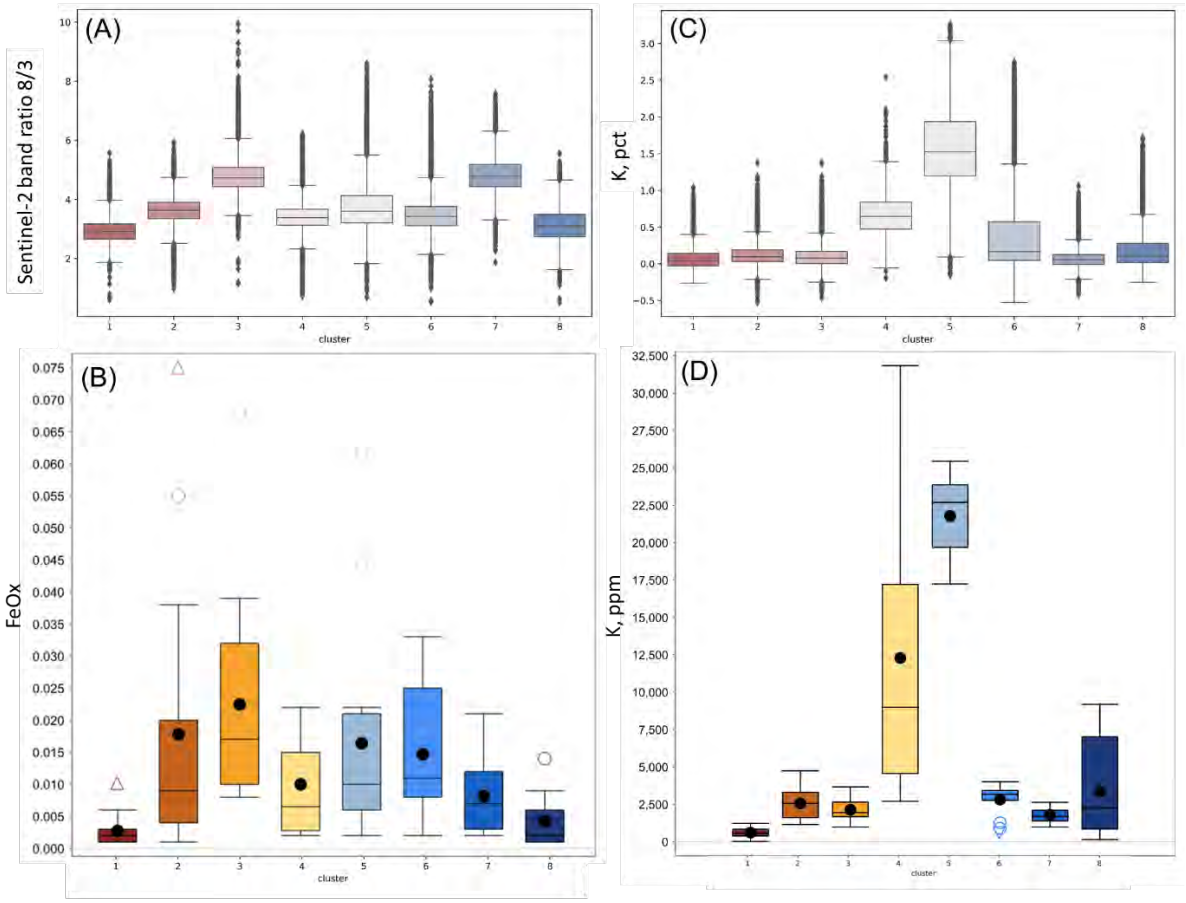



Figure 32: Remotely sensed and point-sample distribution of some properties of each of the eight modelled landscape clusters. (A) Remote-sensed Sentinel-2 band ratio 8/3. (B) Relative iron oxide abundance acquired via hand-held ADS on the <2 mm size fraction of 150 soil samples. (C) Remote-sensed radiometric data of percent K. (D) Potassium abundance acquired via pXRF on the <2 mm size fraction of 150 soil samples.

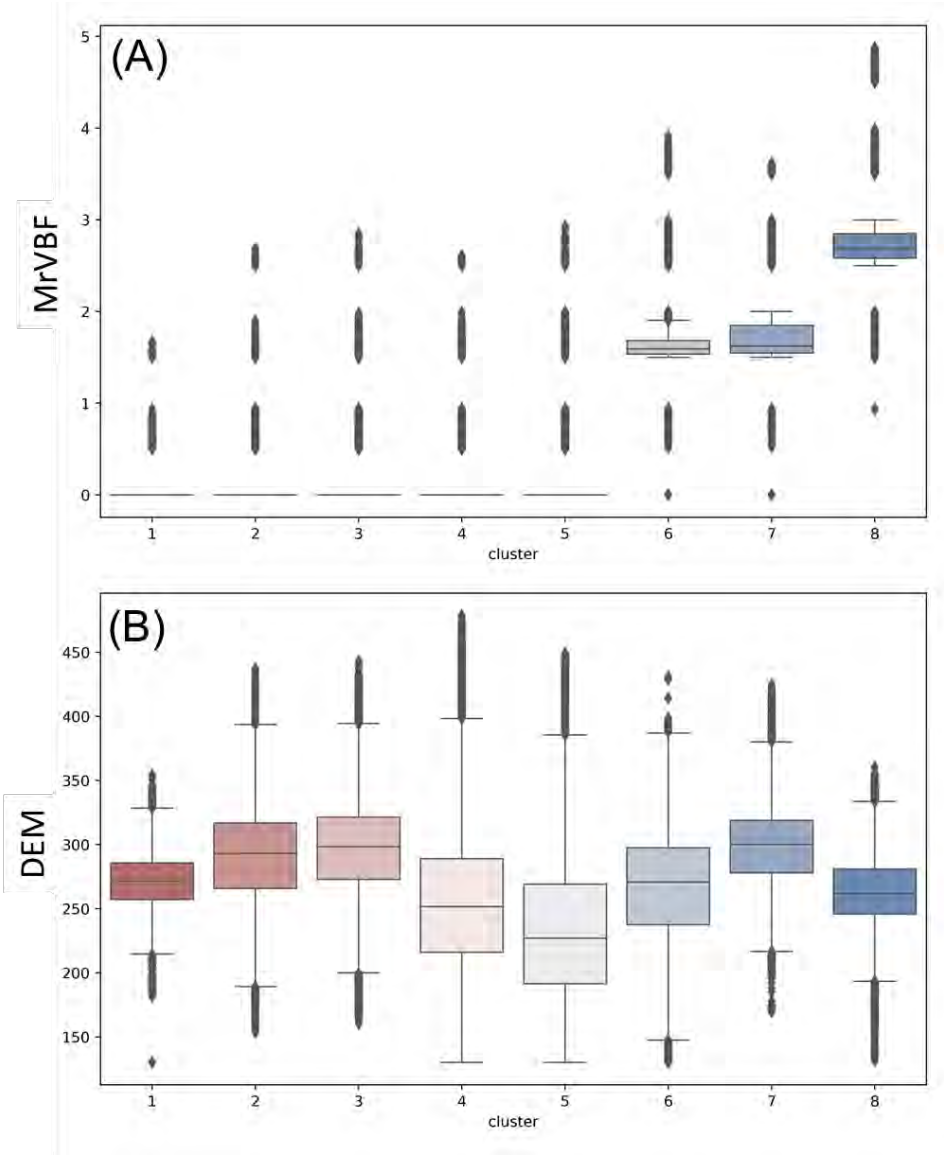



Figure 33: Distribution of MrVBF values (A) and elevation (B) for each modelled landscape cluster.

### 3.2.5 Assessing geochemistry in landscape context

Outliers in soil geochemical data are commonly calculated on all samples (the whole population) collected over an area of interest, regardless of their landscape context. However, the common “all samples” approach ignores the underlying processes that may affect metal dispersion. For example, high metal concentrations may be readily identifiable as outliers in a geochemical dataset where samples were collected over mineralisation with exposed outcrop or shallow residual materials. The same mineralisation would have a much weaker elemental signal in samples collected over moderately thick depositional landscape types such as a sand plain. It is therefore important to consider landscape context when evaluating geochemical outliers. Landscape types in a soil survey area are essentially considered sub-populations within the sample set that can be assessed separately using the regolith proxies generated by the ML models.

The below example of a geochemical survey in machine learned landscape context shows the Bi results of 270 samples over the Wagga Tank project area (Figure 34A; samples submitted by the Geological Survey of New South Wales; Henne et al., 2023b). This is a typical output

for a small landscape model (kmeans4; Figure 34) with few samples. The soil sample population is represented in four landscape clusters ((coloured boxes; Figure 34A). Traditional outliers calculated from all data as a single or whole sample population are shown for comparison in a white box on the left-hand side. Outliers are indicated as triangles to highlight potential anomalies within different landscape settings. For most of these landscape types, outliers (triangles) below the dashed line would be considered unremarkable (background concentration values) if evaluated as part of the whole data set. However, there are outliers from samples in landscape sub-populations in depositional settings (grey-blue and dark blue boxes). The dark brown and orange populations on the other hand were sampled in residual and/or erosional landscape settings and are therefore, as expected, well represented by outliers in the overall sample population (compare white box/triangles with brown and orange triangles above the dashed line). Outliers in the light and dark blue, depositional landscape settings form potential targets in cover and those in the NE represent anomalies formed over or down slope of the Wagga Tank mineralisation highlighting the value of the approach.

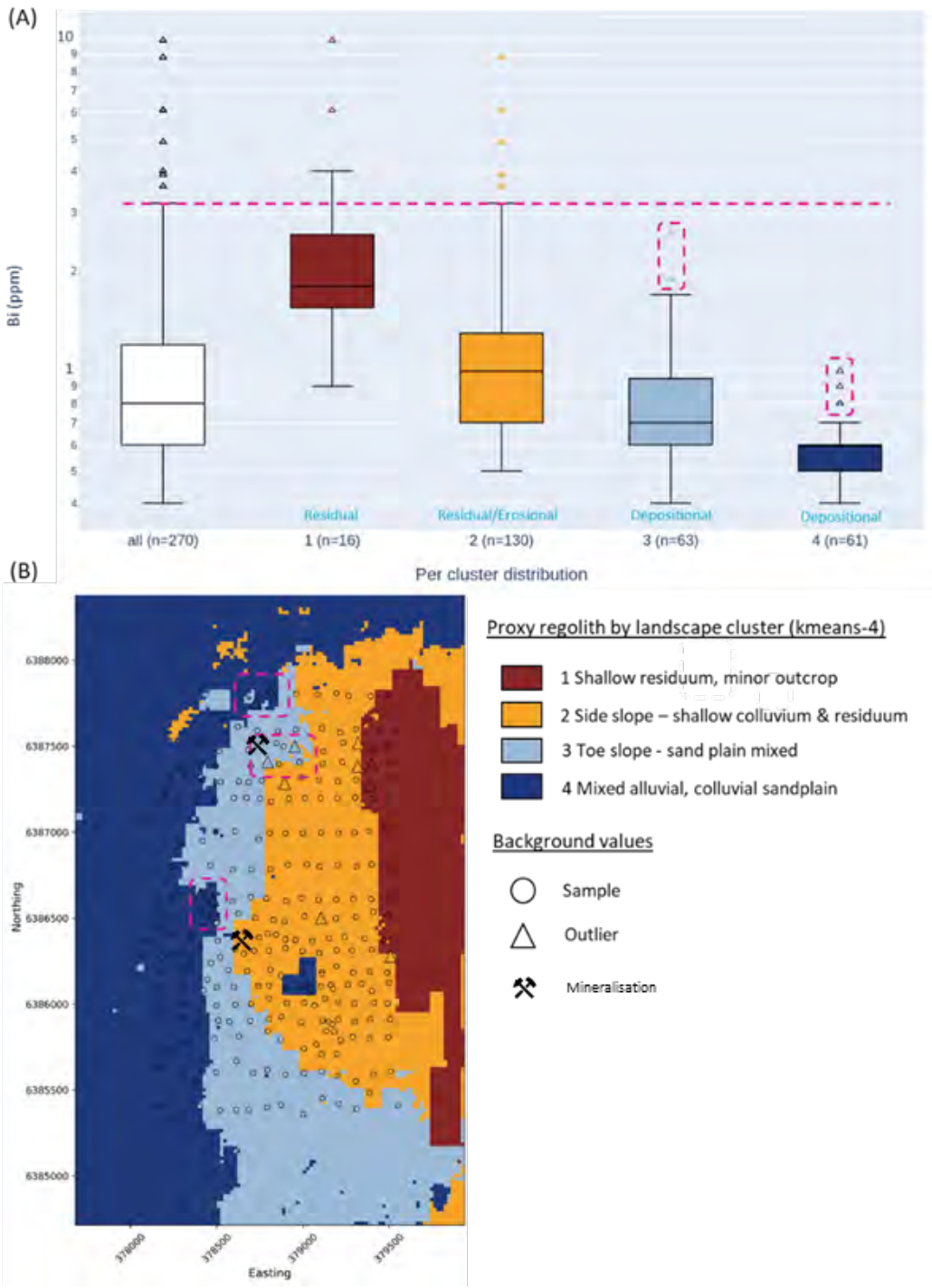



Figure 34:Example of boxplot and map from the Wagga Tank project site in NSW. (A) Boxplots for all Bi data (white box) and by landscape type (coloured boxes). The dashed line indicates the outliers for the whole sample population (1.5x the IQR). Easily observed soil anomalies are samples above the dashed horizontal line. Those shown below the dashed line would not be easily observed without the landscape context. (B) Spatial distribution of Bi outliers (triangles) by proxy regolith type. Outliers in dashed boxes correlate to outliers below the dashed horizontal line in (A) in depositional landscape settings.

While the workflow was designed to identify outliers in transported cover, the Next Gen Analytics workflow accommodates identification of anomalies without landscape settings by also generating outliers for the whole sample population. Hence, original, elevated elemental signals are not lost. The example below is from a large survey area (approximately 15 km x 20 km) with approximately 5,000 samples. The outliers for Zn are calculated on the whole

dataset, with only two values displayed as outliers (Figure 35A). With the same data separated into 12 landscape clusters and outliers calculated, more outliers are identified (Figure 35B). A closer review of the north-eastern outlier (Figure 35 Inset A) as a potential target is essentially extended south and with more anomalous samples when viewed in landscape context, providing more confidence in the target area (Figure 35 Inset B).

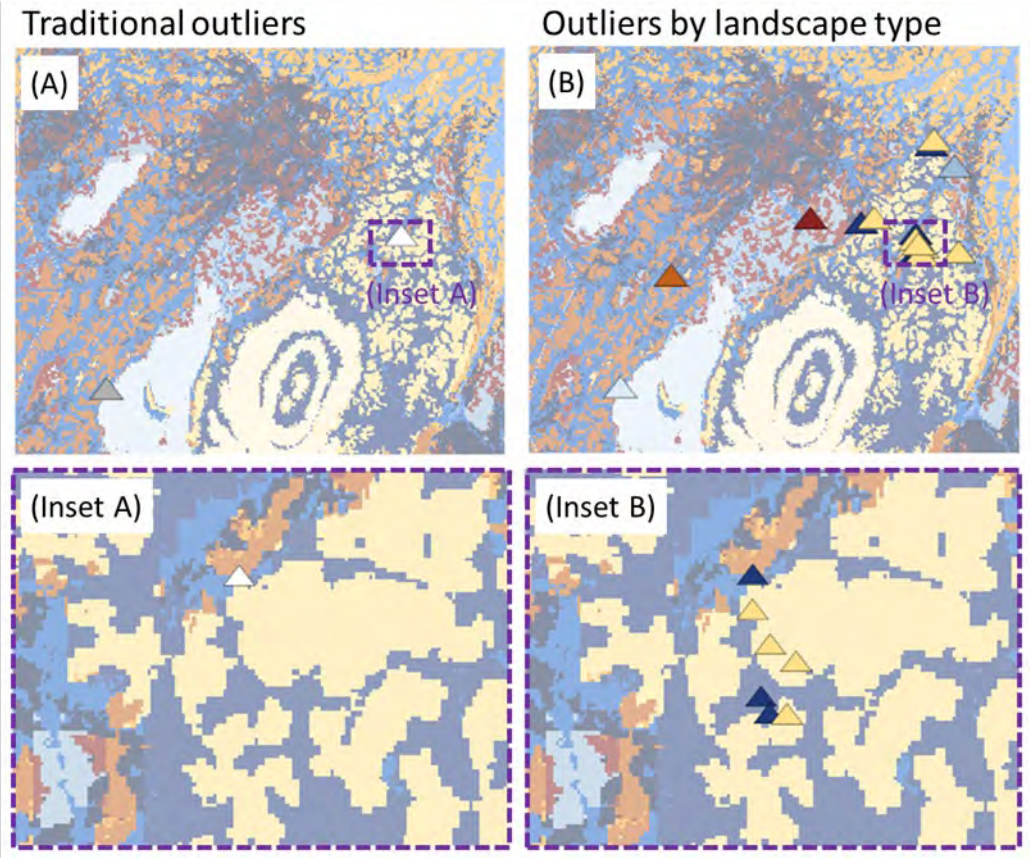



Figure 35. Example of the Next Gen Analytics outputs comparing anomalous Zn outliers with the whole data set (A) and the outliers by landscape type (B) at the broad site scale (top) and the more detailed target area (bottom).

### 3.2.6 Consideration of other spatial layer inputs for greenfields exploration settings

There are many spatial feature layers (both remotely sensed and human-interpreted) which may be useful to generate a proxy landscape type map, depending on site and commodity specific exploration needs. Hence, multiple spatial data feature layers were investigated as to their use in the unsupervised machine learning workflow during the initial testing phase of this research project. These included Advanced Spaceborne Thermal Emission and Reflection Radiometry (ASTER) products, clay content, organic carbon and pH (15-30cm) from the Terrestrial Ecosystem Research Network (TERN), Global 1-km Gridded Thickness of Soil, Regolith, and Sedimentary Deposit Layers, a variety of digital elevation models (SRTM-DEM, COP-DEM and FAB-DEM) and derivatives thereof (e.g., Multi-resolution Valley Bottom Flatness and weathering intensity), multiple Sentinel-2 products (e.g., regolith ratios, 754 and Newton-Boyle), airborne magnetic and gravity surveys, and surface and regolith geology maps produced by the state and territory geological surveys in Australia. Some local exploration company-acquired data was also considered in one case study.

The spatial data layers used in the final workflow (Table 8) 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) depth of transported cover (Multi-resolution Valley Bottom Flatness Index, or MrVBF), and (d) regolith material type (Sentinel-2 derived regolith band ratios), mainly relating to clay and iron oxide content. Important factors that were considered were:

(1) availability of high-quality data with consistent resolution across the entire Australian continent (e.g., ASTER data was excluded due to inconsistent flight lines and coverage, and this also precluded high-quality but spatially limited exploration company-acquired data);
(2) data is minimally interpreted or interpolated by humans (e.g., surface and regolith map products were excluded);
(3) data provides a direct measurement of soil properties (i.e., it is not an interpolated model from limited soil property measurements such as is the case for TERN data);
(4) data provided independent information (e.g., this precluded the use of weathering intensity - a derivative of the SRTM-DEM and radiometric data) to ensure the resulting landscape model was not unduly "skewed" towards one parameter)
(5) data provided information of the sampled medium (approximately up to 30 cm depth; e.g., this precluded airborne magnetic and gravity data).

The final workflow that generated proxy landscape maps for all project sites provides partially independent information concerning regolith composition and include (a) airborne radiometric data to indicate parent/source material, (b) a Digital Elevation Model to indicate general landscape position, and (c) Sentinel-2 derived regolith band ratios to indicate regolith material type (mainly relating to clay and iron oxide content). In addition, and as an exception to (3), (4) and (5) described above, the workflow also includes (d) the Multi-resolution Valley Bottom Flatness (MrVBF) Index as an indication of the relative depth of transported cover to provide further exploration relevant regolith context. While this product is a derivative of the SRTM-DEM produced via an algorithm developed by Gallant et al. 2002, the inclusion of this layer assumes that the detectable mobile metal concentration originating from source lithology at depth will be diluted with increasing depth of cover. By incorporating the relative depth of cover into the unsupervised landscape model residual, erosional and depositional environments may become more apparent, and the dilutionary effect in depositional settings with transported cover may be accounted for, revealing potential additional targets in these depositional settings. This is particularly valuable in areas of sheetwash plains and sand plains and where cover extends across paleochannels. While in areas of residual cover, high-quality, detailed regolith and surface geology maps are commonly very similar to the ML derived proxy landscape maps (see examples previously), the inclusion of the MrVBF differentiates the machine learning derived proxy regolith maps from traditional regolith or surface geology maps which do not capture this information.

In addition to choosing appropriate input layers, several dimensionality reduction and clustering methods were trialled, and outputs tested against input layers, satellite imagery and during ground-truthing studies. A range of algorithms were assessed, including UMAP for dimensional reduction, and several clustering methods including agglomerative hierarchical clustering, k-means and HDBScan. There are multiple parameters that required workflow adjustments and decision making and, overall, over 1000 models were run to derive the final workflow. These models were also interrogated as to the effect of the presence (and masking)

of man-made features (such as open pits, dams, buildings and roads) and the scale of each modelled area. The resulting, final UltraFine+® Next Gen Analytics workflow described in the methods above uses the four publicly available spatial feature layers with man-made features removed to produce landscape clustering models, or proxy regolith types, which represent the regolith types of the respective project area.

## 3.3 Additional Next Gen Analytics Data Package Components

In addition to the refined soil geochemical results, landscape models, and interpretation of soil geochemical results in landscape context, the M0462a project team also coded several other first-pass interpretation outputs into the Next Gen Analytics workflow (Figure 2). These include soil texture analysis, regolith and exploration indices, principal component analysis, dispersion and source direction, and the visualisation of spectral mineralogy data. All of these were generated as part of the Next Gen Analytics Data Package for each project site. Demonstrated examples of each output are outlined below.

### 3.3.1 Regolith Ratios and Exploration Indices

The UltraFine+® method provides a high-quality, multi-element geochemical dataset with less values below the detection limit than many other commercially available analytical methods. With this wealth of information, geochemical data can be interrogated beyond the commodity of interest approach. The standard 52 elements as well as some of the additional soil properties (e.g., total organic carbon and iron-oxide abundance) can be used to assess pathfinder elements associated with the main commodity of interest. This usually requires some general knowledge of mineral associations, dispersion mechanisms and weathering processes within the exploration area. While pathfinder suites can vary dramatically for different deposits, even of the same mineralisation style and within the same mineral province, general mineralisation style-specific regolith ratios and indices have been documented in exploration geochemical literature over decades (e.g., Anand and Paine 2002; Dinis et al. 2020; Gu et al. 2002; Yang et al. 2004a; Yang et al. 2004b; Barnes et al. 2014; Barnes and Brand 1999; Barnes et al. 2004; Halley et al. 2015; Butt et al. 2005; Joyce 1984; Smith and Perdrix 1983). The Next Gen Analytics workflow leverages this knowledge by utilising the geochemical and spectral UltraFine+® analyses to generate a set of ratios and indices for each project site. These indices have been adjusted for available elemental compositions reported by the UltraFine+® method and indices (unlike ratios) are generated on centre-log ratio data. The indices and ratios are not intended to be an exhaustive list. Rather, the Next Gen Analytics workflow provides common regolith ratios and simplified indices as an integrated output of the machine learning workflow for a first-pass assessment, that can be interrogated spatially. The usefulness of a given pathfinder suite depends on ore-forming processes, mineral assemblages and subsequent weathering processes. Ore zonation at depth can further complicate the applicability of fixed pathfinder suites and it is recommended that in addition to using this data, other multi-element data combinations using local knowledge should be considered.

The Next Gen Analytics workflow calculates 12 ratios (Table 14) and 22 indices (

Table 15). Multi-element exploration indices range from gold pathfinder suites, porphyry and CHI6 (chalcophile-6) indices, to pegmatite (Figure 36A) and greenstone indices (Figure 36B). Regolith ratios, such as the Au/Mn ratio, can be used to normalise Au by Mn, as has been done when working with gold exploration in manganese-rich (metal-scavenging) soils (Noble 2012). Calcium ratios for Au in soils are potentially useful when sampling calcrete although

this association is commonly observed in profile and may not be closely associated in shallow surface soils (Lintern 2002).

Table 14. Regolith ratios calculated via the Next Gen Analytics workflow from multi-element geochemical as well as spectral analyses of ultrafine soil samples. Non CLR data = uses the elemental concentrations, e.g., ppm or instrument reported values.

| Regolith Ratio | Calculation | Comments |
|---|---|---|
| **Au/Ca** | Au divided by Ca | Non CLR data |
| **Au/Fe** | Au divided by Fe | Non CLR data |
| **Au/FePhase** | Au divided by Hem:Goeth | Non CLR data; FePhase relates to VNIR data; where FePhase is not detected there is no value and the ratio will not be calculated |
| **Au/OrgC** | Au divided by Organic Carbon | Non CLR data; relates to FTIR data; where TOC is not detected there is no value and the ratio will not be calculated |
| **Au/SSA** | Au divided by Specific Surface Area | Non CLR data; relates to sizing data |
| **Au/Mn** | Au divided by Mn | Non CLR data |
| **Ca/Mg** | Ca divided by Mg | Non CLR data |
| **Pt/Pd** | Pt divided by Pd | Non CLR data |
| **Rb/Sr** | Rb divided by Sr | Non CLR data |
| **Th/U** | Th divided by U | Non CLR data |
| **Ti/Zr** | Ti divided by Zr | Non CLR data |
| **Zr/Sc** | Zr divided by Sc | Non CLR data |

Table 15. Exploration indices calculated via the Next Gen Analytics workflow from multi-element geochemical analyses of ultrafine soil samples. CLR = Centred Log Ratio (compositional data transformation addressing issues with closure, presenting numerically simpler values which are potentially easier to work with).

| Exploration Indices | Calculation | Comments* |
|---|---|---|
| **CHI6** | As+Sb+Bi+Mo+Ag+Sn+W+Se | Addition of CLR values |
| **Gold1** | Au+Ag+As+W+Sb+Bi | Addition of CLR values |
| **Gold2** | Au+As+Bi+Te+S | Addition of CLR values |
| **Gold3** | Au+As+Sb | Addition of CLR values |
| **Greenstone** | Cr+V-U | Addition and subtraction of CLR values |
| **IOCG** | Cu+U+Au+Sm+Ce+La+Y+Sc | Addition of CLR values |
| **IOCG+Br** | Above + Br | Addition of CLR values; only if Br analysed |
| **Komatiite** | Ni-Cr | Subtraction of CLR values |
| **NiPGE** | Ni+Pt+Cu+Co+As+Te | Addition of CLR values |
| **NiPGE+Pd** | Above + Pd | Addition of CLR values; only if Pd analysed |
| **PEG4** | As+Sb+Sn+Nb+Ta | Addition of CLR values |
| **Pegmatite1** | Li+Ta+Cs+Nb | Addition of CLR values |
| **PolyMetal1** | Ag+Au+As+Sb+Mn+Ba+Te | Addition of CLR values |
| **PolyMetal2** | Au+Cu+Ag+Mo+Pb+Zn | Addition of CLR values |
| **PorphyryD** | Tl+Li+Sb | Addition of CLR values |
| **PorphyryL** | Mo+Cu+Ca-As-Mn-Zn-Co-K | Addition and subtraction of CLR values |
| **PorphyryP** | Bi+Te+Mo+Sn+Cu+W+Se | Addition of CLR values |
| **REE4** | Ce+La+Y+Sc | Addition of CLR values |
| **SedCu** | Cu+Zn+Pb+Bi+Co+Ni+As+Ag+Au | Addition of CLR values |
| **SnWGran** | Sn+W+Mo | Addition of CLR values |
| **SnWGran+I** | Above + I | Addition of CLR values; only if I analysed |
| **VMS_MVT** | Ag+Pb+Zn+Cd+Cu+Tl | Addition of CLR values |

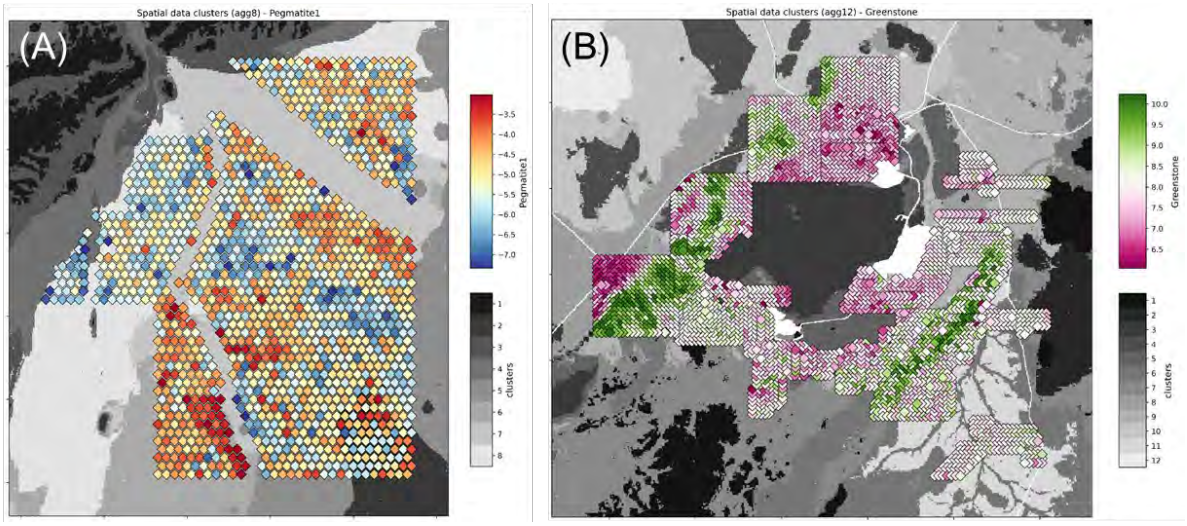



Figure 36. Examples of UltraFine+® Next Gen Analytics exploration indices outputs. Data is displayed over a proxy landscape map for this study area (landscape clusters are shaded in grey tones). (A) The “Pegmatite1” exploration index (Li+Ta+Cs+Nb) within a study area in the Pilbara. (B) The “Greenstone” exploration index (Cr+V-U) indicating lithological changes within a study area in the Pilbara.

### 3.3.2 Principal Component Analysis

Principal Component Analysis (PCA) is generated as part of the Next Gen Analytics package to identify patterns in the geochemical results using multi-element trends. Unlike the regolith ratios and exploration indices PCA identifies multi-element trends without geological knowledge bias. Principal Component Analysis reduces the variables (or dimensions) of a dataset. In the UltraFine+® Next Gen outputs, the results of all analysed elements (usually 52 dimensions) are described in five dimensions while preserving variation in the dataset. The PCA outputs allow for a rapid, first-pass identification of element association and potential exploration indices within the dataset.

The five principal components, PC0 through to PC4, generated by the workflow are displayed in a spider diagram which shows the loadings (weightings) of each element for each of these five principal components (Figure 37). This illustrates the general geochemical affinity of each principal component. The spatial distribution of each of the principal components by sample is also displayed. The explained variance of the principal components decreases from PC0 to PC4. Lower principal components that explain much of the variation in the data will, therefore, often pick out large-scale lithological variation as the major component(s). Principal components with exploration potential usually explain less variance in the data than the major elements related to geological influence. Hence, in many areas, PC2 and PC3 (3$^{rd}$ and 4$^{th}$ components) may represent the more relevant components in the context of mineral exploration (Figure 37). In some cases, signatures relating to mineralisation will constitute relatively low proportions of variance and thus will not be represented by principal components. Eigenvalues that document the proportion of variance in the data explained by each principal component are also provided in the Next Gen Analytics outputs. The further a component loading plots away from the 0 line (dashed black line in Figure 37), the greater the loading for the specific principal component.

Figure 37 shows an example of PCA on an UltraFine+® dataset located in the Yilgarn Craton (Western Australia) that may be useful for gold exploration at this site. While PC1 and PC2 are related to main mineralogy (association with greenstone belts), PC3 shows relatively high loadings for Ag, As, Au, Pd and W (purple diamonds connected by a purple line).

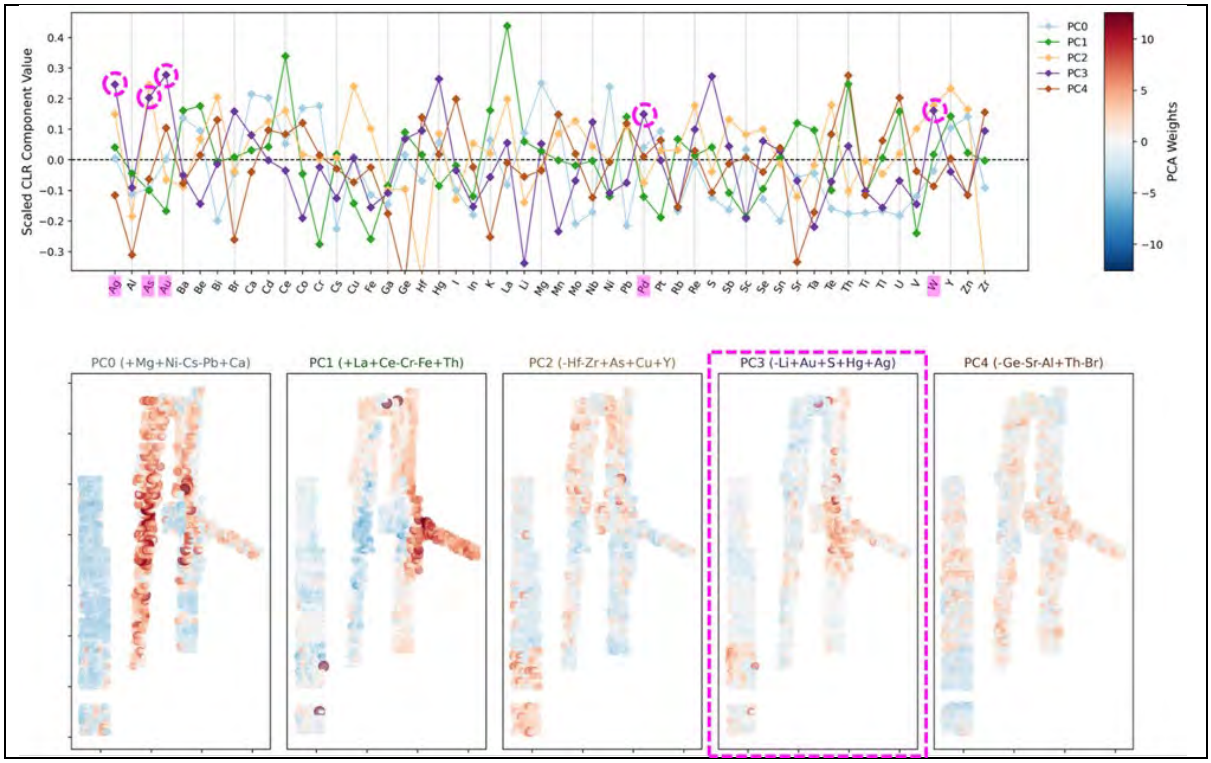



Figure 37. Spider diagram (top) for the 5 components showing positive and negative loadings for each element with the spatial projection of these principal component scores below. The elements noted in pink have strong positive loadings and present a good multielement pathfinder suite for gold exploration. The spatial distribution of each of the five principal components is weighted by both colour and symbol size where red indicates a positive component weight (association) and blue a negative component weight (association). The larger the symbols the stronger the association

Figure 38 shows an example of PCA at the Federation Project site in NSW (full report Henne et al. 2022b) where soil samples were collected over two known copper, lead, zinc, silver and gold prospects, Federation and Dominion. The first three principal components exhibit useful trends that can be related to geological processes. The first principal component (PC0; light blue) explains 46 % of the variability within the dataset and highlights shallow geology and climatic influence (slope aspect or elevation and matching the landscape proxy types/colours in the background). PC1 (green) explains 18 % of the variability within the dataset and is associated with As, Au, Ta and Tl. PC2 only explains 15 % of the variability within the dataset but exhibits exploration potential with the positive association of Ag, As, Cd, Cu, Hg, Ni, S, W and Zn. It therefore has positive loadings with some target and pathfinder elements such as Ag, Cu and Zn, but not with Au or Pb. PC2 (yellow) also is strongly influenced by Mn in the samples. In retrospect, PC1 highlights the Federation mineralisation (downslope from and immediately to the northwest of the Federation prospect) and PC2 highlights the Dominion prospect (positive PC2 scores coincide broadly with the Dominion prospect, but are more subdued around Federation; Figure 38).

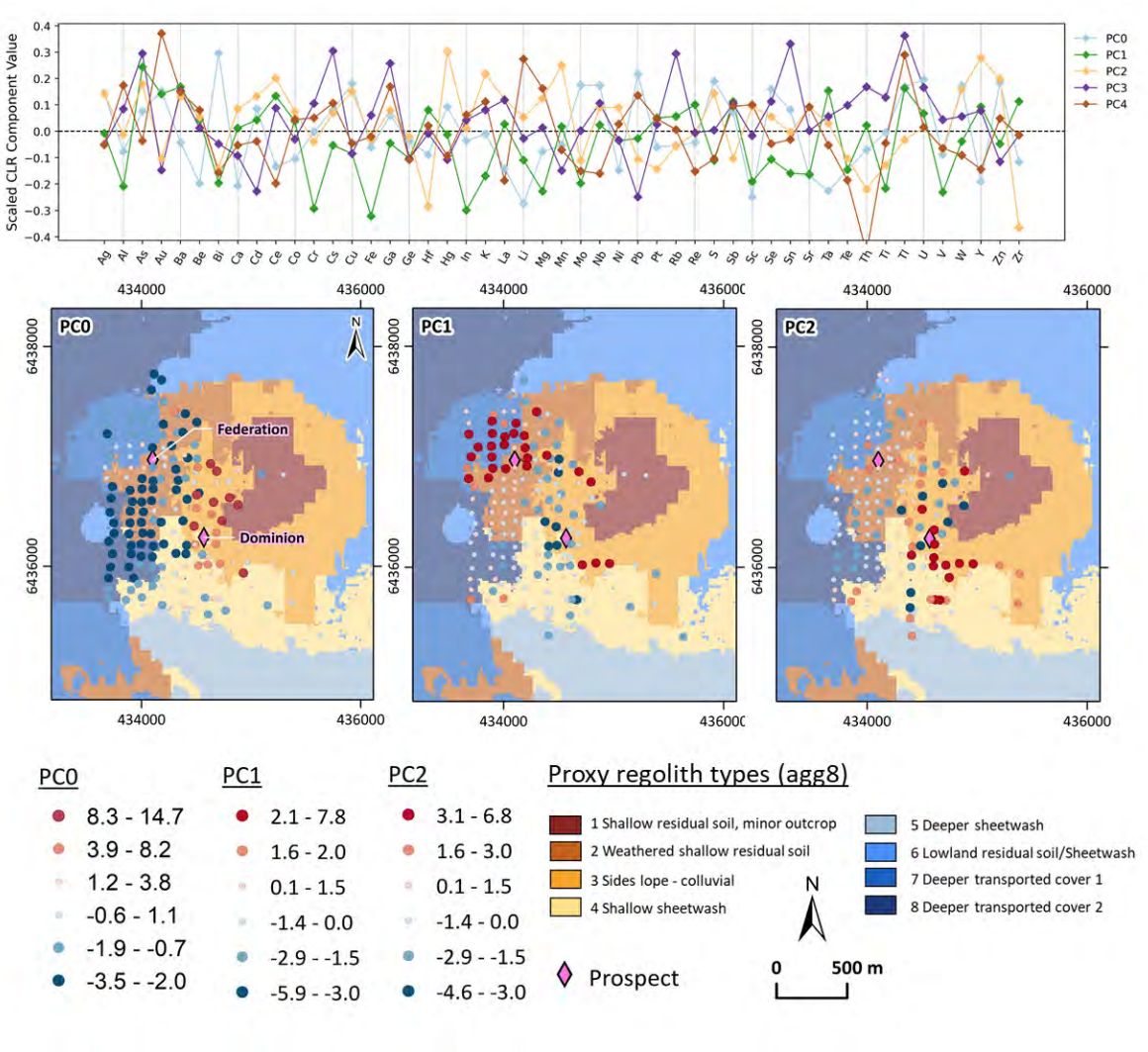



Figure 38: PCA outputs from the UltraFine+® Next Gen Analytics workflow over the Federation project area. Top is the spider plot of elemental loadings for each of the first five principal components. The further away an element plots from the 0 line, the greater the loading for (influence on) the specific principal component. Below is the spatial distribution of principal components weighted by both colour and symbol size (absolute magnitude). The colour red indicates a positive component weight (association); the colour blue indicates a negative component weight (association). The larger the symbols the stronger the association. From left to right boxes display the spatial distribution of principal component 0, 1, and 2 all plotted over the proxy regolith types modelled at this site.

### 3.3.3 Dispersion and Source Direction

The final addition to the Next Gen Analytics Data Package were source and dispersion directions. The source direction is derived for each sample with the arrow pointing up-slope and proportional to the steepness (indicating potentially greater dispersion distance) to enable the visualisation of source consideration for anomalous results (Figure 39A). The accuracy of the source direction depends on the accuracy of the GPS reading of a given soil sample and multiple adjacent sample points should be considered for interpreting the likely source direction of a geochemical anomaly in transported cover. The dispersion direction is calculated on a grid over each project area rather than on each sample point to give a more even overview of the landscape and likely general trends in element dispersion. The arrows point down-slope and are intended to give a first-glance overview of broader dispersion trends within

a project area (Figure 39B). This can be a useful tool for larger survey areas with transported cover and potential for infill sampling.

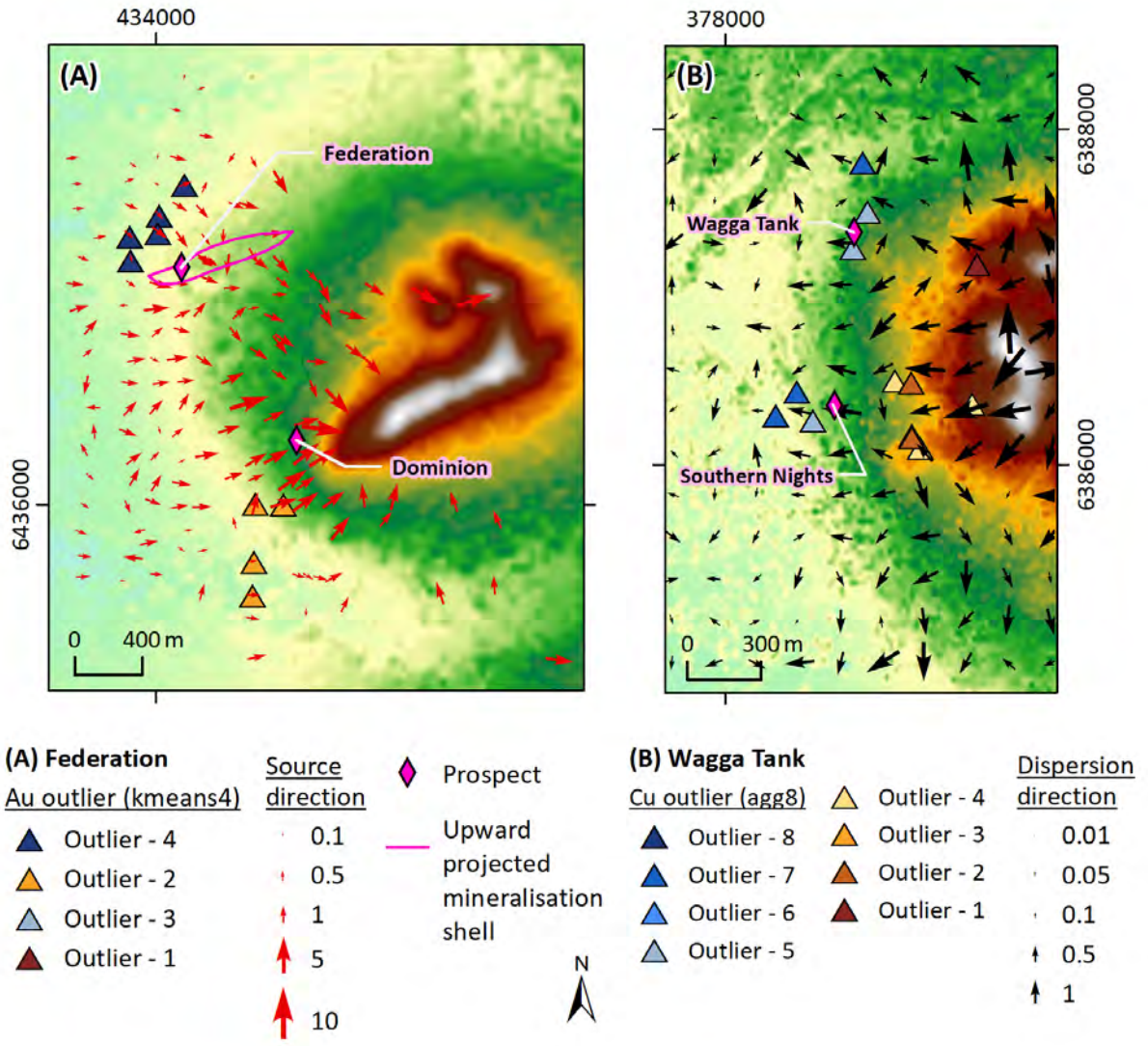



Figure 39: Source and dispersion direction over the Federation and Wagga Tank project areas. (A) Source direction of individual soil sample points and Au outliers by landscape type (kmeans4) within the Federation project area. Source direction is calculated from the DEM (background) but is dependent on accurate GPS readings. (B) Dispersion direction grid and Cu outliers by landscape type (agg8) within the Wagga Tank project area. Dispersion directions indicate broad scale trends. Arrows and numbers are proportional to the slope degree.

Figure 39 shows examples of source and dispersion direction for the Federation and Wagga Tank site where known deposits have been identified and the application of source/dispersion direction could be evaluated (full report Henne et al. 2022b). At both sites, most samples were collected in relatively flat landscape settings. The largest slope was 9.5 degrees within the Federation project area and 5.0 degrees within the Wagga Tank project area. Within the Federation project area, source directions of individual sample locations identified as Au outliers by landscape cluster (kmeans4) to the northwest of the Federation prospect and immediately to the south-southwest of the Dominion prospect point up-slope towards these prospects (Figure 39A).

For the Wagga Tank project area, dispersion directions were calculated on a 250 m grid. Copper outliers by landscape cluster (agg8) to the west of the Southern Nights prospect are

broadly situated in the direction of the overall dispersion trend (down-slope) (Figure 39B). Outliers are also apparent in the direction of dispersion from a radiometric potassium “anomaly” (elevated DEM displayed in white in Figure 39B), and dispersion directions around the Wagga Tank project area are influenced by the very small slope angle. It is important to note that the source and dispersion directions only take current slope directions into account and palaeo-relief was not considered.

### 3.3.4 Application of spectral data analysis to soil samples

The concentration of many metals in soil samples is related to their mineral composition, such as the abundance of iron oxides and clay minerals. VNIR-SWIR parameters can aid the identification of patterns, and in downgrading or upgrading target areas, as well as identifying areas of unique soil types to better interpret potential anomalies. Which parameters to use may be dependent on soil type, landscape context and deposit style.

Many trace metals are readily adsorbed to clay minerals and iron oxides. Given the high adsorption capacity of iron oxides, it is worth considering whether an anomaly of a metal of interest is present in iron oxide-rich or -poor soils and normalise the geochemical data accordingly to up- or downgrade targets. In some cases, adjusting for iron oxide abundance can tighten a geochemical anomaly for gold such as that shown in the northeast part of the soil survey in Figure 40.

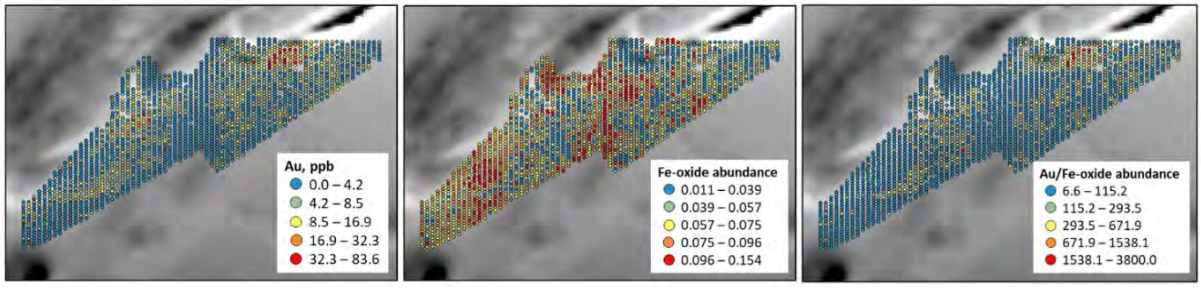



Figure 40: Au concentration, relative iron oxide abundance, and Au normalised by relative iron oxide abundance over gravity (grey scale). The normalisation has effectively reduced the Au target size over a very large survey area.

Other applications of adjusting exploration targets using spectral mineralogy can use the clay information. Clay minerals can be effective metal sinks in soils. Swelling clays (e.g., smectites) have negatively charged interlayer surfaces, which attracts hydrated cations. Non-swelling clays (e.g., kaolinite) do not attract metals into interlayer spaces. In a similar manner to the iron-oxide example previously, it is therefore worth considering whether an anomaly of a metal of interest is present in a swelling or non-swelling clay and normalise the data accordingly. In Figure 41, the data population was separated by Mineral Group 1 classes and Au was z-log normalised for each population individually to create a much stronger target in this section of an UltraFine+® soil survey.

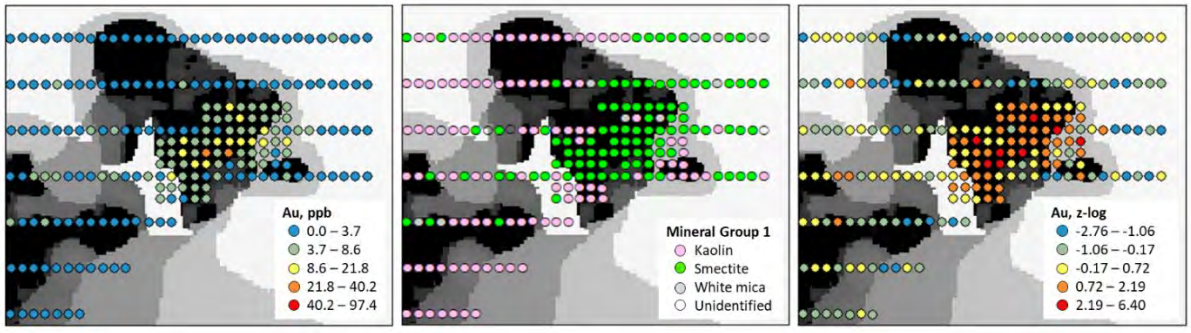



Figure 41: Au concentration (left), Mineral Group 1 (centre) and Au adjusted for Mineral Group 1 (z-log normalised and recombined, right) over depth of cover (grey scale). The recombined data has effectively highlighted the main Au target over the survey area.

While the initial UltraFine+® project work identified clays and iron oxides as adsorptive phases for mobile metals, the other major adsorptive phase is organic matter, hence the project requirement to measure this parameter. These “scavenging” phases such as clays, organic compounds and various oxides/oxyhydroxides may preferentially adsorb metals and may bias concentrations. The concentration of many metals may relate to the abundance of such phases in each soil sample. As a demonstration of how this might be applied in exploration, given the high adsorption capacity of organic matter, it is worth considering whether an anomaly of a metal of interest is present in organic-rich or -poor soils, and normalise the geochemical data accordingly to up- or downgrade targets. Adjusting Au concentrations for organic carbon content of the soil reduced the potential gold targets in one region of a soil survey (Figure 42). This site has not been drilled so it is inconclusive if this was effective, but it does provide options for explorers to test their targets and the robustness of the anomalies.

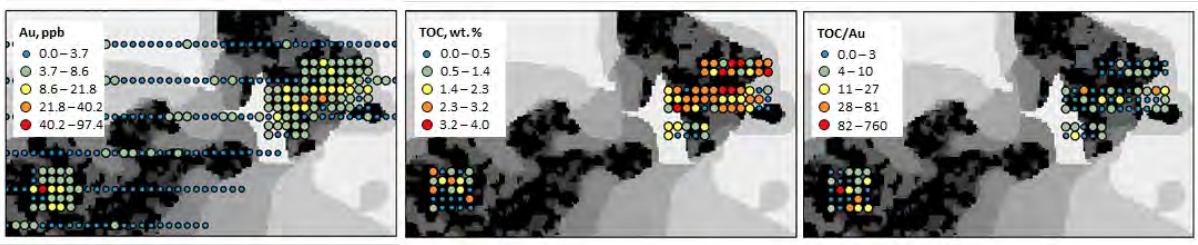



**Figure 42**: Au concentration, total organic carbon content, and Au normalised by total organic carbon content over depth of cover (grey scale). The normalised data has effectively reduced the Au target size in the two main areas of interest.

## 3.4 Additional Research Project Components

The following aspects of mineral surface exploration were interrogated during this research project. As they were outside of the project scope, they were not integrated into the UltraFine+® Next Gen Data Package. Some were only tested on individual sites or cannot be used at the “standard” sites and have not been tested for repeatability and reproducibility (i.e., catchments). Others were generated for future necessity and are presented in brief given they were not within the project scope and were funded through other sources (mainly the CSIRO) and will be published elsewhere (e.g., geochemical backgrounds for WA).

### 3.4.1 Catchment Analysis

Two datasets that were submitted as part of the research project contained stream sediment samples and these required consideration of the sample geochemistry in the context of

catchment area, rather than landscape type. The catchment analysis process consisted of two main steps: (1) delineating catchments for each sample, and (2) normalising geochemical concentrations with respect to properties of the catchment, such as area and lithology. During this research project, an initial draft workflow was developed that is not part of the main python code or data package. Site specific results are confidential, however, the methodology used is outlined below along with some illustrative figures.

#### *3.4.1.1 Catchment delineation*

The catchments for each sample were calculated from hydrological analysis of a digital elevation model. For this purpose, the Hydrologically Enforced Digital Elevation Model (DEM-H) derived from the SRTM acquired by NASA in 2000 (Gallant et al. 2011) was chosen. The following steps were performed using open-source Python libraries (Bartos 2022; Lindsay 2023) to calculate catchments for each sample:

1. Processing of the DEM-H model to fill local pits and depressions
2. Calculate flow direction and flow accumulation
3. Calculate stream network based on a minimum accumulation area
4. Snap sample locations to the nearest location on the stream network
5. Manually review sample locations with respect to satellite imagery and relocate samples onto the stream network location which best represents the sample catchment
6. For each sample location, determine all upstream pixels which together form its complete catchment area
7. Calculate incremental catchments by removing catchments associated with upstream samples
8. Reduce very large catchment areas by placing an upper threshold on flow distance to the outlet point (sample location)

##### 3.4.1.1.1 DEM models

The choice of DEM results in a significant difference in catchment delineations. Three DEM's were considered:

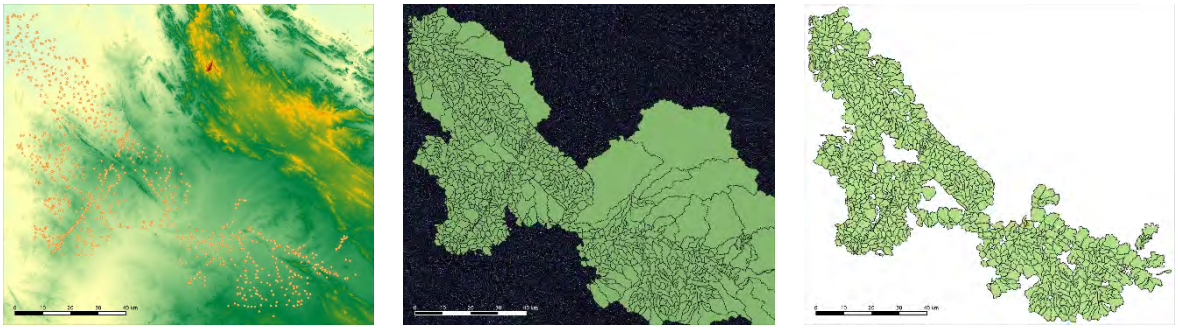
1. Hydrologically Enforced Digital Elevation Model (DEM-H) derived from the SRTM acquired by NASA in 2000 (Gallant et al. 2011).
2. Copernicus GLO-30, a Digital Surface Model (DSM) which represents the surface of the Earth including buildings, infrastructure and vegetation (*Copernicus Digital Elevation Model (DEM)* 2022)
3. FABDEM derived from the GLO-30 dataset with forests and buildings removed (Neal and Hawker 2023)

The DEM-H product has been smoothed and processed to include mapped streamlines for the specific purpose of hydrological modelling. However, in some specific locations, it did not appear to best represent the current stream network. The GLO-30 and FABDEM products are more recent, however, they contain artefacts (from either vegetation or smoothing) which inhibit their use for hydrological modeling. While all these models contain inaccuracies, ultimately, the DEM-H product was chosen as most fitting for catchment analysis.

##### *3.4.1.1.2 Geofabric*

An alternative source of catchments is the Australian Hydrological Geospatial Fabric (Geofabric) database (Bureau of Meteorology 2022) which provides a stream network and three levels of catchment hierarchy (drainage division, drainage basin, and sub-basin drainage

catchments) for all Australia. However, sampling locations must align with each Geofabric catchment outlet for these catchments to be used for analysis.



(a) Digital Elevation Model (b) Flow accumulation and full catchments (c) Final sample catchments

Figure 43: Selected images from the catchment delineation workflow including (a) the processed DEM-H model, (b) the flow accumulation and full catchment boundaries, and (c) the "flow-distance" reduced area catchments used for the analysis of catchment geochemistry.

#### *3.4.1.2 Catchment Geochemistry*

Two methods of analysing stream sediment sample geochemistry were considered to help identify potential anomalous sources within the catchments they represent. Both approaches attempt to account for catchment lithology.

##### 3.4.1.2.1 Uni-element productivity analysis

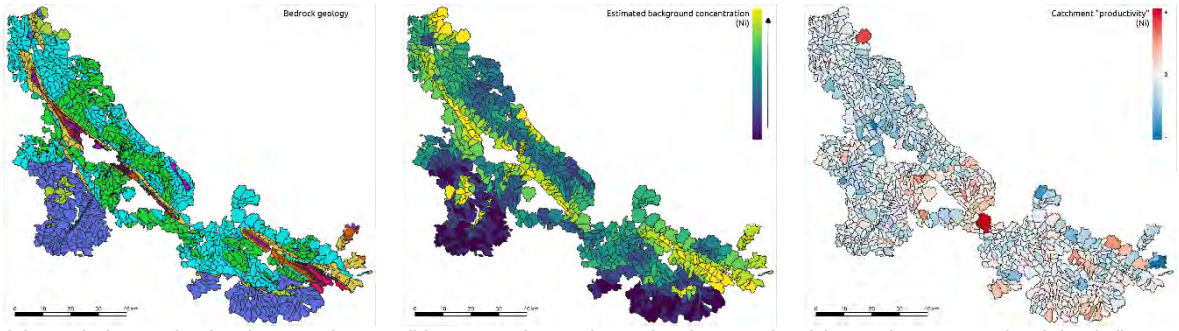
For each catchment a measure of "productivity" for a given element which accounts for downstream dilution is given by (Carranza 2009):

$$Y_a A_a = A_i(Y_i - Y_i') + Y_i' A_a$$

Where $Y_i$ is the uni-element concentration of stream sediment sample $i$ with catchment area $A_i$, $Y_a$ is the uni-element concentration of an anomalous source within the catchment with area $A_a$, and $Y_i'$ represents the uni-element concentration due to background sources within the catchment. For small anomalous sources the productivity of a catchment can be simplified to:

$$Y_a A_a = A_i(Y_i - Y_i')$$

The background geochemistry values, $Y_i'$, depend on lithology. Using an interpreted lithology map of the area, background values for each lithology type were estimated using a non-negative least squares optimisation, using the outlier-removed sample concentrations and the relative proportions of lithologies for each catchment. Background concentrations for each catchment were then estimated by using these regressed values, and the productivity of each catchment with respect to elements of interest can be calculated. Figure 44 shows this process.



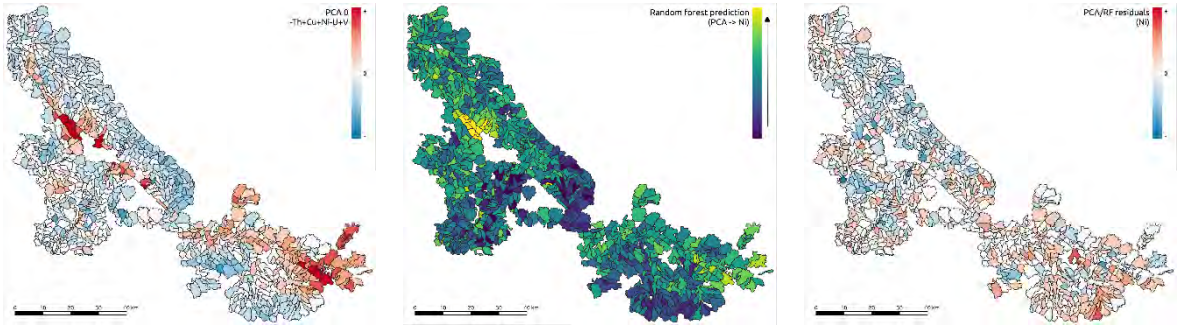

(a) Distinct bedrock geology classes (b) Estimated background concentrations (Ni) (c) Catchment productivity (Ni)

Figure 44: A bedrock geology map is used in conjunction with sample geochemisty to estimate background values for each lithology type and then catchment. From this a uni-element productivity measure can be calculated.

#### 3.4.1.2.2 Principal component analysis

An alternative to using interpreted lithology maps and estimating background concentration for the uni-element analysis as described above is to consider all the geochemistry using PCA. By capturing correlated elements among the suite of geochemistry data, principal components can reflect bedrock lithology. A method for normalising for lithological variation is to regress each element against the principal components and plot the residuals for each catchment (Arne et al., 2018).

Random forest regressor was trained on the first five principal components to predict geochemical concentrations. For elements where a relationship to principal components was observed, the residuals (difference between measured and predicted concentrations) were then plotted for each catchment, highlighting catchments which are significantly above or below background (Figure 45).




(a) First principal component (b) Random forest predictions (Ni) (c) Concentration residuals (Ni)

Figure 45: Principal components of the sample geochemistry are used to train a per-element random forest model which captures lithological variation. Differences between the model and measured concentrations highlight catchments of interest.

### *3.4.1.3 Further development*

Presented here are two among many possible approaches for analysing stream sediment sample geochemistry, aiming to indicate anomalous sources. A crucial task remains in initial catchment delineation, which requires an accurate DEM and manual verification with on-the-ground drainage patterns in the area. All three DEMs considered contained some inaccuracies when compared to satellite imagery.

The analysis has many potential areas for improvement which are common challenges encountered in state-of the art research as well as during industry application, including:

- reconsidering catchment size and the possibility of including the entire upstream catchment,
- modelling relationships between samples taken along the same stream (i.e., the catchment hierarchy),
- better modelling of the dilution effect by considering e.g., terrain slope/sinuosity (Sadeghi et al. 2021) and erosion rates as a function of geology, and
- validating models with any known mineralisation in the region.

Once sample catchments and lithology are accurately defined, there is potential to automate much of the analysis and plot generation via a machine learned workflow.

### 3.4.2 MrVBF concept testing for future application

The Multi-Resolution Valley Bottom Flatness (MrVBF) index is a key spatial feature layer used in the Next Gen Analytics workflow. This product is publicly available for the Australian continent (Gallant & Dowling 2003). However, no MrVBF is available for sites outside of Australia which affected one project site in New Zealand. The MrVBF is derived from DEM products and the legacy code to derive it is available. However, this legacy code used to create the original product is no longer functional, and it was necessary to reimplement the algorithm for the application of the workflow abroad. This was tested on a case study site in Western Australia by recreating the original map by Gallant and Dowling (2002), and more specifically reproducing the product generated by Gallant et al. (2012). The Python reimplementation of the MrVBF algorithm was developed using existing legacy ESRI Arc Macro Language (AML) and C scripts as a guide. The Python package was designed to replicate the existing algorithm (specifically the version `mrvbf6g-a3`) as closely as feasible, while taking advantage of scalable tools within the modern scientific Python ecosystem (including Dask; Dask Development Team 2016; Rocklin 2015). This implementation allows the algorithm to be used in a flexible manner (including as part of the existing Python-based workflows), reparameterised, and scaled to run in cloud infrastructure to allow larger product generation.

#### *3.4.2.1 Comparison to Existing Product*

As part of this research project, a test site in southwest Western Australia was used to compare generated products with the original MrVBF product. At the conclusion of this project, a similar but not identical result was generated with the Python reimplementation. As such it is not recommended to directly compare numerical values to existing MrVBF data, although very similar relative information is provided by the two products. Generally, it is expected that the algorithm/workflow may need specific tuning in different landscape settings as well as with the use of different DEM products that are available around the globe. The level of equivalence achieved in this implementation should be sufficient for the clustering workflow used in the Next Gen Analytics workflow (where values from different regions are not directly compared) although this warrants additional testing.

While the existing AML running on ArcGIS® provided a framework to reimplement the algorithm for the MrVBF, it remains unknown whether this corresponded to the published national MrVBF product, and thus comparison of outputs to this product has been required to validate the reimplementation. The existing product was generated from the smoothed SRTM DEM[11], and this DEM has been used as the input for the comparison exercise. Some modification of parameterization relative to that recorded in the AML script has also been

[11] https://ecat.ga.gov.au/geonetwork/srv/eng/catalog.search#/metadata/72759

required to provide comparable outputs; the likely values of two of these parameters were taken from the SAGA GIS MrVBF script (accessible via QGIS).

Currently, the large-scale structure of the MrVBF output can be replicated, but several differences remain related to the finer scale structure, and the relative proportions of pixels corresponding to different MrVBF values (Figure 46; i.e., the histogram of pixel values, Figure 47). Note that edge-effects are expected from the multi-scale nature of the algorithm, and these can clearly be observed towards the edges of the map below. Further work is required to determine whether a numerically comparable output is attainable via optimization of parameterization.

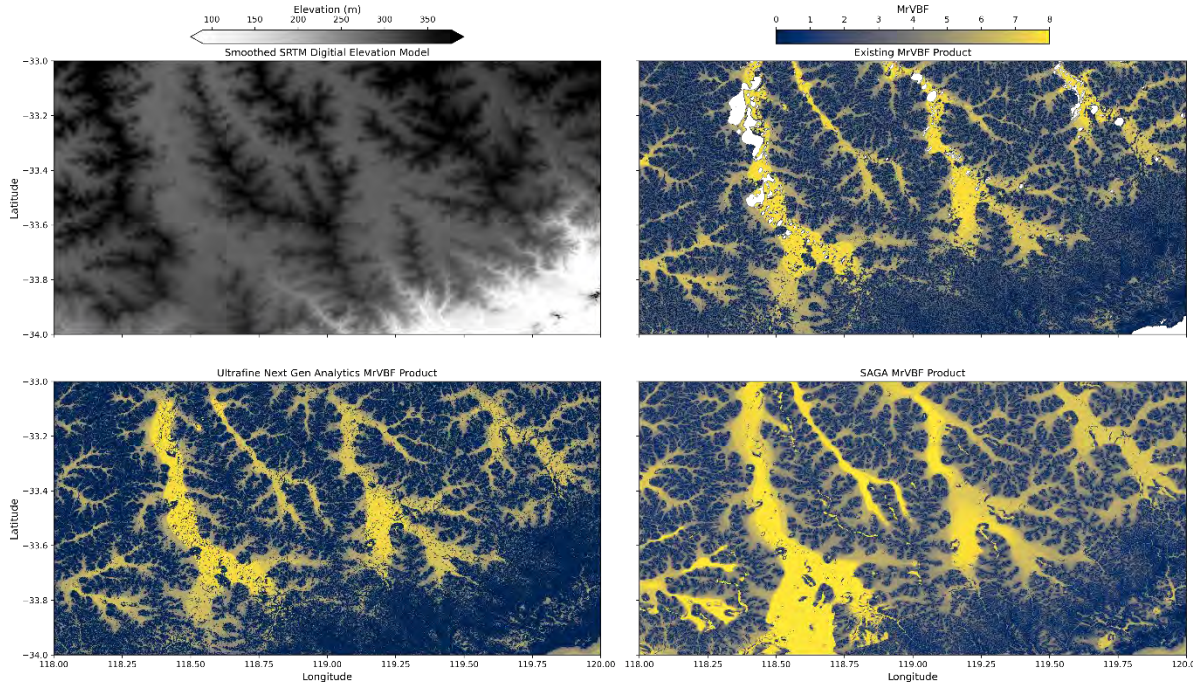



Figure 46: Comparison of the existing MrVBF product to the recreated output for a region of South-West Western Australia, highlighting some inconsistencies.

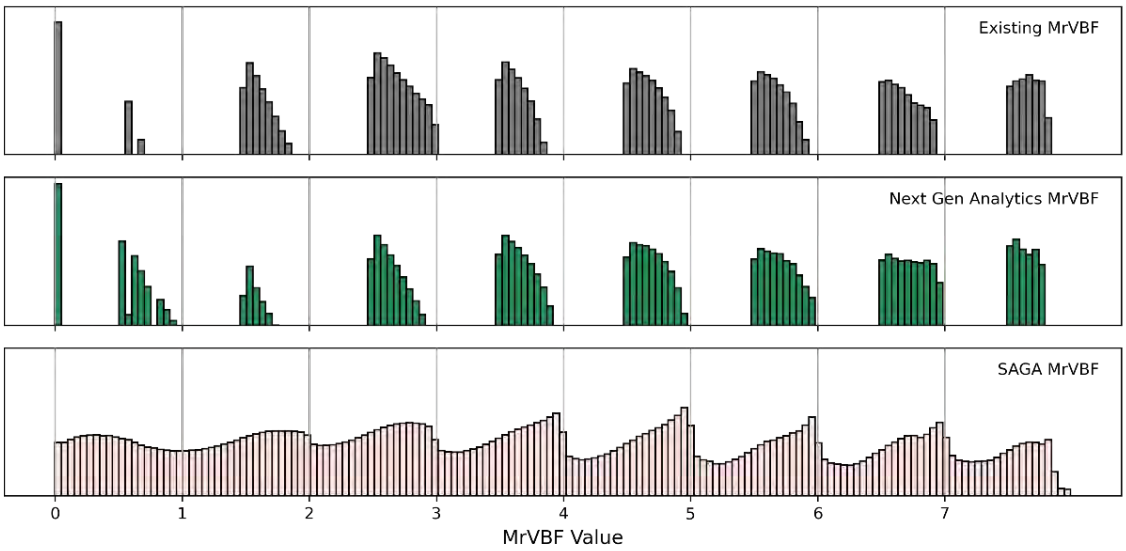



Figure 47: Comparison of histograms for the existing MrVBF product, the current product produced for Next Gen Analytics, and the default SAGA MrVBF output over the same area (restricted to an area bounded by [118.3, -33.75, 119.7,-33.25] to avoid known edge effects). Histograms are log-scaled to allow visualization of the proportion of pixels at zero (note that only the existing MrVBF product has been masked for water bodies, so some residual differences are expected, particularly in the highest MrVBF sections).

#### *3.4.2.2 Beyond Australia: MrVBF Product for New Zealand*

To generate MrVBF products overseas, the principal requirement is an available DEM at a suitable resolution (typically, on the order of 30m). An Ultrafine Next Gen Analytics site located in New Zealand (Reefton) provided an opportunity to test this in a new and contrasting environment (with the output over the relevant area shown below in Figure 48). For the New Zealand site, DEM data was sourced from the New Zealand Land Resource Information Systems Portal[12]. Selected tiles of the 25m DEM from the South Island dataset (Barringer, 2018) were merged to generate a seamless 25m DEM GeoTIFF with GDAL in QGIS (GDAL/OGR contributors, 2022). The original coordinate reference system was retained during this operation (NZGD49; EPSG:27200). The Python implementation of the MrVBF algorithm was applied to this DEM product using the original parameterisation of the algorithm (i.e., it was not tuned specifically to adjust for differences in the Australian and New Zealand landscapes).

Additional processing may be required at either end of the workflow, either to generate smoothed versions of digital elevation models (such as used for the original product), or to mask water bodies. In the case of the New Zealand site, a water mask was generated from polygon data for lakes, rivers and coastlines available via Land Information New Zealand's Data Portal[13], but quality water masks are not generally available in every country.

[12] lris.scinfo.org.nz
[13] data.linz.govt.nz

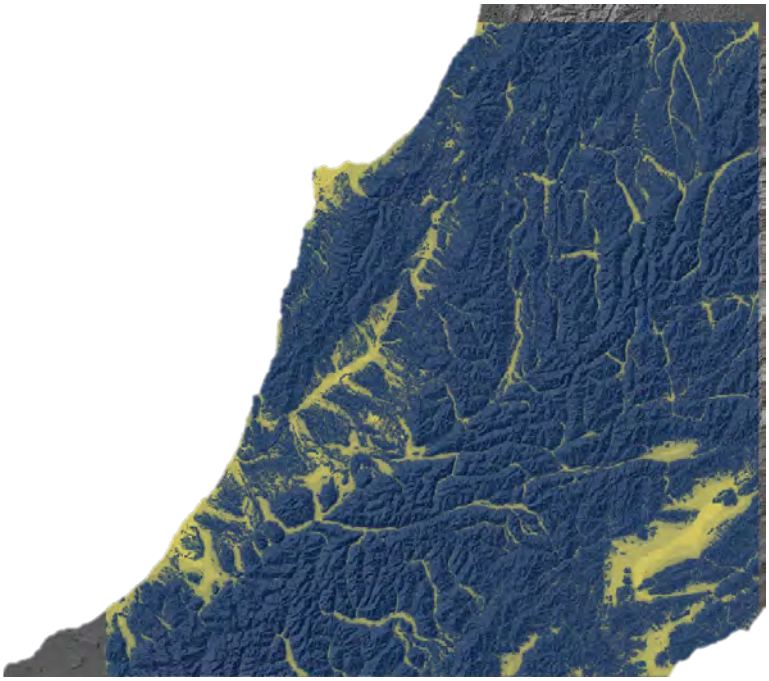


Figure 48: MrVBF created for the Reefton site on the South Island of New Zealand, shown underlain with a hillshaded DEM (high MrVBF values that represent deeper cover are shown in yellow colours).

### 3.4.3 Geochemical background

Over 74,000 soil samples were collected by numerous exploration companies and Australian state and territory geological surveys during this research project. A follow-on project funded by the CSIRO aimed to utilise this data to provide first-pass indications on regional background concentrations for future use by the exploration industry and has been submitted for publication (Henne et al. 2023b accepted pending revisions). The following is a short summary of initial findings.

The samples collected across Western Australia were grouped into subsets by their location over main tectonic units of general interest to mineral exploration using the digital 1:10 000 000 simplified tectonic units of Western Australia (GWSA 2022) to provide statistical values including mean, median, minimum, maximum and 99$^{th}$ percentile values for each of the dataset groupings. All values below detection limits were replaced with half the respective detection limit reported prior to calculations.

Most of the collected samples were part of commodity-specific exploration surveys that were based on various levels of previous geological information. Therefore, survey design and sample spacing limit the development of geochemical baselines for UltraFine+® geochemistry. In addition, elemental abundances in soils reflect a multitude of processes such as soil type, depth of sampling, cover type, depth of cover, the presence of and distance to nearby mineralisation, the type and potential alteration of this mineralisation, as well as a host of other processes affecting soil during erosion, weathering and (bio)geochemical cycling (Rose et al. 1979; Anand et al. 2014, 2016).

However, some overarching patterns and regional variations were observed that suggest the influence of underlying parent materials on the trace element composition found within the clay fraction of soil samples. This underscores the importance of providing estimated background abundance data specific to the primary tectonic units. Given the dataset's inherent

focus on exploration-centric regions, it offers explorers a prudent method for establishing realistic potential threshold values. In the absence of comprehensive, large-scale systematic background studies, this dataset serves as an initial reference point for the UltraFine+ technique for surface mineral exploration in Western Australia. Results for some key target and pathfinder elements are presented in Table 16, reproduced from Henne et al. 2023b. The accepted publication with summary statistics for the full suite of elements will be published in an open access journal and linked on the UltraFine webpage[14] .

[14] https://research.csiro.au/ultrafine/publications/

Table 16. Summary statistics for precious, base and some critical metals by tectonic unit. Values in italics indicate values below the detection limit (replaced with half the detection limit).

| | Tectonic unit | n* | Min | Max | Mean | Med | 95th | Min | Max | Mean | Med | 95th | Min | Max | Mean | Med | 95th | Min | Max | Mean | Med | 95th |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | **Ag, mg/kg** | | | | | **Au, µg/kg** | | | | | **Pd, µg/kg** | | | | | **Pt, µg/kg** | | | | |
| **Precious metals** | **All data** | 74,004 | *0.0015* | 9.23 | 0.07 | 0.056 | 0.173 | *0.25* | 7116 | 4.61 | 2.6 | 12.8 | *0.5* | 2540 | 13.16 | 4 | 89 | *0.5* | 357 | 4.17 | 3 | 11 |
| | **Youanmi Terrane** | 17,467 | *0.005* | 2.1 | 0.08 | 0.052 | 0.23 | *0.25* | 7116 | 5.59 | 3.3 | 14 | *0.5* | 212 | 41.34 | 8 | 125 | *0.5* | 186 | 6.37 | 5 | 17 |
| | **Eastern Goldfields** | 17,693 | *0.005* | 0.823 | 0.05 | 0.038 | 0.098 | *0.25* | 331 | 5.93 | 3.9 | 18.3 | *0.5* | 184 | 5.67 | 3 | 14 | *0.5* | 58 | 3.19 | 3 | 7 |
| | **Pilbara Craton** | 14,379 | *0.005* | 9.23 | 0.11 | 0.09 | 0.226 | *0.25* | 190 | 3.58 | 2.4 | 9.5 | *0.5* | 79 | 8.14 | 6 | 20 | *0.5* | 290 | 5.02 | 4 | 11 |
| | **Gascoyne Province** | 12,206 | 0.006 | 0.799 | 0.07 | 0.068 | 0.139 | *0.25* | 44.6 | 2.04 | 1.5 | 5.1 | *0.5* | 33 | 3.48 | 3 | 8 | *0.5* | 13 | 1.85 | 2 | 4 |
| | **Lamboo Province** | 5,167 | 0.011 | 4 | 0.09 | 0.068 | 0.204 | *0.25* | 1243 | 9.48 | 2.7 | 29.42 | *0.5* | 2540 | 19.32 | 5 | 55 | *0.5* | 357 | 5.54 | 2 | 17 |
| | **Aileron Province** | 547 | *0.005* | 0.08 | 0.03 | 0.03 | 0.05 | *0.25* | 3.1 | 0.81 | 0.8 | 1.5 | N/A | N/A | N/A | N/A | N/A | *0.5* | 3 | 0.86 | 0.5 | 2 |
| | **Sylvania Inlier** | 4,560 | *0.0015* | 0.566 | 0.05 | 0.037 | 0.127 | *0.25* | 111.8 | 1.88 | 1.5 | 4.195 | *0.5* | 41 | 5.48 | 4 | 15 | *0.5* | 32 | 2.77 | 2 | 5 |
| | **Bryah Basin** | 54 | 0.04 | 0.11 | 0.05 | 0.05 | 0.08 | 0.5 | 4.9 | 1.71 | 1.45 | 3.825 | N/A | N/A | N/A | N/A | N/A | 4 | 21 | 9.20 | 9 | 16 |
| | **Earaheedy Basin** | 1,931 | 0.009 | 0.233 | 0.03 | 0.031 | 0.048 | *0.25* | 52.5 | 1.89 | 1.4 | 4.7 | *0.5* | 34 | 7.55 | 7 | 14 | *0.5* | 9 | 1.72 | 2 | 4 |
| | | | **Cu, mg/kg** | | | | | **Ni, mg/kg** | | | | | **Pb, mg/kg** | | | | | **Zn, mg/kg** | | | | |
| **Base metals** | **All data** | 74,004 | 1.6 | 8650 | 73.69 | 55.7 | 168.75 | 2.8 | 6690 | 90.16 | 68.5 | 188 | 1.11 | 7080 | 26.71 | 22.7 | 48.6 | 3.5 | 3300 | 86.03 | 81.5 | 143 |
| | **Youanmi Terrane** | 17,467 | 1.6 | 1110 | 98.03 | 73 | 250 | 5.5 | 2040 | 104.35 | 94.1 | 228 | 2.43 | 706 | 23.06 | 22 | 37.3 | 6 | 3300 | 91.56 | 85.5 | 156 |
| | **Eastern Goldfields** | 17,693 | 6 | 702 | 55.60 | 51.4 | 93.8 | 7.1 | 1990 | 84.45 | 68.7 | 165 | 1.47 | 328 | 20.26 | 19.2 | 34.2 | 11.6 | 528 | 87.91 | 86.4 | 127 |
| | **Pilbara Craton** | 14,379 | 9.8 | 4960 | 73.09 | 58 | 132 | 17.2 | 2120 | 112.24 | 106 | 180 | 4.12 | 7080 | 37.86 | 26.4 | 71.1 | 3.5 | 2070 | 85.78 | 70.7 | 143 |
| | **Gascoyne Province** | 12,206 | 3.6 | 585 | 54.69 | 51.5 | 87.9 | 2.8 | 155 | 41.90 | 41.5 | 59.9 | 4.45 | 609 | 30.77 | 28.3 | 52.7 | 8.6 | 354 | 99.68 | 99 | 145 |
| | **Lamboo Province** | 5,167 | 2.4 | 8650 | 139.68 | 95.5 | 373 | 7.4 | 6690 | 178.96 | 67.2 | 786.6 | 1.11 | 101 | 15.49 | 13.8 | 32.46 | 5.2 | 608 | 50.76 | 46 | 96.46 |
| | **Aileron Province** | 547 | 5.6 | 90.5 | 16.52 | 15.7 | 27.38 | 8 | 147 | 23.74 | 23 | 36 | 8.3 | 79 | 19.51 | 18.6 | 26.7 | 13.4 | 132 | 31.06 | 28.3 | 50.74 |
| | **Sylvania Inlier** | 4,560 | 7.4 | 290 | 54.91 | 50.6 | 93.59 | 4.9 | 1940 | 51.49 | 43 | 99.385 | 3.32 | 1390 | 33.36 | 27 | 66.295 | 7.6 | 1130 | 87.66 | 76.45 | 166 |
| | **Bryah Basin** | 54 | 48.3 | 90.5 | 59.53 | 58.4 | 81.6 | 26 | 70 | 35.22 | 33 | 53 | 20 | 34.3 | 24.74 | 24.85 | 29.25 | 67.4 | 350 | 116.27 | 103.5 | 233.75 |
| | **Earaheedy Basin** | 1,931 | 9 | 106 | 28.35 | 26.3 | 43.42 | 10 | 246 | 28.84 | 27.4 | 42 | 12.4 | 149 | 26.63 | 26.3 | 34 | 4.8 | 203 | 39.52 | 36.5 | 64.48 |
| | | | **Co, mg/kg** | | | | | **Cs, mg/kg** | | | | | **La, mg/kg** | | | | | **Li, mg/kg** | | | | |
| **Critical elements** | **All data** | 74,004 | 1.02 | 489 | 30.41 | 27 | 61 | 0.07 | 256 | 5.58 | 4.68 | 11.575 | *0.025* | 651 | 27.99 | 21 | 65.7 | *0.025* | 382 | 38.30 | 33.9 | 78.4 |

| | Tectonic unit | n* | Min | Max | Mean | Med | 95th | Min | Max | Mean | Med | 95th | Min | Max | Mean | Med | 95th | Min | Max | Mean | Med | 95th |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Youanmi Terrane** | 17,467 | 1.99 | 285 | 31.48 | 27.3 | 72.9 | 0.17 | 256 | 6.44 | 4.75 | 15.4 | 1.01 | 372 | 19.10 | 17.6 | 33.5 | 1.52 | 265 | 34.54 | 31.3 | 70.4 |
| | **Eastern Goldfields** | 17,693 | 2.6 | 146 | 24.13 | 21.6 | 45.6 | 0.39 | 26.1 | 3.91 | 3.9 | 5.5 | 0.7 | 378 | 22.31 | 19 | 45.83 | 1.45 | 167 | 28.93 | 27.1 | 49.5 |
| | **Pilbara Craton** | 14,379 | 4.62 | 271 | 35.91 | 33.1 | 64.6 | 0.76 | 102 | 6.31 | 5.61 | 12.6 | *0.025* | 651 | 25.20 | 21.2 | 46.7 | 4.6 | 382 | 57.15 | 50.9 | 102 |
| | **Gascoyne Province** | 12,206 | 1.02 | 98.8 | 29.94 | 28.4 | 49.465 | 0.28 | 43.6 | 7.25 | 6.55 | 12.7 | 7.4 | 496 | 55.63 | 40.8 | 157 | 3.65 | 151 | 47.49 | 44.1 | 80.4 |
| | **Lamboo Province** | 5,167 | 6.68 | 489 | 43.76 | 34 | 106.6 | 0.29 | 37.2 | 3.30 | 2.36 | 8.912 | 0.65 | 423 | 24.47 | 19 | 57.3 | *0.025* | 113 | 25.14 | 23.4 | 47.96 |
| | **Aileron Province** | 547 | 3.8 | 38.7 | 11.42 | 10.8 | 18.76 | 0.7 | 8.4 | 3.21 | 3.1 | 5.7 | 7.01 | 71.7 | 20.75 | 20.4 | 33.4 | 9.7 | 64.4 | 28.27 | 27.7 | 45.66 |
| | **Sylvania Inlier** | 4,560 | 1.78 | 350 | 27.85 | 25.8 | 55 | 0.07 | 17.1 | 4.71 | 4.66 | 7.0895 | 1.04 | 358 | 28.85 | 24.6 | 63.295 | 2.63 | 112 | 25.25 | 23.8 | 44.5 |
| | **Bryah Basin** | 54 | 6.3 | 41.1 | 12.55 | 9.6 | 28.55 | 4.8 | 7.1 | 5.74 | 5.7 | 6.825 | 23.8 | 70 | 41.99 | 39.3 | 60.125 | 7.9 | 36.5 | 17.83 | 16.85 | 28.775 |
| | **Earaheedy Basin** | 1,931 | 3.2 | 209 | 16.65 | 15.6 | 29.08 | 2.8 | 14.2 | 5.92 | 5.87 | 7.7 | 5.61 | 179 | 15.50 | 14 | 29.12 | 3.29 | 86.5 | 29.13 | 29.2 | 42.74 |

*Refer to Henne et al. 2023b for full details.

#### 3.4.3.1 Example application of background ranges – Li, Pilbara Craton

Although limitations from processes that can influence “background” geochemistry exist, the soil geochemistry dataset presented provides background estimates by tectonic unit that can be used for first-pass interpretation. Rather than minimum, maximum or even median values, the 95th (or higher) percentiles may be a more conservative approach for use in exploration settings. For example, a soil sampling campaign in the Pilbara Craton targeting Au and Li noted “multiple large lithium targets” associated with Cs, Ta and Sn had been identified using the UltraFine+® method (Figure 49a). Lithium abundances in this survey with ~800 samples ranged from 28.5 mg/kg to 382 mg/kg, which is strongly elevated compared to publicly available background values from the National Geochemical Survey of Australia dataset (11 samples within the Pilbara Craton with a maximum value of 20.1 mg/kg in the < 2 mm size fraction analysed via aqua regia; de Caritat & Cooper, 2011). However, the < 2 µm size fraction is dominated by clay phases and Li ions can adsorb to external clay surfaces (e.g., kaolinite) and accumulate in clay interlayers (e.g., smectite; Li & Liu, 2022). Hence, the abundances of Li analysed with this method are expected to be elevated and it could be argued that the targets are only slightly greater than the enhanced background related to the method itself.

However, by using the new background UltraFine+® data and the 95$^{th}$ and 99$^{th}$ percentile values calculated from the over 14,000 analyses in the Pilbara Craton the interpretation is improved. The dataset shows the 95$^{th}$ and 99$^{th}$ percentile values of 102 mg/kg and 147 mg/kg for Li and 12.6 and 19.4 mg/kg, for Cs, respectively. This provides more confidence in the potential targets identified by the company (Figure 49a,c) as much of the Li abundances within the survey area are above the 95$^{th}$ and 99$^{th}$ percentile for the greater region, and similar trends were confirmed in other elemental abundance ranges (e.g., Cs; Figure 49b, d). It is important to note that no details on potential mineralisation at depth are known at this stage, and without these data it is impossible to confirm whether soil anomalies are related to mineralisation, but at least the initial assessment that the region has anomalous Li can be substantiated without having to collect more samples from background areas or via incorrect comparisons to other soil data. This is a common challenge in industry as numerous different methods and laboratories are used that do not match publicly available data (and their methods). With the addition of this UltraFine+® data, industry is better positioned to benefit in future exploration surveys. It is hoped that these resources will continue to build for other areas with the assumption that data sharing will be more prevalent (and data more interoperable).

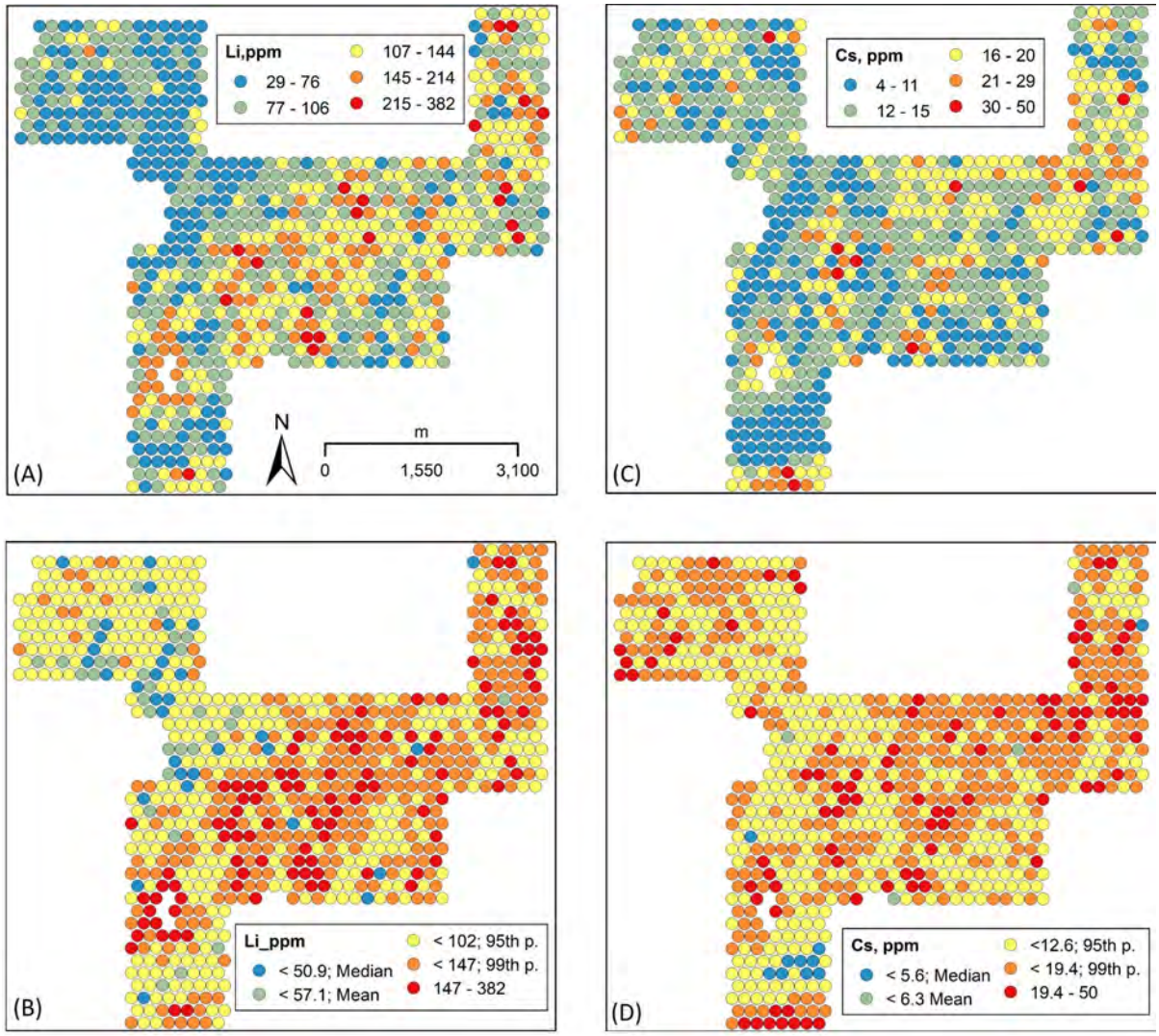



Figure 49. Lithium and Cs abundances in soil geochemical samples from a survey in the Pilbara Craton. (**a**) Li abundances, colours relate to bins defined by natural breaks. (**b**) Cs concentrations, colours relate to bins by natural breaks. (**c**) Li abundances, colours relate to bins defined by Median, Mean, 95$^{th}$ percentile, and 99$^{th}$ percentile of the 14,379 samples collected within the Pilbara Craton. (**d**) Cs abundances, colours relate to bins defined by Median, Mean, 95$^{th}$ percentile, and 99$^{th}$ percentile of the 14,379 samples collected within the Pilbara Craton. Please note that for confidentiality reasons, the location of samples and further details cannot be disclosed. Refer to Henne et al. 2023 for more details.

# 4 Conclusions

With the development of a robust set of measurable parameters and new data products to fully assess underappreciated soil properties, the UltraFine+® soil analytical method and the Next Gen Analytics landscape modelling workflow provide next generation analytical tools for mineral explorers to make qualified decisions on when and where to direct further exploration, delivered via the UltraFine+® Next Gen Analytics Data Package.

The M0462a UltraFine+® Next Gen Analytics research project was conducted by the CSIRO in collaboration with LabWest and 22 industry sponsors, five State and Territory geological surveys and MRIWA. The aim of the research project was to facilitate a paradigm shift for precious, base, and critical metals exploration in Australia by combining UltraFine+® soil analytical methods with purpose-built data integration tools, the Next Gen Analytics, to add value to routine soil sampling in frontline exploration, shaping mineral exploration approaches for the coming decades. The Next Gen Analytics workflow was designed primarily for greenfield exploration areas in transported cover with 100s to 1,000s of soil samples and to generate interpretation of these samples prior to significant ground disturbance.

The project successfully achieved the pairing of the UltraFine+® method and the integration of machine learning, delivering actionable tools via the output data package, Digital Sample Observer and web-based guidance resources. At the outset, the data outputs were expected to comprise a simple spreadsheet of results and a georeferenced map of the landscape proxies generated by the machine learning approach. Instead, the final delivered output for each site was a comprehensive data package that delivers three different landscape models (4, 8 and 12 classes) with dozens of figures, files and formats that can be rapidly reviewed. The project team has produced these models and outputs for over 30 sites across Australia for the numerous sponsors and collaborative partners. More importantly, the refinements to the commercially available soil analytical method and the developed machine learning workflow can be used in future exploration programs to accelerate exploration and decision making.

While the term “paradigm shift” is overused, in exploration geochemistry this project has demonstrated significant advances and prototyped analysis and data delivery for future near surface exploration that has never been seen prior – a paradigm shift for explorers and the analytical laboratory providing the UltraFine+® results. This is akin to major changes in this research field such as instrumental analysis with ICP-MS and portable XRF. It is likely that the research developed in this project supersedes previous major extractive changes such as the partial extractions like MMI™ and its competition, that may be argued to represent the last significant analytical change in surface geochemical exploration science.

In addition to the site results generated for industry partners, the project has delivered approximately 20 publications, 35 presentations including sponsor data presentations, supported a student project (with an additional student project to draw on data acquired during this project) and much more (refer to available online resources[15]). The impact to industry and the transformation of LabWest as the service provider of UltraFine+® is evident in their web presence[16], and particularly in the ~190 ASX releases[17] that capture the interest and

[15] https://research.csiro.au/ultrafine/
[16] https://www.labwest.net/ultrafine/
[17] https://www.labwest.net/asx-releases-referencing-uff/

successes generated through applying the approach developed in the MRIWA M0462 and M0462a projects.

The original project workplan aimed to address the following key points, some of which evolved further during the project:

(1) Results will be characterised and assessed with respect to known mineralisation to understand mobility in a landscape evolution context.
(2) Orientation samples in different cover environments will be compared (representative soils) to understand variance in the sample media and regolith settings to inform supervised ML options and CSIRO IP regolith geoscience algorithms. All results will be placed in an exploration targeting context with landscape models.
(3) In association with the orientation work and in-kind samples provided by industry or the Surveys, freely-available data will be incorporated into several ML outputs (e.g. compare random forests, neural networks, gaussian process regressor) to determine the best fit-for-purpose technique(s).
(4) CSIRO expertise in interpretation will be added (satellite imagery bands, element ratios, landscape position estimates and then respective comparison of like types of data) as a demonstration of potential future automated outputs.
(5) Application of spatial statistics will be developed to inform variation and uncertainty estimates at various sites as a demonstration of potential future automated outputs.
(6) Proxies for organic C and broader wavelength spectral mineral proxies will be built into the early stage analytical development. Use of historic agricultural data (20-30k sample results) will be utilised to achieve a robust test of organic C estimates.

With respect to (1) above, where study sites had known mineralisation, this information was incorporated. UltraFine+® has successfully located mineralisation at sites such as the DeGrussa Mine, haloes around the Bibra Mine, and clearly demonstrates dispersion down slope near Federation and Dominion mineralisation as noted in this report and Henne et al. (2022b). Published results in Noble et al. (2020), and many ASX releases that highlight anomalous results of target and pathfinder metals in the expected areas of interest attest to the potential of the approach. However, due to limited drilling information this cannot be expanded on further at this stage. The uptake of the truly greenfield approach was exceptional but limited the metal dispersion study component due to a lack of known mineralisation in many study areas. Orientation studies were minimal for known mineralisation, but the vast uptake allowed a very large range of results and soils to be tested. Results from organic-rich younger soils in New Zealand compared to the extremely weathered and older soils of the Yilgarn and Pilbara Cratons provided ample test cases and overall, the generated outputs added value regardless of setting and size. Supervised ML options were limited to the known mineralisation targets and using supervised ML is likely better suited to future projects. Much of the ML approaches were unsupervised, building the baseline for future supervised tests as noted in (2) above.

Item (3) was completed early in the evolution of the Next Gen Analytics workflow with the general results presented and published by Williams et al., 2023 and the final output documented in section 1.4.3, whereas (4) was fully realised in the Digital Sample Observer generated for the sponsor data package and the related components (section 1.4.9). This vastly exceeded the earlier plans for the project and is now progressing towards a commercially available product. The project team developed a prototype data delivery package and is now moving to progressing this to become commercially available.

Spatial statistics and uncertainty estimates (5) were tested in a number of projects with industry such as landscape models and predictive models with Carnavale (Williams & Noble 2021) and with Iceni Gold (unpublished) looking at analytical variation in batches over time. This is also captured in a review of the many laboratory jobs done by LabWest and the standards QC 320 and QC 422 published in Henne et al. (2023b) looking at background geochemical values. Sample design has also demonstrated how spatial data and uncertainty can be used for more efficient sampling. This was part of the ground-truth sampling in Mundaring presented to sponsors as part of the student projects (Henne et al., 2023c) as well as other work conducted for the Geological Survey of South Australia (unpublished). The further application of these aspects into data outputs is ongoing with significant scope for more research in this area.

Finally, the proxies for organic C and broader wavelength spectral mineral proxies have become part of the standard UltraFine+® output (6). FTIR analysis produced 5 new features for the method and the results for organic C were generally superior to the organic C estimates using NIR as tested in other projects outside the scope of the work (i.e., carbon fixation for $CO_2$ capture in agricultural settings; Birchall et al., 2022). This is just one of the examples of how the work completed during this project has had flow on effects into other applications.

## 4.1 Project benefits and impacts

The project benefits for the industry partners in the proposal were listed as follows:

1. Higher probability of success using surface geochemistry;
2. Increased use and value of state and territory geological survey data;
3. Access to the newest version of the UltraFine+® soil methodology and analysis via Next Gen Analytics machine learning algorithms;
4. Potential methodology to explain or distinguish false positives from true anomalies;
5. Greater benefit for explorers with limited capability in geochemical interpretation and machine learning integration;
6. Generate new search space/refine old targets;
7. First access to new regional and orientation data for different commodities, regolith regimes and climates;
8. Regolith-landform based approach for generating and interpreting data for a range of soil environments;
9. Create a new method and demand from industry that is exportable globally;
10. Keep current, and attract new, exploration expenditure in Australia.

It is evident that the benefits of the project have been achieved above and beyond expectations given the uptake of the method (>200,000 samples run with the technique; pers. com. LabWest) and the interpretation of survey data in landscape context at approximately 40 sites which informed the development of the standardised and semi-automatic Next Gen Analytics machine learning workflow. In addition, a data package that delivers a vast number of products in various data formats (csv, tiff, png and shapefiles) as well as a html data viewer that enables the interrogation of both the soil sample results and the landscape context was developed. These developments were complemented by the generation of explanatory notes that were published concurrently with development to ensure knowledge transfer to the sponsors and the wider exploration community. However, there are aspects (i.e., benefit #4; Potential methodology for distinguishing false positives from true anomaly) that were not clearly established in this project. It was hoped that the integration of landscape analysis and the additional soil properties would aid in the identification of such false positives. However, due to the lag between generating targets, drilling and determining whether an anomaly was

false or not, as well as the lack of surveys over clearly defined, known and economic deposits at reasonable depth (<30 m), this could not be achieved at the time of writing this report. Anecdotal evidence suggests subeconomic successful targets and some false positives have been drilled and the project team continues to work with the industry partners to follow up on these results as it is critical to continue to refine the interpretation of exploration outputs.

The SME industry were the dominant class of supporters in this project principally because they lack internal access to expertise and ML with limited staff numbers. Hence, benefit #5 and #6 have been realised. Some older surveys were reprocessed in the State Survey reports (Henne et al., 2022a; Noble et al., 2023 unpublished) and these have generated some different targets to some of the original work. These results were also rapidly reported to project sponsors as they were released providing first access for the project sponsors. It is clear from the Next Gen Analytics data package that the benefits of the landscape context have been successfully generated.

Finally, with benefits #9 and #10 above, sponsors provided many samples to the project, well above expectations (10x initial plans) with 5000-8000 samples becoming over 80,000 samples in this project. The uptake by the local industry continues with more than 180 projects completed by LabWest. In some cases, the companies that have used UltraFine+® have international interests and this has meant the UltraFine+® method has been used internationally, too, with a vast opportunity to expand globally. This project tested samples from New Zealand and the successful quarantine heat treatment testing means there are now no limitations to importing samples. UltraFine+® has been tested in Finland, Japan, Tanzania, Burkina Faso, New Zealand, and Namibia etc (*pers. comm.* Blake Stacey, LabWest) with samples from these countries shipped to WA and processed. Coupling the local WA success, increased national uptake and international interest, the project has certainly achieved the larger outcome of keeping current and attracting new exploration interest to Australia, as well as building a product that will attract international exploration expenditure to Australia through the WA service provider.

In addition to the M0462a project, a separate impact study that focused on the previously developed UltraFine+® method was conducted for CSIRO by external consultants Tractuum[18]. The M0462a project has assisted in building these impacts further. A summary of these impacts includes:

*UltraFine+® has seen a surge in demand since its roll-out and has been mentioned in many ASX releases. It has been used by 163 mineral exploration companies within Australia and making its way into global markets. The adoption of technique has delivered a range of incremental economic, social, and environmental impacts compared to traditional soil sampling methods, including (but not limited to):*

*– reduced risks, and costs, and improved productivity for the mineral exploration sector. Quicker ability to determine the commercial viability of deposits*

*– greater demand and new revenue streams for LabWest (an SME). Growth in revenue of up to 80% and creation of new jobs*

*– relatively lower drilling footprint (vs traditional technique)*

*– new CSIRO sources of revenue and collaborations (internal and external)*

[18] https://research.csiro.au/ultrafine/wp-content/uploads/sites/463/2023/11/Ultrafine-Impact-Assessment-FY23-Public-version.pdf

*Cost Benefit Analysis conducted from Australia's perspective for period between FY2020-FY2033 estimates:*

*– BCR of 36 and NPV of ~49 mil AUD (in present value) without deadweight loss*

*– BCR of 30 and NPV of ~48 mil AUD (in present value) with deadweight loss*

*Other prospective outcomes*

*LabWest is working towards driving the adoption of UltraFine+® in overseas markets and has already entered six countries (Finland, Japan, Tanzania, Burkina Faso, New Zealand, and Namibia). There are future prospects for export income and the adoption of technique for the detection of critical metals and high-value rare earth elements in emerging markets.*

*UltraFine+® Next Gen Analytics prototype workflow development is underway (not yet commercialised) and a product of collaboration between CSIRO Minerals and Data61 BUs, post the success of foundational UltraFine+® R&D. This prototype integrates Machine Learning (ML) with decades of regolith expertise to deliver results that can be viewed via a HTML-based dashboard to fast-track analysis. There is significant enthusiasm within the industry to test and try the capabilities of this new platform to prioritise target areas for exploration, simplify and accelerate the exploration workflows*.

## 4.2 Future research

Projects such as M0462a and its precursor M0462 are built upon years of incremental developments. This project has achieved major steps in changing the way industry collects, receives, and reviews soil geochemical data. It has also demonstrated that there is more potential to improve. Several scientific objectives have been identified by the project team that will be pursued in the future. The ML approaches that support the landscape context work, as well as technical lessons developed in the project, will also be part of future research. Proposed publications from this project include journal articles furthering the discussion of geochemical backgrounds, more detailed site studies and follow up around false positives that include tracking metal mobility through cover or understanding how geochemical variables are altered with changes in cover thickness and separating detrital, allochthonous, and transported signatures. There is a big opportunity to further investigate the "other" parameters like pH and spectral mineralogy in relation to regolith settings. Some further method comparisons, including spectral mineralogy and the quarantine heat treatment studies are also being investigated for publication. This also includes more detailed work around applying the UltraFine+® method to other critical metals that were not the target of the R&D when the project started. This requires continued development of the QA/QC outputs especially the reference standards such as QC 320 and QC 422 and others that will be required.

A "Click and collect" delivery of the Next Gen Analytics product was the long-term objective that was not funded via a CRC-P bid and this component was not part of outside of this specific project scope. However, the project team has progressed steadily towards this goal and CSIRO is actively investigating the commercialisation of the product or service delivery. A future data delivery product is a reality, with the timeframe dependent on technical capacity and funding support. Market push is likely to accelerate the process compared to the research pull. In addition to the currently developed Next Gen Analytics outputs, there is clear scope for expansion with a "build your own landscape" model and the ability to integrate other data (agnostic to UltraFine+® geochemistry). Adaptations of the workflow are already being used internally at CSIRO for targeting Li pegmatites, updating regolith maps for transported versus residual cover, and adjusting soil survey designs, among other things. However, these are

currently boutique, tailored R&D solutions, rather than an established ML workflow on an automated platform for a standard user (e.g., the exploration geologist). Opportunities for adaptation of the landscape modelling workflow outside of mineral exploration are also evident. Aspects such as ecological management, land use optimisation or mapping wine terroir are all viable future research areas on the basis of the developments achieved in the M0462a project, and it will continue to evolve with other spatial data analytics and modelling research.

Finally, the UltraFine+® workflow has mostly been applied in Western Australia. The vast expanses of shallow (<20 m) cover hide much mineral wealth globally and the future will see the R&D established here expanded into other regions. However, some further adjustments will be required for areas with glacial till cover (e.g., Canada, Finland) to improve the Next Gen Analytics workflow and outputs. Significant scope exists to expand this research for applications in these areas. Key developments such as generating the MrVBF or similar derivatives for other countries or factoring in ice flow direction are in the early test work stages that would need development. The current project team has made incremental improvements toward these objectives. However, these were outside the scope of this project.

Overall, the future for UltraFine+® Next Gen Analytics is positive. A paradigm shift for precious, base, and critical metals exploration in Australia was successfully delivered during the M0462a project through the combination of the UltraFine+® soil analytical method with the Next Gen Analytics tools. The project has added value to routine soil sampling in frontline exploration, shaping mineral exploration approaches for the coming decades.

# 5 References


A guide to public domain spatial data is available at: https://spatialreserves.wordpress.com/.

Anand R.R., & Paine, M., 2002. Regolith geology of the Yilgarn Craton, Western Australia: Implications for exploration, Australian Journal of Earth Sciences, 49:1, 3-162, DOI: 10.1046/j.1440-0952.2002.00912.x

Anand, R., Lintern, M., Noble, R., Aspandiar, M., Macfarlane, C., Hough, R., Stewart, A., Wakelin, S., Townley, B. & 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., Aspandiar, M. & 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, https://doi.org/10.1016/j.oregeorev.2015.06.018

Arne, D. & 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

Arne, D.C., Mackie, R., Pennimpede, C., Grunsky, E, & Bodnar, M. (2018). *Integrated Assessment of Regional Stream-Sediment Geochemistry for Metallic Deposits in North-western British Columbia (Parts of NTS 093, 094, 103, 104)* (Geoscience BC Report 2018-14, 94 p.).

Barnes, S.J. & Brand, N.W., 1999. The distribution of Cr, Ni, and chromite in komatiites, and application to exploration for komatiite hosted nickel sulfide deposits. *Economic Geology* 94: 129–132.

Barnes, S.J., Hill, R.E.T., Perring C.S. & Dowling, S.E., 2004. Lithogeochemical exploration for komatiite-associated Ni-sulfide deposits: strategies and limitations. *Mineralogy and Petrology* 82: 259– 293.

Barnes, S.J., Fisher, L.A., Anand, R.A. & Uemoto, T., 2014. Mapping bedrock lithologies through in-situ regolith using retained element ratios: a case study from the Agnew-Lawlers area, Western Australia, Australian Journal of Earth Sciences, 61:2, 269-285, DOI:10.1080/08120099.2014.862302

Barringer, J., 2018. NZDEM South Island 25 metre [Data set]. Manaaki Whenua Landcare Research. https://doi.org/10.7931/L1R94

Bartos, M., 2022. *pysheds: simple and fast watershed delineation in python* (0.3.3) [Computer software]. https://doi.org/10.5281/ZENODO.3822494

Birchall, R., Pejcic, B., Williams, M., Cole, D., Pinchand, T., Shelton, T., Noble, R. & Henne, A., 2022. In-situ infrared sensing for low-cost, accurate detection of soil organic carbon - National Soil Carbon Innovation Challenge – Feasibility Study. Perth, CSIRO Report EP2022-4343. p.28.

Bureau of Meteorology, 2022. *Australian Hydrological Geospatial Fabric (Geofabric)* (Version 3.3). http://www.bom.gov.au/water/geofabric/download.shtml

Butt, C.R.M., Robertson, I.D.M., Scott, K.M. & Cornelius, M., (Editors) 2005. Regolith expression of Australian ore systems. CRC LEME, Perth 431p.

Butt, C.R.M., 2016. The development of regolith exploration geochemistry in the tropics and sub-tropics. Ore Geology Reviews, Volume 73, Part 3, Pages 380-393,ISSN 0169-1368, https://doi.org/10.1016/j.oregeorev.2015.08.018.

Carranza, E. J. M.,2009. *Geochemical anomaly and mineral prospectivity mapping in GIS* (1st ed). Elsevier.

*Copernicus Digital Elevation Model (DEM)*. (2022). https://registry.opendata.aws/copernicus-dem

Copernicus GLO-30 Digital Elevation Model https://registry.opendata.aws/copernicus-dem

Dask Development Team, 2016. Dask: Library for dynamic task scheduling. http://dask.pydata.org

de Caritat, P. &Cooper, M., 2011. National Geochemical Survey of Australia: The Geochemical Atlas of Australia. Record 2011/20. Geoscience Australia, Canberra, http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_71973

Digital Earth Australia https://explorer.sandbox.dea.ga.gov.au/products/s2_barest_earth

Dinis, P.A., Garzanti, E., Hahn, A., Vermeesch, P. & Cabral-Pinto, M., 2020. Weathering indices as climate proxies. A step forward based on Congo and SW African river muds, Earth-Science Reviews, Volume 201, 103039, ISSN 0012-8252, https://doi.org/10.1016/j.earscirev.2019.103039.

Dunster, J.N., Mügge, A.E., 2001. Stream sediment survey of western MacDonnell Ranges – statistical and GIS-based interpretation. Northern Territory Geological Survey, Digital Information Package DIP 002.

Gadikota, G., Zhang, F. & Allen A., 2017. *In-Situ* Angstrom-to-Micrometer Characterization of the Structural and Microstructural Changes in Kaolinite on Heating using Ultra-Small-Angle, Small-Angle, and Wide-Angle X-ray Scattering (USAXS/SAXS/WAXS). Ind Eng Chem Res. 2017;56:10.1021/acs.iecr.7b02810. doi: 10.1021/acs.iecr.7b02810. PMID: 31092971; PMCID: PMC6512850.

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

Gallant, J., Wilson, N., Dowling, T., Read, A., & Inskeep, C., 2011. *SRTM-derived 1 Second Digital Elevation Models Version 1.0*. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/72759

Garousi Nezhad, S., Mokhtari, A. R., & Roshani Rodsari, P., 2017. The true sample catchment basin approach in the analysis of stream sediment geochemical data. *Ore Geology Reviews*, *83*, 127–134. https://doi.org/10.1016/j.oregeorev.2016.12.008

GDAL/OGR contributors, 2022. GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation. https://doi.org/10.5281/zenodo.5884351

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

Gray, D.J., Wildman, J.E. & 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.

Grunsky, E.C. & de Caritat, P., 2020. State-of-the-art analysis of geochemical data for mineral exploration. Geochemistry: Exploration, Environment, Analysis, 20: 217–232 https://doi.org/10.1144/geochem2019-03

Gu, X.X., Liu, J.M., Zheng, M.H., Tang, J.X., & Qi, L., 2002. Provenance and tectonic setting of the Proterozoic turbidites in Hunan. South China: geochemical evidence J. Sediment. Res., 72: 393-407

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

Halley, S., Dilles, J.H. & Tosdal, R.M., 2015. Footprints: Hydrothermal Alteration and Geochemical Dispersion Around Porphyry Copper Deposits, SEG Discovery, 100, 91–170. doi: https://doi.org/10.5382/SEGnews.2015-100.fea

Henne A., Noble R.R.P., Huang, F., Cole, D., Williams, M., Ibrahimi, T. & Lau, I., 2022a UltraFine+® Next Gen Analytics. Northern Territory Geological Survey – MacDonnell Ranges. CSIRO, Australia. CSIRO Report EP2022-2484, CSIRO, Australia.

Henne A., Noble, R.R.P., Huang, F., Cole, D., Williams, M., Ibrahimi, T., Lau, I. & Pejcic B., 2022b. UltraFine+® Next Gen Analytics. Geological Survey of New South Wales – Cobar Projects. CSIRO Report EP2022-3306, CSIRO, Australia.

Henne, A., Noble, R.R.P., Huang, F., Cole, D., Williams, M., Lau, I.C. & Ibrahimi, T., 2023a. UltraFine+® Next Gen Analytics. Geological Survey of Queensland – Jericho. CSIRO Report EP2023-0486, CSIRO, Australia

Henne, A., Noble, R.R.P. & Williams, M., 2023b, *accepted*. Multi-element geochemical analyses on ultrafine soils in Western Australia - Towards establishing abundance ranges in mineral exploration settings. *Geochemistry: Exploration, Environment, Analysis*.

Henne, A., Noble, R., Hutcheon, S., Cole, D., Huang, F., Williams, M., Lau, I., Pejcic, B., P, T. & Shelton, T., 2023c. Machine v Human – Testing a machine-learning derived landscape model in the field. Abstracts of the Australian Earth Science Convention. The Geological Society of Australia, Perth.

Heon, D., (compiler) 2003. Yukon Regional Geochemical Database 2003 in Arne, D. & 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

IAEA, 2003. Guidelines for radioelement mapping using gamma ray spectrometry data, International Atomic Energy Agency, IAEA-TECDOC-1363.

Jackson, M.L., 1964. VIII Soil clay mineralogical analysis. In Rich, C.I., and Kunze, G.W., (Eds). Soil Clay Mineralogy. University of North Carolina Press, Chapel Hill, N.C., USA. p.245-294.

Joyce, A.S., 1984. Geochemical exploration. Glenside, S. Aust: Australian Mineral Foundation.

Li, W. & Liu, X.-M. 2022. Mineralogy and fluid chemistry controls on lithium isotope fractionation during clay adsorption. Science of The Total Environment, Vol. 851, Part 1, 158138, ISSN 0048-9697. https://doi.org/10.1016/j.scitotenv.2022.158138.

Lindsay, J., 2023. *WhiteboxTools Open Core* (v2.3.0) [Computer software]. https://github.com/jblindsay/whitebox-tools

Lintern, M.J., 2002. Chapter 6: Calcrete sampling for mineral exploration, *In* Chen, X.Y., Lintern, M.J., Roach, I.C. (eds). Calcrete: characteristics, distribution and use in mineral exploration. Intant colour press, Belconnen, ACT. 160 p.

Martín-Fernández, J.A., Barceló-Vidal, C. & Pawlowsky-Glahn, V., 2003. Dealing with zeros and missing values in compositional data sets using nonparametric imputation. Mathematical Geology, 35, 253–278, https://doi.org/10.1023/A:1023866030544

Minasny, B. & McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Comput. Geosci. 32, 9: 1378–1388. DOI:https://doi.org/10.1016/j.cageo.2005.12.009

Neal, J. & Hawker, L., 2023. *FABDEM V1-2*. University of Bristol. https://doi.org/10.5523/BRIS.S5HQMJCDJ8YO2IBZI9B4EW3SN

Noble, R.R.P, 2012. Transported cover in northwestern Victoria, Australia - an impediment to geochemical exploration for gold. Journal of Geochemical Exploration, 112:139-151. https://doi.org/10.1016/j.gexplo.2011.08.006

Noble, R.R.P., Lau, I.C., Anand, R.R. & Pinchand, G.T., 2019. Refining fine fraction soil extraction methods and analysis for mineral exploration. Geochemistry: Exploration, Environment, Analysis, https://doi.org/10.1144/geochem2019-008

Noble, R.R.P., Lau, I.C., Anand, R.R. & Pinchand, G.T., 2020. Refining fine fraction soil extraction methods and analysis for mineral exploration. Geochemistry; Exploration, Environment, Analysis 20(1):113-128. https://doi.org/10.1144/geochem2019-008

Noble R.R.P., Henne, A., Williams, M., Cole, D., Huang, F. & Ibrahimi, T., 2023 unpublished. UltraFine+® Next Gen Analytics. Geological Survey of Western Australia – West Arunta. CSIRO Report draft.

Radiometric Grid of Australia (Radmap) http://dx.doi.org/10.26186/5dd48d628f4f6; http://dx.doi.org/10.26186/5dd48e3eb6367; http://dx.doi.org/10.26186/5dd48ee78c980

Rocklin , M., 2015. Dask: Parallel computation with blocked algorithms and task scheduling. In K. Huff & James Bergstra (Eds.), Proceedings of the 14th python in science conference (pp. 130–136).

Rose, A.W., Hawks, H.E. & Webb, J.H., 1979. Geochemistry in Mineral Exploration. Academic Press, New York.

Sadeghi, B., Yilmaz, H., & Pirajno, F., 2021. Weighting of BLEG data with drainage and catchment properties to enhance Au anomalies. *Geochemistry*, *81*(2), 125733. https://doi.org/10.1016/j.chemer.2020.125733

Sanford, R.F., Pierson, C.T. & Crovelli, R.A., 1993. An objective replacement method for censored geochemical data. Mathematical Geology, 25, 59–80, https://doi.org/10.1007/BF00890676

Smith, R.E. & Perdrix, J.L., 1983. Pisolitic laterite geochemistry in the Golden Grove massive sulphide district. Western Australia, J. Geochem. Explor., 18: 131-164.

Williams M.J. & Noble R.R.P., 2021. Grey Dam & Mt. Alexander Ultrafine+ Kickstart Report: Utilizing clays and iron oxide geochemistry to explore for battery metal orebodies in the Kurnalpi region of Western Australia. CSIRO, Australia. 21 p.

Williams, M., Cole, D., Huang, F., Henne, A. & Noble, R., 2023. New Horizons for UltraFine+® Analytics: Enhancing Tools for Landscape Context. Abstracts of the Australian Earth Science Convention. The Geological Society of Australia, Perth.

Yang, S.Y, Jung, H.S., & Li, C., 2004a Two unique weathering regimes in the Changjiang and Huanghe drainage basins: geochemical evidence from river sediments Sediment. Geol., 164: 19-34

Yang, S.Y., Li, C.X., Yang, D.Y., & Li, X.S, 2004b. Chemical weathering of the loess deposits in the lower Changjiang Valley, China, and paleoclimatic implications Quat. Int., 117: 27-34

# 6 Appendices

A) UltraFine Next Gen data package example from the Geological Survey of Western Australia containing the outputs of the final version of the machine learning workflow created during the M0462a project.

   Please note that this data package does not contain vis-NIR or FTIR analyses as these are not available for this site. To view examples of these analyses please refer to earlier published reports for the Geological Survey of New South Wales (EP2022-3306), Queensland (CSIRO Report EP2023-0486) or the Northern Territory Geological Survey published (CSIRO Report EP2022-2484).