



Australia's National
Science Agency

UltraFine+[®] Next Gen Analytics

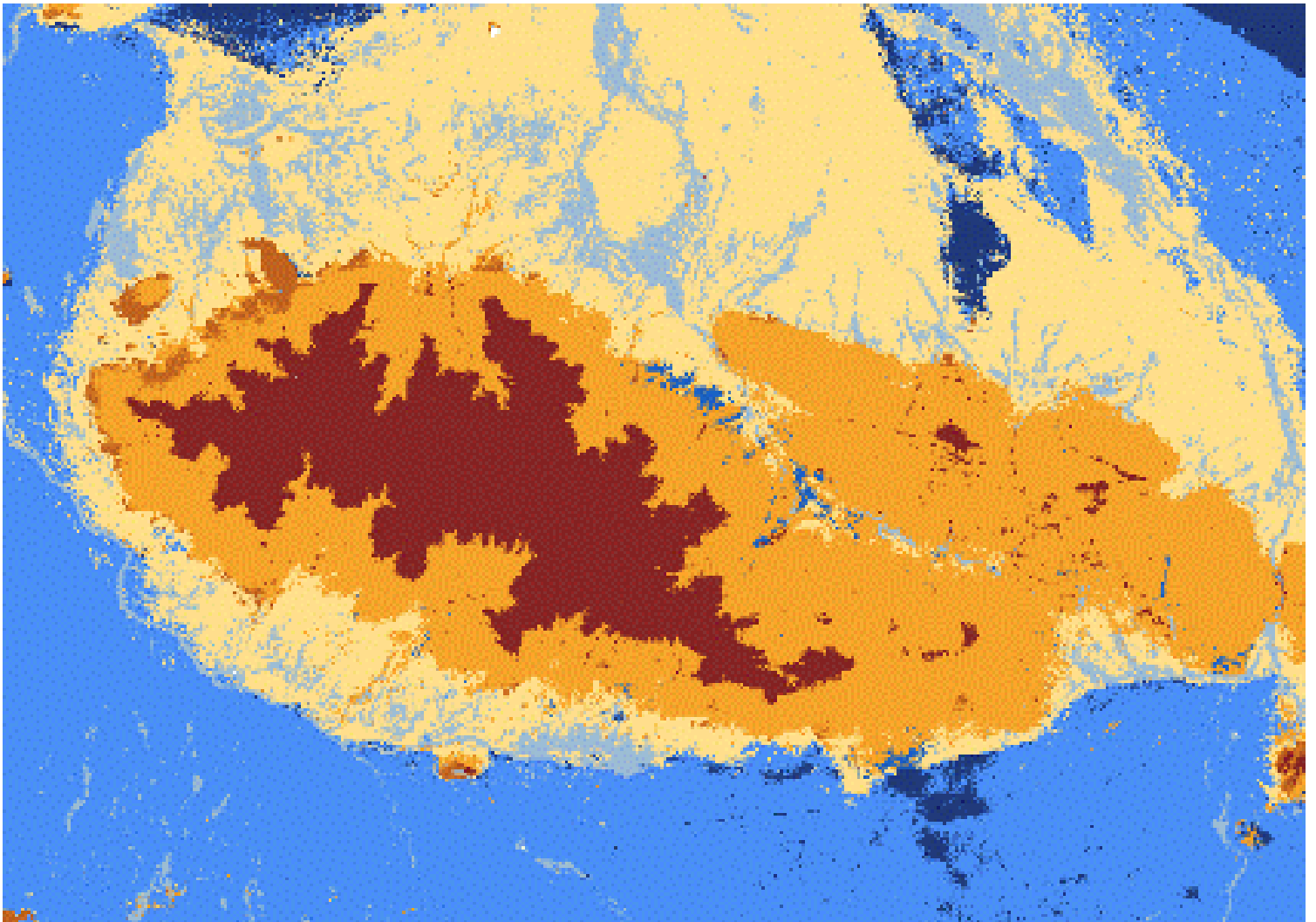
Northern Territory Geological Survey – MacDonnell Ranges

Anicia Henne, Ryan R P Noble, Fang Huang, Dave Cole, Morgan Williams, Tania Ibrahimi, Ian Lau

EP2022-2484

22 July 2022

Northern Territory Geological Survey



Citation

Henne A, Noble RRP, Huang F, Cole D, Williams M, Ibrahimi T, Lau I (2022) UltraFine+® Next Gen Analytics. Northern Territory Geological Survey – MacDonnell Ranges. CSIRO, Australia.

Copyright

© Commonwealth Scientific and Industrial Research Organisation 2022. To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

Important disclaimer

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

CSIRO is committed to providing web accessible content wherever possible. If you are having difficulties with accessing this document, please contact csiroenquiries@csiro.au.

UltraFine+® Next Gen Analytics Sponsors



Barton Gold



UltraFine+® Next Gen Analytics for Discovery

The Northern Territory Geological Survey's MacDonnell Ranges project is part of the UltraFine+® Next Gen Analytics research project conducted by the CSIRO in collaboration with LabWest and over 20 industry sponsors and geological surveys. The aim of this research project is to facilitate a paradigm shift for precious, base and critical metals exploration in Australia by combining the UltraFine+® soil analysis method with intelligent data integration tools, adding value to routine soil sampling in frontline exploration and shaping mineral exploration approaches for decades.

It has been common practice to use soil geochemistry with little regard for physicochemical soil parameters or landform settings and how these relate to buried mineralisation. The UltraFine+® Next Gen Analytics research project addresses this challenge by delivering an analytical refinement of the UltraFine+® soil analysis method and adding relevant mineral proxies via spectral mineralogy and the soil properties pH, EC and particle size distribution to the workflow and interpretation. UltraFine+® Next Gen Analytics utilises machine learning approaches to integrate these soil parameters with spatial data and regolith landform models. This improves our ability to identify targets and false positives as well as understand the spatial variance and influence of regolith types. With the development of a robust set of measurable parameters and new data products to fully assess underappreciated soil properties, UltraFine+® Next Gen Analytics provides the latest iteration of analytical tools for mineral explorers using soil samples to make qualified decisions on when and where to explore further.

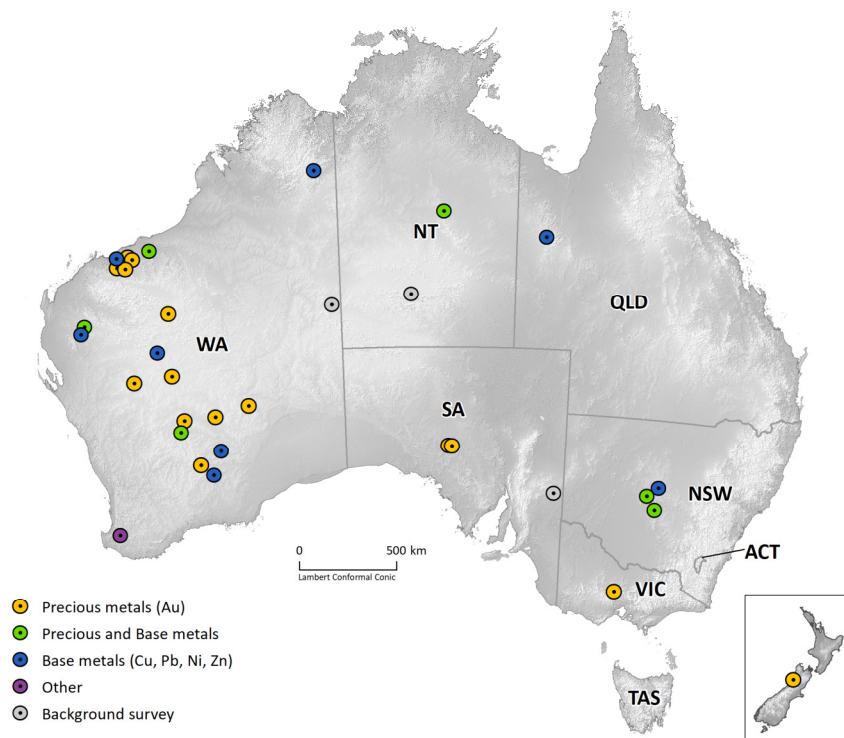


Figure 1: UltraFine+® Next Gen Analytics for Discovery project sponsors' site locations (as of 18 May 2022).

Contents

UltraFine+® Next Gen Analytics Sponsors.....	ii
UltraFine+® Next Gen Analytics for Discovery	i
Acknowledgments.....	vii
Executive summary	viii
1 The UltraFine+® Next Generation Analytics for Discovery Research project	1
1.1 Northern Territory Geological Survey – MacDonnell Ranges project.....	1
2 The UltraFine+® Next Gen Analytics workflow.....	4
2.1 Sample collection	4
2.2 UltraFine+® laboratory soil analyses	4
2.3 Automated QAQC.....	6
2.4 Machine Learning - Spatial data integration and clustering.....	7
3 The UltraFine+® Next Gen Analytics outputs	11
3.1 Landscape clusters	11
3.2 Comparison of UltraFine+® to traditional soil survey analyses	18
3.3 Outliers by landscape type	25
3.4 Exploration Indices	31
3.5 Other soil properties	34
4 Summary.....	38
Appendix A	40
Appendix B	41
Appendix C	42
References	43

Figures

Figure 1: UltraFine+® Next Gen Analytics for Discovery project sponsors' site locations (as of 18 May 2022).....	i
Figure 2: MacDonnell Ranges project area (dashed box), sample locations and known mineral occurrences (Northern Territory Mineral Occurrences) near Alice Springs in the Northern Territory, Australia. Samples submitted for this project were part of a stream sediment survey carried out in 2001 by the Northern Territory Geological Survey (Dunster and Mügge 2001).....	2
Figure 3: Surface Geology (Raymond et al. 2012) and Regolith Geology (Craig 2006) in the MacDonnell Ranges project area. Refer to Appendix A for surface geology legend.	3
Figure 4: Jitter-box plot of pixel values embedded in a 3-dimensional latent space (a representation of data compressed to 3 dimensions, in which similar data points are closer together in space) using the dimensionality reduction algorithm UMAP (Uniform Manifold Projection and Approximation). Points are coloured using an RGB of the axes (u0, u1, u2) values.....	8
Figure 5: Clustered data points in the 3-dimensional latent space grouped by cluster colour to visualise data separation. (A) Data points clustered into four clusters using the k-means algorithm (k-means4). (B) Data points clustered into eight clusters by an agglomerative clustering algorithm (agg8). Refer to Appendix B for interactive plot.	10
Figure 6 (previous page): Proxy regolith clustering outputs (A, B), input layers (C-F), and comparison to traditional geological maps (G, H). (A) Proxy regolith types derived via clustering algorithm agg8, plotted in spatial context. (B) Proxy regolith types derived via clustering algorithm k-means4, plotted in spatial context. White line in E and F is a stitch line between two sentinel 2 satellite image tiles. For subsequent models the workflow will utilise bare earth Sentinel data avoiding such artifacts. (C) 1-second DEM SRTM. (D) Continuous MrVBF. (E) RGB image of radiometric grid of Australia. (F) RGB image of sentinel-2 regolith ratios. (G) Surface Geology (H) Regolith Geology.....	13
Figure 7: Correlation of proxy regolith types with landscape clusters generated by an agglomerative clustering algorithm with 8 clusters. Please refer to the text for more details on proxy regolith types.	13
Figure 8: Assessment of landscape clusters indicative of exposed outcrops and residual 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), radiometrics (G) and regolith ratio (H).....	14
Figure 9: 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), radiometrics (G) and regolith ratio (H).....	15
Figure 10: Assessment of landscape clusters indicative of active alluvial channels 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), radiometrics (G) and regolith ratio (H)..... 16

Figure 11: 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), radiometrics (G) and regolith ratio (H)..... 17

Figure 12: Comparison of landscape clusters 2, 4 and 6 with weathering intensity over the MacDonnell Ranges project area. (A) Landscape cluster types derived via the clustering algorithm agg8 plotted in spatial context. (B) Weathering intensity (Wilford and Roberts 2019) derived from radiometrics and DEM. 17

Figure 13: Boxplots of analyses for Au, Cu, Zn and Pb. The 2001 survey results are displayed in red and the 2020 results from re-analysis via UltraFine+® are displayed in green..... 21

Figure 14 (previous page): Comparison of the NTGS’s 2001 survey and the re-analysed results in 2020 displayed on the same scale. (A) Spatial distribution of Au analysed via 24-hour cyanide leach. (B) Spatial distribution of Au analysed via UltraFine+®. (C) Spatial distribution of Cu analysed via four-acid digest. (D) Spatial distribution of Cu analysed via UltraFine+®. (E) Spatial distribution of Pb analysed via four-acid digest. (F) Spatial distribution of Pb analysed via UltraFine+®. (G) Spatial distribution of Zn analysed via four-acid digest. (H) Spatial distribution of Zn analysed via UltraFine+®. 23

Figure 15 (next page): Comparison of the NTGS’s 2001 survey and the re-analysed results in 2020 displayed on scales with natural breaks within each dataset. (A) Spatial distribution of Au analysed via 24-hour cyanide leach. (B) Spatial distribution of Au analysed via UltraFine+®. Note that Au values below the detection limits have been replaced with half the detection limits (0.005 in the 2001 survey data and 0.25 in the 2020 UltraFine+® data). (C) Spatial distribution of Cu analysed via four-acid digest. (D) Spatial distribution of Cu analysed via UltraFine+®. (E) Spatial distribution of Pb analysed via four-acid digest. (F) Spatial distribution of Pb analysed via UltraFine+®. (G) Spatial distribution of Zn analysed via four-acid digest. (H) Spatial distribution of Zn analysed via UltraFine+®. 23

Figure 16: Ni outliers by landscape cluster over the MacDonnell Ranges project area. (A) Boxplots for all data (white box) and by landscape type (coloured boxes). Dashed mauve line indicates the upper 25 % boundary for the whole sample population. Commonly observed soil anomalies are samples above the dashed line (white triangles). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)). (B) Spatial distribution of Ni outliers (triangles) by proxy regolith type. White dashed boxes indicate outliers by landscape type below the dashed mauve line in (A). 26

Figure 17: Comparison of Au outliers in the whole sample population to Au outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Au outliers for all data. (B) Spatial distribution of Au outliers by landscape cluster. Background value samples plotted in grey. The known Au-related mineral occurrence is labelled. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). Commonly observed soil anomalies are samples above the dashed horizontal line (white triangles in (A)). Those

shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)).27

Figure 18: Comparison of Bi outliers in the whole sample population to Bi outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Bi outliers for all data. (B) Spatial distribution of Bi outliers by landscape cluster. Background value samples plotted in grey. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). Commonly observed soil anomalies are samples above the dashed horizontal line (white triangles in (A)). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)).....28

Figure 19: Comparison of Zn outliers in the whole sample population to Zn outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Zn outliers for all data. (B) Spatial distribution of Zn outliers by landscape cluster. Background value samples plotted in grey. The known Zn-related mineral occurrences are labelled. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). Commonly observed soil anomalies are samples above the dashed horizontal line (white triangles in (A)). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B))......29

Figure 20: Comparison of Pb outliers in the whole sample population to Pb outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Pb outliers for all data. (B) Spatial distribution of Pb outliers by landscape cluster. Background value samples plotted in grey. The known Pb-related mineral occurrences are labelled. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). Commonly observed soil anomalies are samples above the dashed horizontal line (white triangles in (A)). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)).30

Figure 21: Comparison of Cu outliers in the whole sample population to Cu outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Cu outliers for all data. (B) Spatial distribution of Cu outliers by landscape cluster. Background value samples plotted in grey. The known Cu-related mineral occurrences are labelled. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). No outliers by landscape type were identified.31

Figure 22: Elemental loadings for each of the first five principal components for the MacDonnell Ranges project. The further away an element plots from the 0 line, the greater the loading for (influence on) the specific principal component.....32

Figure 23: Automated output of the spatial distribution of principal components weighted by both colour and symbol size (absolute magnitude). The top five elemental loadings (highest influence) for each principal component are indicated as headings. The colour red indicates a positive component weight (association); the colour blue indicates a negative component weight. The larger the symbols the stronger the association. (A) Spatial distribution of principle component 0 weightings. (B) Spatial distribution of principle component 1 weightings. (C) Spatial distribution of principle component 2 weightings. (D) Spatial distribution of principle component 3 weightings. (E) Spatial distribution of principle component 4 weightings.33

Figure 24: Spatial distribution of principal component 2 coincides broadly with known Au and Zn-Cu-Pb mineral occurrences Haasts Bluff, Ulparuta, Mount Larrie and Stokes Yard, but not with Glen Helen II and Glen Helen III.34

Figure 25: Variation in sample pH (A) appears to broadly correlate with the variation of underlying parent material captured in principal component 1 (B).35

Figure 26: Spatial distribution of VNIR results over the MacDonnell Ranges project area. (A) Primary mineral group (contributes >51 % to the spectral unmixing algorithm). (B) Secondary mineral group. (C) Relative iron oxide abundance. (D) Relative kaolinite abundance. (E) Kaolinite crystallinity.36

Figure 27: Example of normalising geochemical data with VNIR analyses. (A) Relative iron oxide abundance. (B) Spatial distribution of Zn concentrations. (C) Spatial distribution of Zn normalised to iron oxides. Dashed white boxes indicate areas of interest.37

Tables

Table 1: Number of samples analysed with the UltraFine+® workflow for the MacDonnell Ranges project. *Added to the workflow after MacDonnell Ranges Project data acquisition.....5

Table 2: Geospatial covariates used for landscape clustering.8

Table 3: Comparison between the 2001 analyses and the UltraFine+® method in 2020 with available analyses, detection limits and Pearson’s correlation coefficient *r*. All analyses are reported in ppm, except for Au and Pt which are reported in ppb. Since the data acquisition in 2020, the element Pd has been added to the UltraFine+® element suite.18

Acknowledgments

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

Most critically, we thank the in-kind support of the above contributors who provided UltraFine+[®] analytical results of 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

Assessing geochemical data in mineral exploration often focuses on understanding outliers, such as elevated Au, Cu or Zn. Commonly the largest concentrations are followed up, but landscape (soil) types can significantly influence concentrations. For example, high metal concentrations may be readily identifiable as outliers in a geochemical dataset where samples were collected over mineralisation in shallow residual soils, while the same mineralisation would have a much weaker elemental signal in samples collected over thicker depositional landscapes. With the ability to approximate landscape types from spatial data via machine learning, the UltraFine+® Next Gen Analytics workflow was developed to improve outlier identification within landscape types. As part of the UltraFine+® Next Gen Analytics for Discovery research project, the MacDonnell Ranges project site was chosen as a first-generation, large-scale trial site with a focus on principal functionality of the UltraFine+® Next Gen Analytics workflow. For this purpose, the Northern Territory Geological Survey provided 785 historic stream sediment samples, and we present some example outputs of the workflow here.

These outputs include proxy regolith landscape clusters to provide context for geochemical samples, maps and boxplots of elemental outliers by landscape type, and exploration indices for a rapid, first-pass identification of element association and potential exploration indices. The data package also contains geochemistry, VNIR (visible to near infra-red spectral mineralogy) and pH as shapefiles. This provides a basic, first-pass interpretation of geochemical samples by proxy regolith type and the identification of otherwise “overlooked” potential anomalies.

The data presented herein was analysed in September 2020. Since then, the components of the UltraFine+® workflow have undergone continuous improvements especially with regards to the consistency of pH and VNIR (visible near-infrared spectroscopy) measurements, as well as the addition of soil sizing, FTIR (Fourier-transform infrared spectroscopy) and Pd analyses. Therefore, some analysis results and related outputs, such as soil texture diagrams, dispersion direction, regolith indices and catchment analyses are not available for this project. However, where data was available, it was reprocessed with the latest workflow updates in February and April 2022.

A variety of clustering methods were trialled to generate appropriate proxies for regolith types and the recommended outputs for the MacDonnell Ranges project are those produced via an agglomerative algorithm with eight landscape clusters (agg8). Due to the large size of the MacDonnell Ranges project area and the resulting complexity of the landscape an even larger number of clusters would result in a more detailed approximation of the complex regolith types. However, the number of landscape clusters are the result of a balanced approach to represent the major landscape types of the area while also enabling meaningful interpretation of geochemical data. Despite these constraints, the resulting output for the MacDonnell Ranges project area provides a more detailed landscape context than publicly available, interpreted regolith products.

The samples submitted to the research project were previously analysed via four-acid digestion and a 24 hour cyanide leach (Au only). A comparison between these methods and the UltraFine+® method confirmed that the UltraFine+® method improves the recovery of trace metals (2 to >5 times higher for Au, Cu, Pb and Zn) and the resolution of concentrations near the detection limit,

which enables the delineation of subtle geochemical enrichments for elements. The detection of these subtle variations is particularly relevant for exploration through transported cover and the results of this project provide a good background data set for future exploration activities in the region and surrounding areas.

While the UltraFine+® Next Gen Analytics workflow is designed for soil samples and not presently developed for stream sediment analysis and interpretation, the workflow provides a general, large-scale landscape context of the catchment within which the stream sediment samples were collected, and the MacDonnell Ranges project site was crucial in developing the approach, workflow and outputs for the UltraFine+® Next Gen Analytics for Discovery research project.

1 The UltraFine+® Next Generation Analytics for Discovery Research project

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+® workflow which is optimised for multielement analysis of this ultrafine soil fraction. This improved soil geochemistry workflow generates results with more contrast and increased concentrations of Au, Cu and Zn, and removes the nugget effect (coarse grains of gold in a sample leading to the overestimation of the average metal content), thereby enhancing reproducibility and reliability of results (Noble et al. 2020).

The UltraFine+® Next Gen Analytics research project leverages and expands on this workflow by adding relevant soil parameters including spectral mineral proxies, pH, EC and particle size distribution to the UltraFine+® workflow. This provides a wealth of additional data to the standard soil sample exploration package which enables exploration geologists to investigate the relationships between soil geochemistry and other physicochemical soil parameters and how these relate to buried mineralisation. In addition, the UltraFine+® Next Gen Analytics workflow utilises machine learning approaches to produce landscape context for, and first-pass data interpretation of, these soil sample analyses.

The components of the UltraFine+® workflow are undergoing continuous improvements over the course of the UltraFine+® Next Gen Analytics research project (conducted from April 2020 to April 2023) which is reflected in improved detection limits and refined outputs. The data presented herein was analysed in September 2020. However, for this report, the VNIR spectra were reprocessed with the latest VNIR TSG™-processing template update from November 2021 and the UltraFine+® Next Gen Analytics outputs presented in this report were re-processed in February 2022.

1.1 Northern Territory Geological Survey – MacDonnell Ranges project

The Northern Territory Geological Survey's (NTGS) MacDonnell Ranges project covers an area of approximately 21,000 km² over the West MacDonnell Ranges and extends for some 50 km just west of Alice Springs in the Northern Territory, Australia (Figure 2). The NTGS submitted 785 samples with an average spacing of one sample per 5.6 km² as part of the UltraFine+® Next Gen Analytics research project. These samples were part of a larger background study on historical stream sediments (Dunster and Mügge 2001) and are not intended for detailed exploration. While the UltraFine+® Next Gen Analytics workflow is designed for soil samples and not presently developed for stream sediment analysis and interpretation, a detailed surface geology map

generated by the NTGS exists over the area, and the MacDonnell Ranges project site was chosen as a first-generation, large-scale trial site with a focus on principal functionality of the UltraFine+® Next Gen Analytics workflow rather than targeting tangible exploration outcomes. For the MacDonnell Ranges project, the workflow provides a general, large-scale landscape context of the catchment within which the stream sediment samples were collected.

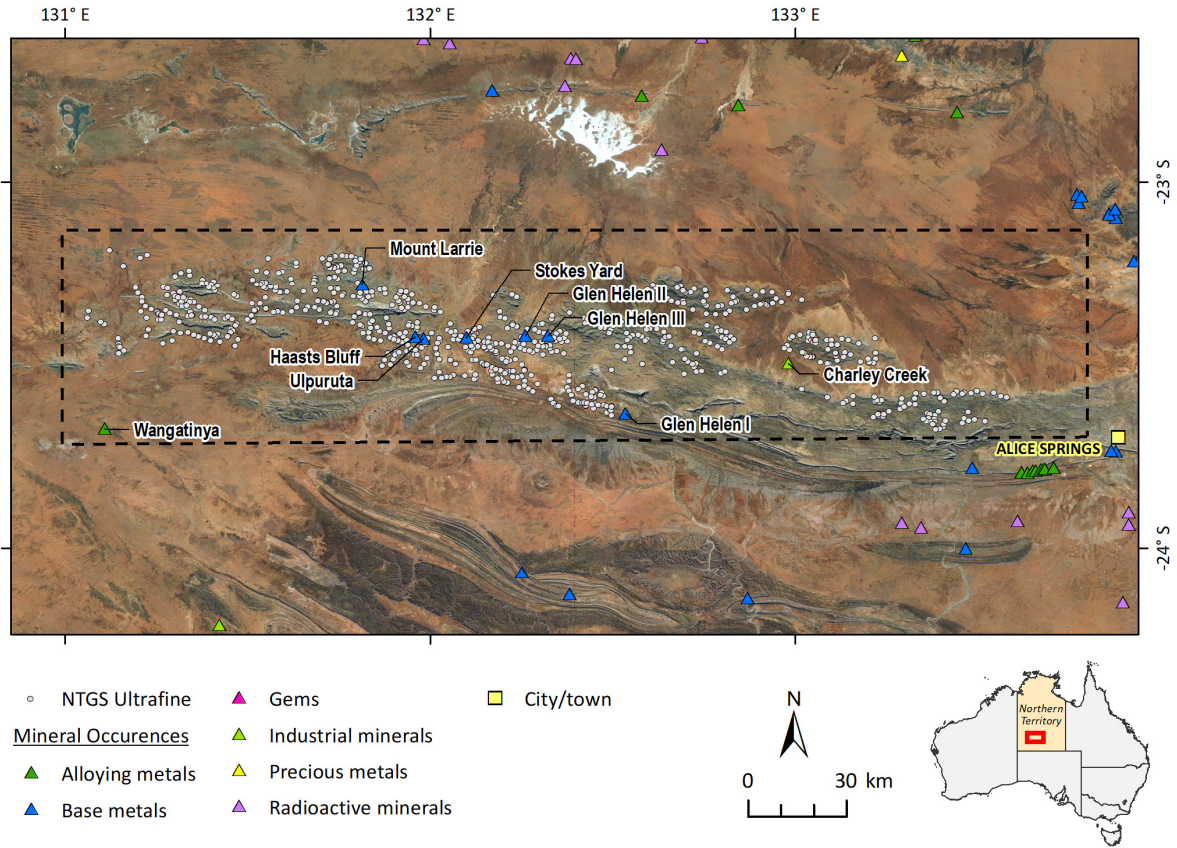


Figure 2: MacDonnell Ranges project area (dashed box), sample locations and known mineral occurrences (Northern Territory Mineral Occurrences) near Alice Springs in the Northern Territory, Australia. Samples submitted for this project were part of a stream sediment survey carried out in 2001 by the Northern Territory Geological Survey (Dunster and Mügge 2001).

The MacDonnell Ranges project area is situated in the arid parts of central Australia and is dominated by grassland with hot, dry summers and cold winters, an average annual rainfall of ≈ 280 mm and an average annual evapotranspiration rate between 300 to 400 mm/year. The mean minimum and maximum temperatures are 28.9 and 32.2°C, respectively (Bureau of Meteorology 2022). The groundwater ranges from mainly fresh and marginal in the northwest to mainly brackish to saline in the southeast of the project area, with generally circum-neutral pH (Gray et al. 2019; Bardwell and Gray 2016).

The study area comprises parts of the West MacDonnell Mountain ranges, prominent parallel ridges of quartzite and sandstone with a maximum elevation of 1,531 m. Detailed geological mapping has been carried out by the NTGS (Figure 3A; Warren and Shaw 1995; Craig 2006). The MacDonnell Ranges outcrop as variably weathered, *in situ* bedrock (saprolite and saprock) within the project area (Figure 3B). Transported cover is dominated by aeolian sediments in the north

and west of the project area that are characterised by quartzose sands. Alluvial sediments dominate the east of the project area as overbank and channel deposits on alluvial plains and floodplains (Figure 3B).

The major soil types are dominated by tenosols and rudosols, with smaller areas of kandosols and chromosols based on the Australian Soil Classification System (Isbell, 2021). These soils are very common for the arid interior of Australia. Commonly, these soils have very weak (B) horizon formation with the exception of the kandosols and chromosols that have some clay formation and evident pedogenesis. However, it is important to note that the samples analysed in this research project are stream sediments and are a composite (loose, alluvial) mix of soil types.

Historic sampling and minor auger and drilling campaigns identified several mineral occurrences in the Tjoritja West MacDonnell Ranges National Park. These include the base metal occurrences Glen Helen I (carbonate-hosted Pb-Zn-Cu), Glen Helen II and Glen Helen III (vein-hosted Cu-Pb-Zn-Ag), Stokes Yard (pegmatitic Zn-Cu-Pb-Ag), Mount Larrie (pegmatitic Cu veins), Haasts Bluff (Fe-oxide associated Cu-Au) and Ulpuruta (vein hosted Zn-Cu-Pb with associated oxidised secondary minerals; variably referred to as Haasts Bluff I or Nickel Hill) as well as a REE mineral occurrence at Charley Creek and sedimentary Mn at Wangatinya (Northern Territory Mineral Occurrences) (Figure 2).

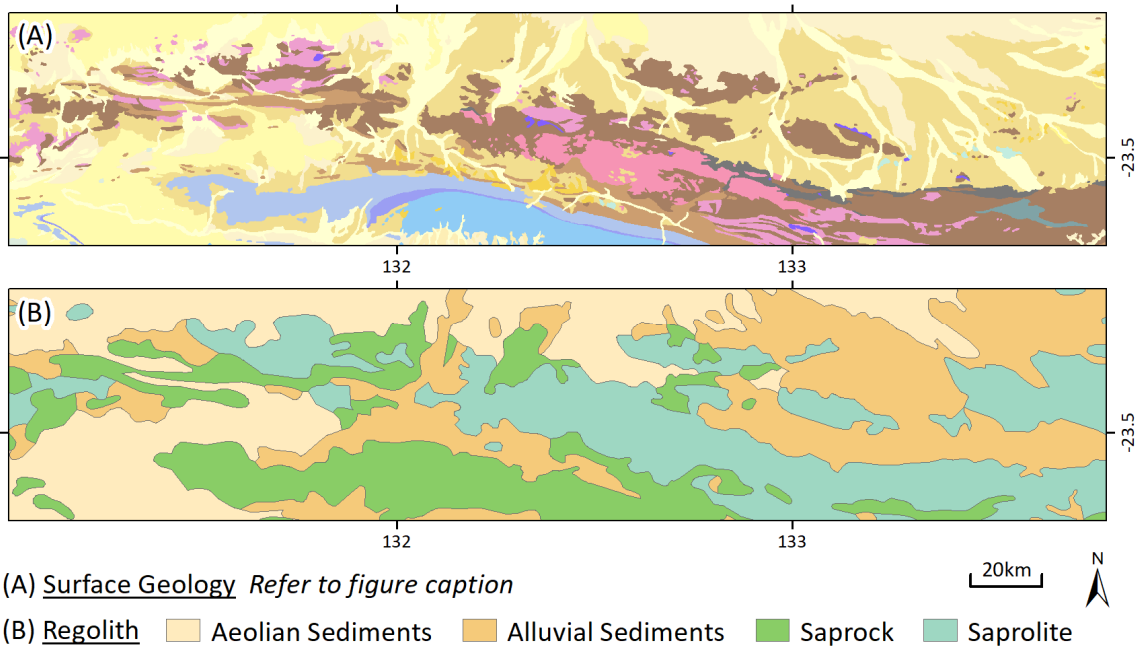


Figure 3: Surface Geology (Raymond et al. 2012) and Regolith Geology (Craig 2006) in the MacDonnell Ranges project area. Refer to Appendix A for surface geology legend.

2 The UltraFine+[®] Next Gen Analytics workflow

The UltraFine+[®] Next Gen Analytics workflow is based on the UltraFine+[®] soil analysis method, previously developed by the CSIRO in cooperation with LabWest (Noble et al. 2020). The main focus of the UltraFine+[®] Next Gen Analytics project is to provide landscape context for the geochemical data acquired with the UltraFine+[®] soil analysis method and a basic, first-pass data interpretation by incorporating semi-automated machine learning into the workflow. In addition, the UltraFine+[®] Next Gen Analytics project aims to improve the UltraFine+[®] method by lowering detection limits and adding valuable soil properties such as spectral mineralogy to the standard analyses.

2.1 Sample collection

Soil samples were collected by the Northern Territory Geological Survey as part of a stream sediment survey conducted in 2000 (Dunster and Mügge 2001). These historic samples had been sieved to <6.5 mm prior to submission to this research project. Analyses for 791 samples, from two Laboratory Jobs (ALW005556 and ALW005556A) were received by CSIRO. Of these, six samples were geochemical standards (QC_UFF_320), 7 samples were excluded from further analyses due to missing coordinates, and geochemical results were not available for a further four samples. Coordinates for another 12 samples were submitted to the CSIRO, but no analyses were available either due to missing samples or potentially due to samples with size fractions that were too coarse to recover ultrafine material. These samples were not included (refer to NTGS_MacDonnell_FinalData.xlsx in Appendix B).

2.2 UltraFine+[®] laboratory soil analyses

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 EC measurements, separation of the ultrafine (<2 µm) size fraction, elemental analyses, and spectral reflectance mineralogy. The components of the UltraFine+[®] workflow are undergoing continuous improvements over the course of the UltraFine+[®] Next Gen Analytics research project, which is reflected in improved detection limits and refined outputs over time. The data presented herein was analysed in September 2020. However, VNIR spectra were reprocessed with TSG™ 3.1 (version 3.1 of the VNIR The Spectral Geologist-processing template) for this report, which is the latest workflow update from November 2021.

2.2.1 Bulk soil properties

Electrical conductivity (EC) 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 Soil sizing

No sizing data was collected for the MacDonnell Ranges Project.

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 (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:HNO₃. 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: Number of samples analysed with the UltraFine+® workflow for the MacDonnell Ranges project. *Added to the workflow after MacDonnell Ranges Project data acquisition.

ANALYSES	NUMBER OF SAMPLES	OUTPUTS
Microwave-assisted aqua regia on ultrafine fraction (LabWest MAR-04)	775	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
Other bulk soil properties	779	EC, pH
Particle size distribution on bulk sample (LabWest SIZE-01)	Not analysed	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
Other bulk soil properties	779	EC, pH
Visible-near-infrared (VNIR) on ultrafine fraction (LabWest VNIR/SWIR)	778	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; Munsell colour; hue; saturation; intensity
Fourier transform infrared spectroscopy (FTIR)* on ultrafine fraction (LabWest FTIR)	Not analysed	Clay, quartz, and carbonate abundances; total organic carbon; gibbsite index

2.2.5 Visible near-infrared reflectance (VNIR)

Visible near-infrared reflectance measurements were acquired on the ultrafine fraction (< 2 µm) using a Panalytical ASD FieldSpec4 standard-resolution spectroradiometer (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 32.8 nm at 700 nm, 8 nm at 1500 nm and 610 nm at 1400 and 2100 nm. The spectral bandwidths of the FieldSpec4 for the RS-3500 are 1.43 nm at 350–1000 700 nm, 3.5 nm at 1500 nm and 1.12.3 nm at 1001– 21500 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 410 scans, averaged into a single measurement. Spectra were processed using The Spectral Geologist (TSG™) software to extract the main features reported as part of the UltraFine+® output (Table 1).

Final data for NTGS was reprocessed with TSG™ version 3.1 of the VNIR TSG-processing template and may differ to previously reported, unrefined outputs.

2.2.6 Fourier transformed infrared spectroscopy (FTIR)

No FTIR data was available for the MacDonnell Ranges project.

2.3 Automated QAQC

The UltraFine+® Next Gen Analytics workflow is developing automated QAQC on standards and duplicates for all available analyses, whereby analysis batch quality and internal reproducibility as recorded by standard and duplicate measurements can be assessed via a “traffic light system” which indicates the level of accuracy (standards) and precision (standards and duplicates) of the analyses. The QAQC automation requires consistent formatting of input data which was not available at the time the MacDonnell Ranges data was submitted. Therefore, QAQC was carried out manually for this project.

2.3.1 Standards

The incorporation of the CSIRO UltraFine+® reference material QC_UFF_320 (Appendix C) is routinely recommended every 50 samples for all UltraFine+® Next Gen analysis batches (MAR-04, VNIR/SWIR and FTIR) excepting SIZE-01, to assess accuracy. The MacDonnell Ranges project analyses included 6 QC_UFF_320 for 785 analysed samples.

The analyses for QC_UFF_320 slightly underestimated Cu and Ni in batch ALW005556A and slightly overestimated these elements in batch ALW005556. More resistate elements (Ti, Zr, W, Mo, Sn) are expected to show higher than average variance, but are generally within acceptable variation ranges (refer to NTGS_MacDonnell_FinalData.xlsx in Appendix B).

2.3.2 Duplicates

The incorporation of field duplicates is routinely recommended every 20 samples for all UltraFine+® Next Gen analyses (MAR-04, VNIR/SWIR, FTIR, SIZE-01) to assess precision.

No duplicates were supplied for the MacDonnell Ranges project.

2.4 Machine Learning - Spatial data integration and clustering

The UltraFine+® Next Gen Analytics workflow applies unsupervised dimensional reduction and clustering methods to spatial data to derive proxy landscapes. These and a range of other spatial outputs are designed to aid a first pass interpretation of soil analysis results in a broader landscape context.

2.4.1 Spatial data clustering

The UltraFine+® Next Gen Analytics workflow uses semi-automated, unsupervised machine learning methods to cluster publicly available spatial data to produce proxy regolith types. The spatial data layers used in this workflow were specifically selected for their relationship to regolith landforms and provide information on landscape position, depth of transported cover and parent material, while minimising the introduction of human interpretation, such as would be the case by including surface geology or regolith landscape type maps. Spatial data layers included in the workflow are Digital Elevation Model Shuttle Radar Topography Mission (DEM SRTM, Multi-resolution Valley Bottom Flatness (MrVBF), Radiometric K (ppm) %, Th (ppm) %, U (ppm) %, and Sentinel-2 satellite data (Table 2). Sentinel-2 satellites, launched via EU Copernicus Program's Earth observation mission, collect high-resolution imagery data of the Earth every 3-5 days (see Digital Earth Australia data product documentation). The Sentinel-2 data used in this workflow is cropped from the Sentinel 2A Analysis Ready Data based on the soil sample locations. Eleven out of thirteen spectral bands are collected: B01, 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 Wollrych and Batty, 2007; see USGS comparison of band equivalence here), with respective RGB bands calculated using $R=B11/B12$, $G=B08/B12$ and $B=B08/B03$.

The total area modelled for the MacDonnell Ranges project was 20,953 km² and all spatial data layers were re-gridded to the highest input layer resolution (20 m). The dimensionality reduction algorithm UMAP (Uniform Manifold Projection and Approximation; McInnes et al. 2018) was applied to transform the data to a three-dimensional latent space to aid more efficient clustering and provide a framework for visualisation. The UMAP algorithm was used as it captures non-linearities within the data and preserves local structures. 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.

Table 2: Geospatial covariates used for landscape clustering.

GEOSPATIAL COVARIAT	LANDSCAPE INFORMATION	RESOLUTION [M]	DATA SOURCE
DEM SRTM	Landscape Position	29	Geoscience Australia Gallant, J., Wilson, N., Dowling, T., Read, A., Inskip, C. 2011. SRTM-derived 1 Second Digital Elevation Models Version 1.0. Record 1. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/72759
MrVBF*	Depth of transported cover	86	CSIRO 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 ppm	Parent material	104	Geoscience Australia Poudjom Djomani, Y., Minty, B.R.S. 2019. 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	104	Geoscience Australia Poudjom Djomani, Y., Minty, B.R.S. 2019. 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	104	Geoscience Australia Poudjom Djomani, Y., Minty, B.R.S. 2019. 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	20	Digital Earth Australia https://docs.dea.ga.gov.au/notebooks/DEA_datasets/Sentinel_2.html

*At the time of modelling the MacDonnell Ranges project area, the UltraFine+® Next Gen Analytics workflow used integer MrVBF. Subsequent comparison of multiple project sites and parameters concluded that the future workflow will utilise continuous (float-valued) MrVBF for improved outcomes. In addition, future outputs will assign landscape cluster colours based on MrVBF-inferred depositional depth from dark brown (relatively shallow/outcrop) to dark blue (relatively deeper cover).

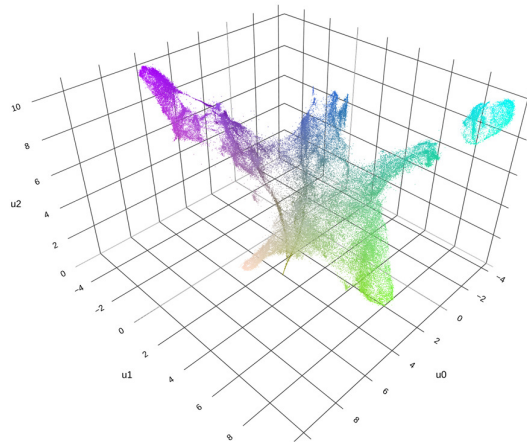


Figure 4: Jitter-box plot of pixel values embedded in a 3-dimensional latent space (a representation of data compressed to 3 dimensions, in which similar data points are closer together in space) using the dimensionality reduction algorithm UMAP (Uniform Manifold Projection and Approximation). Points are coloured using an RGB of the axes (u0, u1, u2) values.

Two clustering algorithms are used in the UltraFine+® Next Gen Analytics workflow to cluster locations with similar spatial data signatures: k-means and agglomerative (hierarchical) clustering. Both algorithms force all sample points into a given number of 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 area sizes (1 km² to 50,000 km²) the number of clusters has been predetermined for the UltraFine+® Next Gen Analytics workflow. Four clusters are routinely used for k-means (for project areas <20 km²) and eight clusters are routinely used for agglomerative clustering (for project areas >20 km²).

K-means is an iterative clustering algorithm that randomly selects a cluster centre from a dataset to compute distances of all datapoints from this selected centre from which it calculates a new centre point. This is repeated until four centroids (geometric centre) with maximum distance from each other have been sampled. All data points are subsequently assigned to the closest cluster centre based on their Euclidean distance. The algorithm then calculates the average of all points in each cluster and moves the centre point to this location.

Agglomerative clustering is a hierarchical clustering algorithm that successively merges individual data points. The UltraFine+® Next Gen Analytics workflow agglomerative clustering model uses Ward linkage criterion (scikit-learn default) to determine the distance metrics between clusters for merging. Due to the nature of the hierarchical clustering algorithm, training an agglomerative clustering model uses large computational resources for sites with a large area size, such as the MacDonnell Ranges site, and the trained model could not be used for predicting new samples. Therefore, we used 1) a subset of the samples to train an agglomerative clustering model and predict the labels; 2) these samples and their predicted cluster labels were used to train a Random Forest (RF) model; and 3) we used the RF model to assign the remaining 99.9% of samples to corresponding clusters. At step one, only 0.1 % of samples (total number 10,768) were randomly selected and the remaining samples (total number 10,768,658) were predicted in step three. The out of bag (OOB) score of the RF model is 0.993, indicating a good representation of the agglomerative clustering model.

Once all data points are clustered, a colour is assigned to each cluster with similar projected spatial features (Figure 5). These clusters can be plotted by cluster colour according to their corresponding spatial reference (Figure 6G, H) to generate maps which are used to evaluate the proxy regolith type each cluster corresponds to. These maps are output as TIFF and PNG files (see Appendix B and Section 3.1).

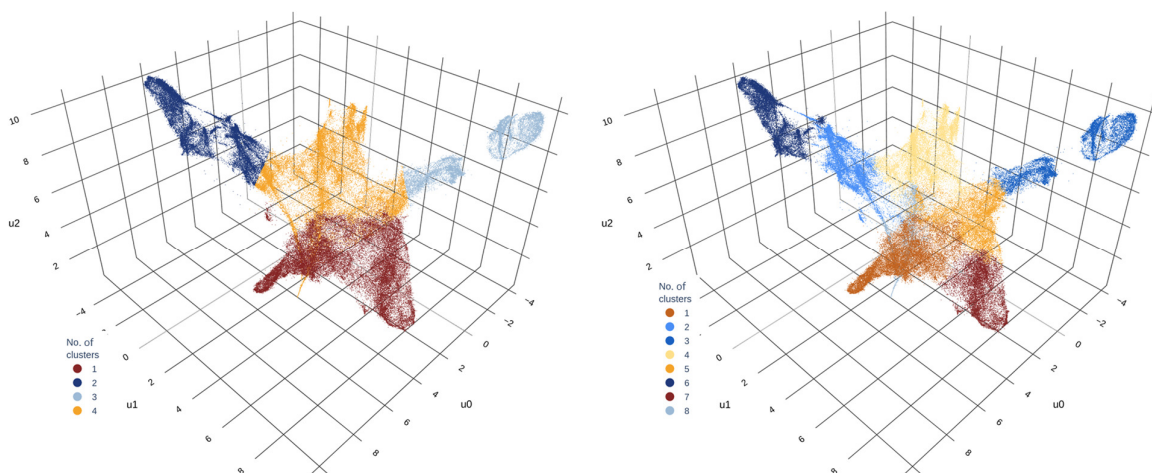


Figure 5: Clustered data points in the 3-dimensional latent space grouped by cluster colour to visualise data separation. (A) Data points clustered into four clusters using the k-means algorithm (k-means4). (B) Data points clustered into eight clusters by an agglomerative clustering algorithm (agg8). Refer to Appendix B for interactive plot.

2.4.2 Geochemical outliers by landscape type

Prior to identifying outliers within the geochemical dataset, all values below the detection limit were replaced with half the detection limit value for each respective element. Outliers were calculated on log transformed data and are defined as values that are greater than 1.5 times the interquartile range outside the first and third quartiles. Geochemical data were also grouped by their corresponding landscape cluster and outliers calculated for each of these clusters. The workflow automatically produces boxplots by element and landscape type, as well as outliers in spatial context plotted over a landscape cluster map for each element (see section 3.3). In addition, shapefiles with outliers grouped by element and landscape type are produced (see Appendix B). At the time of modelling the MacDonnell Ranges project area, the UltraFine+® Next Gen Analytics workflow created outputs of geochemical outliers by landscape only. Subsequent usability testing for multiple project sites concluded that the future workflow will also produce outputs for outliers from the entire submitted dataset for ease of comparison.

2.4.3 Principal Component Analysis

Principal Component Analysis (PCA) is performed on centred-log ratio and quantile-normalised transformed geochemical data for each soil sample. All values below the detection limit were replaced with half the detection limit value for the respective element. The workflow reduces each data point (soil sample) from n dimensions (number of analysed elements) into five principal components (PC0 to PC4) while preserving the maximum of information. The explained variance (importance) of the principal components decreases from PC0 to PC4. For each principal component, the loading (influence) of each element is plotted on a spider diagram to illustrate the general geochemical affinity (see Section 3.4). The spatial distribution of samples coloured by weight of the principal component is another automatic output (see Section 3.4). 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 is missing, the element is included in the PCA analysis, but no principal components are calculated for the affected sample.

3 The UltraFine+® Next Gen Analytics outputs

The UltraFine+® Next Gen Analytics workflow uses machine learning to integrate spatial data and soil properties in several derived outputs. These outputs include proxy regolith landscape clusters for a given project area, maps and boxplots of elemental outliers by landscape type, exploration indices, soil texture diagrams, dispersion direction, regolith indices and catchment analyses. Dispersion direction, regolith indices and catchment analyses are still under development and limited by data availability including FTIR. As a result, these parameters will not be reported for the NTGS data. Most outputs are available in TIFF, PNG, shapefile and CSV formats (Appendix B)

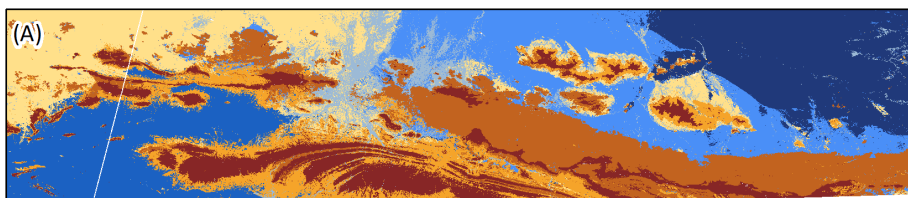
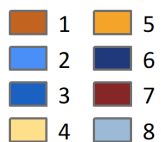
3.1 Landscape clusters

As part of the UltraFine+® Next Gen Analytics workflow, spatial data clustering is used to derive proxy regolith types to provide landscape context for geochemical samples. A variety of clustering methods with varying numbers of clusters (proxy landscape types) have been trialled in several locations across the Australian continent, with the aim to apply landform mapping over a wide range of settings and project sizes. The current output of the UltraFine+® Next Gen Analytic includes a four-cluster k-means clustering (k-means4) for small areas (< 20 km²) and an eight-cluster agglomerative hierarchical clustering (agg8) for larger areas (> 20 km²). All outputs of the UltraFine+® Next Gen Analytics workflow are automatically generated for both algorithms and can be found in Appendix B for the MacDonnell Ranges project. Optimising and ground-truthing the UltraFine+® Next Gen Analytics workflow for application on different scales and for a variety of landscapes across the Australian continent as well as a trial site in New Zealand (Figure 1) is part of ongoing research.

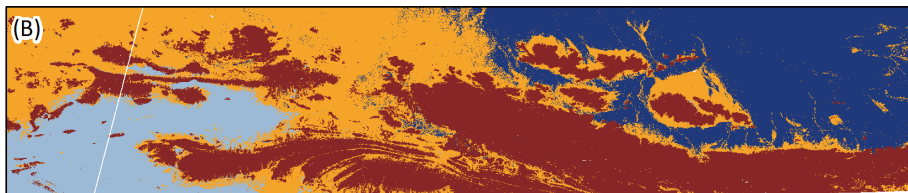
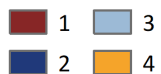
3.1.1 Comparison of unsupervised clustering methods

For the MacDonnell Ranges project area, both clustering algorithms, k-means (with four clusters) and agglomerative (with eight clusters), captured the main features (maximum variation) of the input layers (Figure 6A, B). However, due to the large size (20,953 km²) of the area and the resulting complexity of the landscape, the larger number of dense clusters (agg8) resulted in a more differentiated output. The appropriate number of clusters depends largely on area size, the resolution of the input layers, and the complexity of the respective landscape. In the case of the MacDonnell Ranges project, a comparison of our landscape clusters with input layers (Figure 6C-F) as well as surface geology (Figure 6G) show that an even larger number of clusters would have resulted in a more detailed approximation of the complex regolith types. However, to generate landscape context for geochemical samples, the number of samples across a given area was considered, and eight clusters were deemed appropriate to avoid small clusters with few or no samples. The resulting landscape clusters are therefore a balanced approach to represent the major landscape types of the area while also enabling meaningful interpretation of geochemical data. Despite these constraints, the resulting output provides a more detailed landscape context than publicly available interpreted regolith products (Figure 6H).

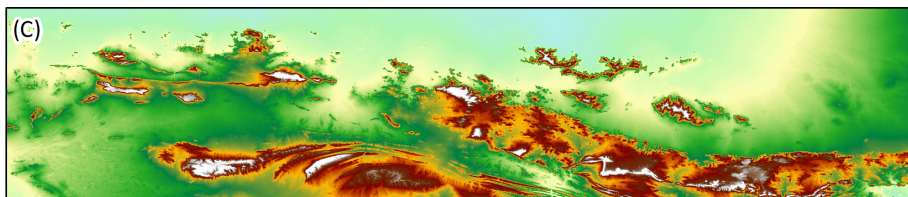
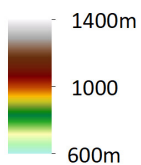
(A) agg8



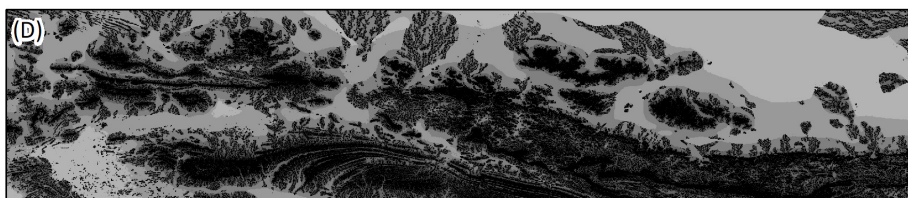
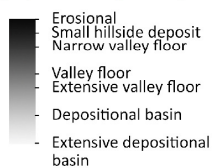
(B) kmeans4



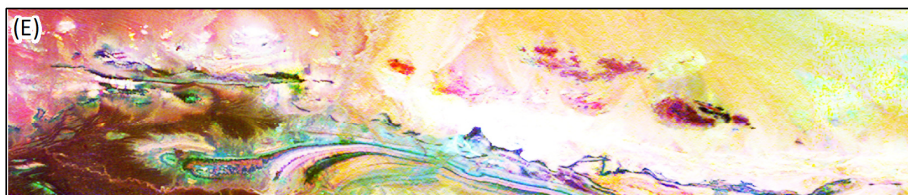
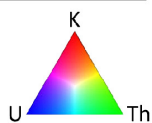
(C) DEM 1-sec



(D) MrVBF (1-sec)

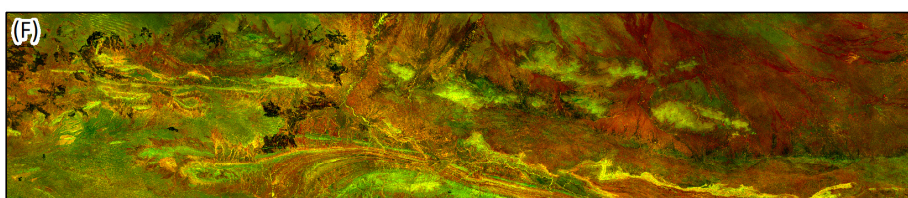
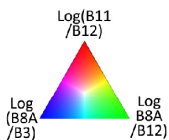


(E) Radiometrics



(F) Regolith Ratio

Sentinel-2



(G) Surface Geology

Refer to figure caption

(H) Regolith

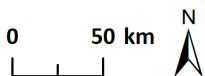
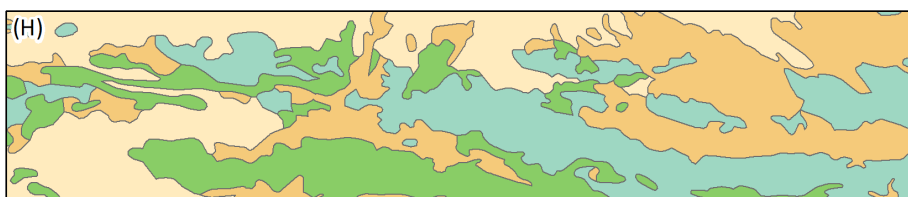
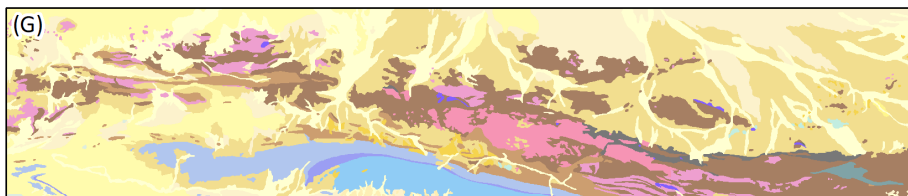


Figure 6 (previous page): Proxy regolith clustering outputs (A, B), input layers (C-F), and comparison to traditional geological maps (G, H). (A) Proxy regolith types derived via clustering algorithm agg8, plotted in spatial context. (B) Proxy regolith types derived via clustering algorithm k-means4, plotted in spatial context. White line in E and F is a stitch line between two sentinel 2 satellite image tiles. For subsequent models the workflow will utilise bare earth Sentinel data avoiding such artifacts. (C) 1-second DEM SRTM. (D) Continuous MrVBF. (E) RGB image of radiometric grid of Australia. (F) RGB image of sentinel-2 regolith ratios. (G) Surface Geology (H) Regolith Geology.

3.1.2 Landscape types

While the unsupervised landscape clustering approach does not assign a particular class to a cluster (see three-dimensional plot in Figure 4), the spatial representation of these clusters corresponds well with broad landscape patterns (Figure 6). Therefore, proxy regolith types (Figure 7) can be interpreted from the landscape clusters by comparing each cluster to the input layers. In addition, the clusters were also compared to publicly available interpreted surface geology (Figure 6G) and regolith geology (Figure 6H). It is important to note that these are generalised descriptors aligning with the majority of the landscape features within a cluster and no ground-truthing was undertaken for the MacDonnell Ranges project. However, we highlight below several examples of the overall positive alignment of landscape clusters with broad patterns of many mapped units as well as input layers.

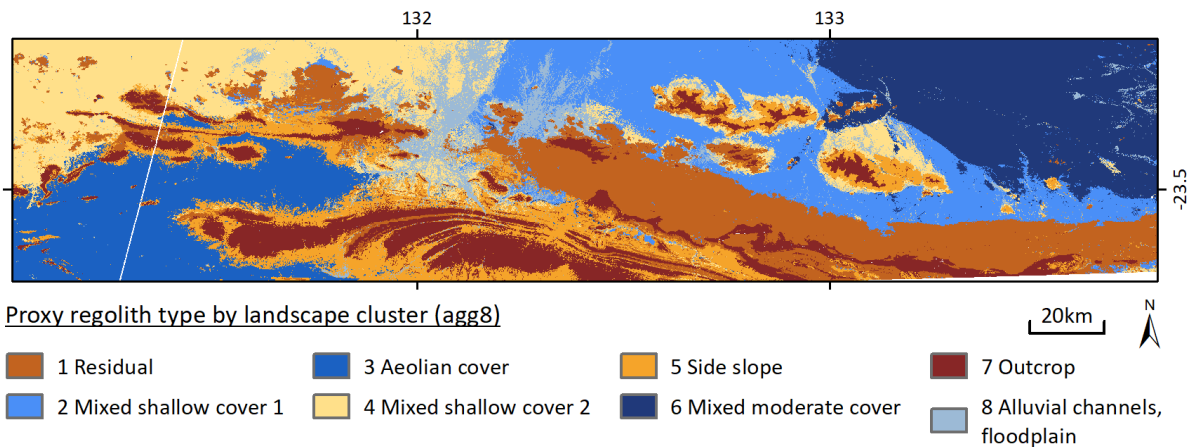


Figure 7: Correlation of proxy regolith types with landscape clusters generated by an agglomerative clustering algorithm with 8 clusters. Please refer to the text for more details on proxy regolith types.

Outcrop, residual and erosional landscape settings

Exposed outcrops and residual landforms evident in DEM and MrVBF input layers align well with landscape clusters 1 (light brown) and 7 (dark brown) on a broad scale (Figure 6). The differentiation between these two clusters, while not readily discernible in MrVBF and DEM, appears to be linked to the parent material highlighted in radiometric data (Figure 8G; white corresponds with cluster 1) and the exposed nature of landscape cluster 7 is readily evident in satellite imagery (Figure 8B). Cluster 1 corresponds well with granitic units and weathered bedrock indicated in the surface geology map (Figure 8C) and cluster 1 and 7 also correlate well with saprolite and saprock, respectively, as indicated in the regolith map (Figure 8D). The spatial distribution of landscape cluster 5 (orange) is closely associated with exposed outcrops (landscape cluster 7; dark brown; Figure 9H) and is likely indicative of erosional landscape settings as indicated in satellite imagery (Figure 9B).

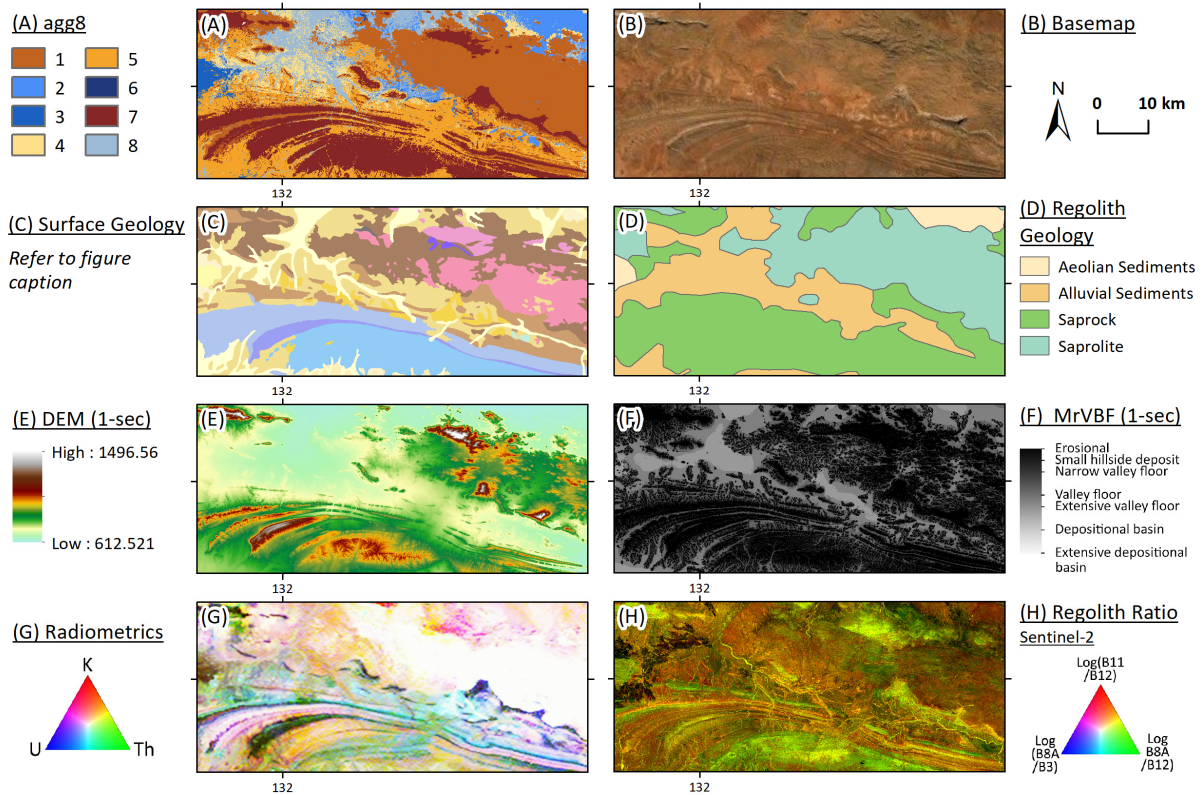


Figure 8: Assessment of landscape clusters indicative of exposed outcrops and residual material produced by agglomerative clustering (agg8) in (A) compared to satellite imagery in (B) and traditional map outputs (surface geology in (C) and regolith geology in (D)). Landscape clusters were produced from the input layers DEM (E), MrVBF (F), radiometrics (G) and regolith ratio (H).

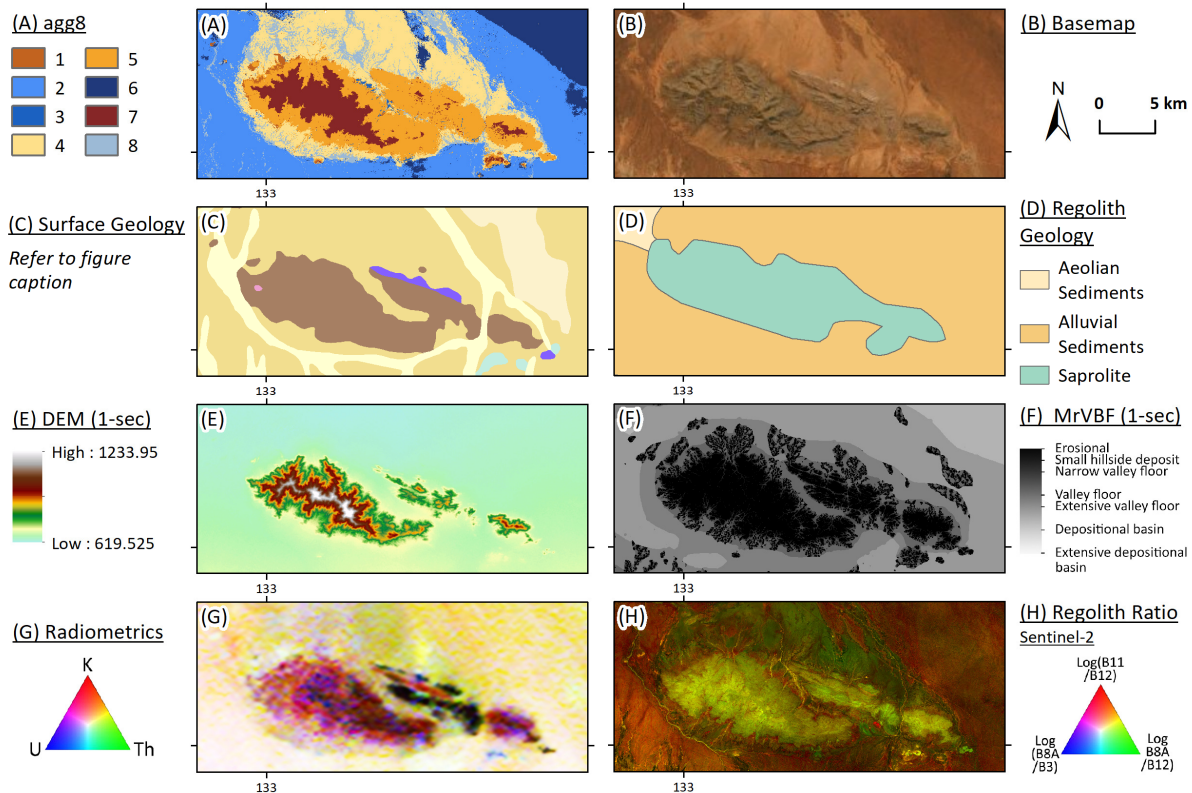


Figure 9: Assessment of landscape clusters indicative of sideslope material produced by agglomerative clustering (agg8) in (A) compared to satellite imagery in (B) and traditional map outputs (surface geology in (C) and regolith geology in (D)). Landscape clusters were produced from the input layers DEM (E), MrVBF (F), radiometrics (G) and regolith ratio (H).

Cover materials in depositional landscape settings

Landscape cluster 8 (light blue) spatially coincides with active alluvial channels and floodplains readily evident in the regolith ratio RGB image (Figure 10H) and is also visible in true-colour satellite imagery (Figure 10B). However, in areas with deeper cover, for example in the north-eastern edge of the modelled area (landscape cluster 6, dark blue; Figure 7) where similar units are mapped in the surface geology, this unit is not captured as distinctly different. This might be due to the lack of contrast in the regolith ratio input layer as well as a lack of contrast in elevation (DEM input layer) and cover thickness (MrVBF input layer).

The remaining landscape clusters 2 (royal blue), 3 (medium blue), 4 (light yellow) and 6 (dark blue) spatially correlate well with different types of cover materials. Landscape cluster 4 (light yellow) is readily distinguishable from exposed outcrop (landscape cluster 7; dark brown) and sideslope materials (landscape cluster 5; orange) in satellite imagery, DEM and MrVBF (Figure 9B, E, F), as well as from alluvial channels (landscape cluster 8; light blue) as evident in RGB imagery of regolith ratios (Figure 10H). Similarly, landscape cluster 3 (medium blue) displays a unique “signature” in radiometric data RGB imagery (Figure 6E) that is different to any other landscape cluster. This landscape cluster correlates very well with mapped aeolian cover (regolith map; Figure 6F) and features such as sand dunes are closely defined and differentiated from surrounding landscape types (Figure 11).

However, in the northern half of the modelled area the landscape model appears to class points associated with depositional landscape settings from east to west into clusters 4 (light yellow), 2 (royal blue) and 6 (dark blue) that do not appear readily distinguishable when viewing input layers and satellite imagery alone. These landscape clusters have been termed Mixed Shallow Cover 1 (landscape cluster 2; royal blue), Mixed Shallow Cover 2 (landscape cluster 4, light yellow) and Mixed Moderate Cover (landscape cluster 6; dark blue) as all three landscape clusters appear to be associated with aeolian as well as alluvial transported cover and sheetwash. While there appear to be only subtle differences between clusters 2 (royal blue), 4 (light yellow) and 6 (dark blue), the MrVBF indicates cover depth differences, with clusters 2 and 4 being shallower cover than cluster 6 (Figure 6B). However, comparing these UltraFine+® landscape clusters to a weathering intensity map (Wilford and Roberts 2019; Figure 12B) which combines only two of the inputs layers (radiometrics and DEM) indicates a correlation between different intensities of weathering. This could be linked to the age of transported cover, with landscape cluster 2 (royal blue) being distinctly younger than landscape cluster 4 (light yellow) and 6 (dark blue).

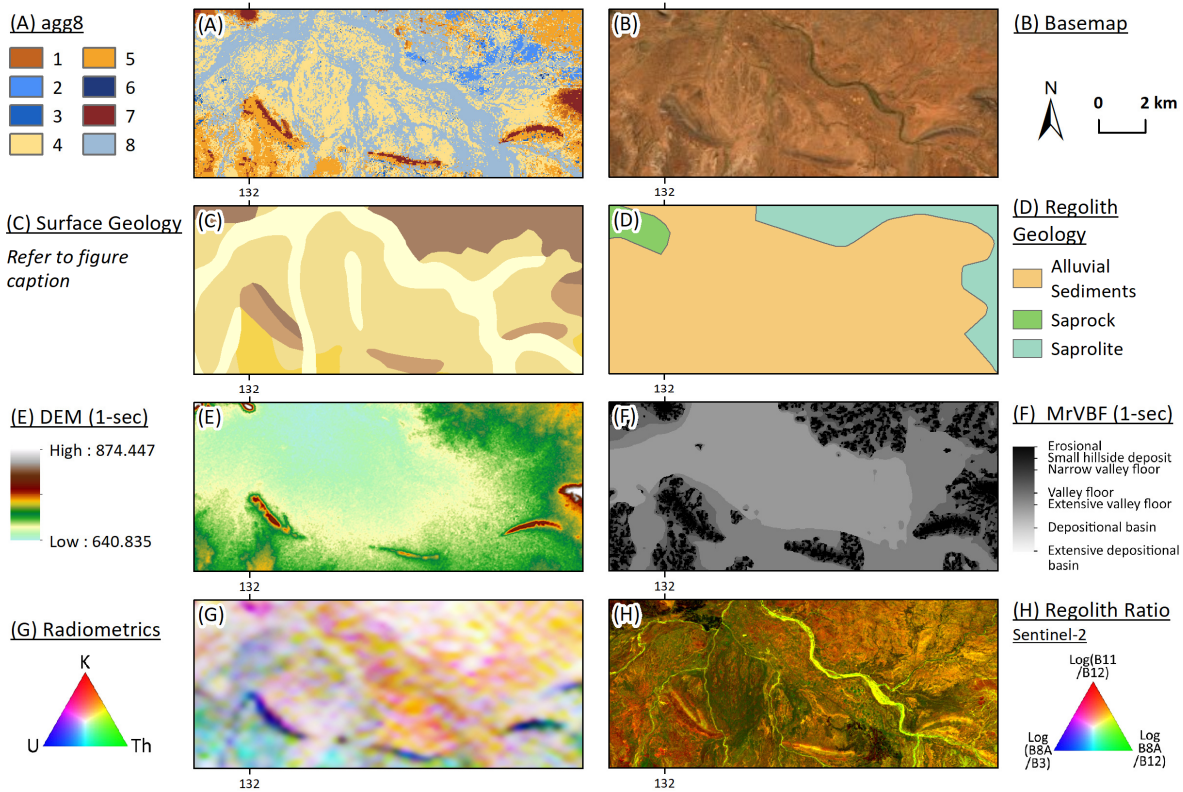


Figure 10: Assessment of landscape clusters indicative of active alluvial channels produced by agglomerative clustering (agg8) in (A) compared to satellite imagery in (B) and traditional map outputs (surface geology in (C) and regolith geology in (D)). Landscape clusters were produced from the input layers DEM (E), MrVBF (F), radiometrics (G) and regolith ratio (H).

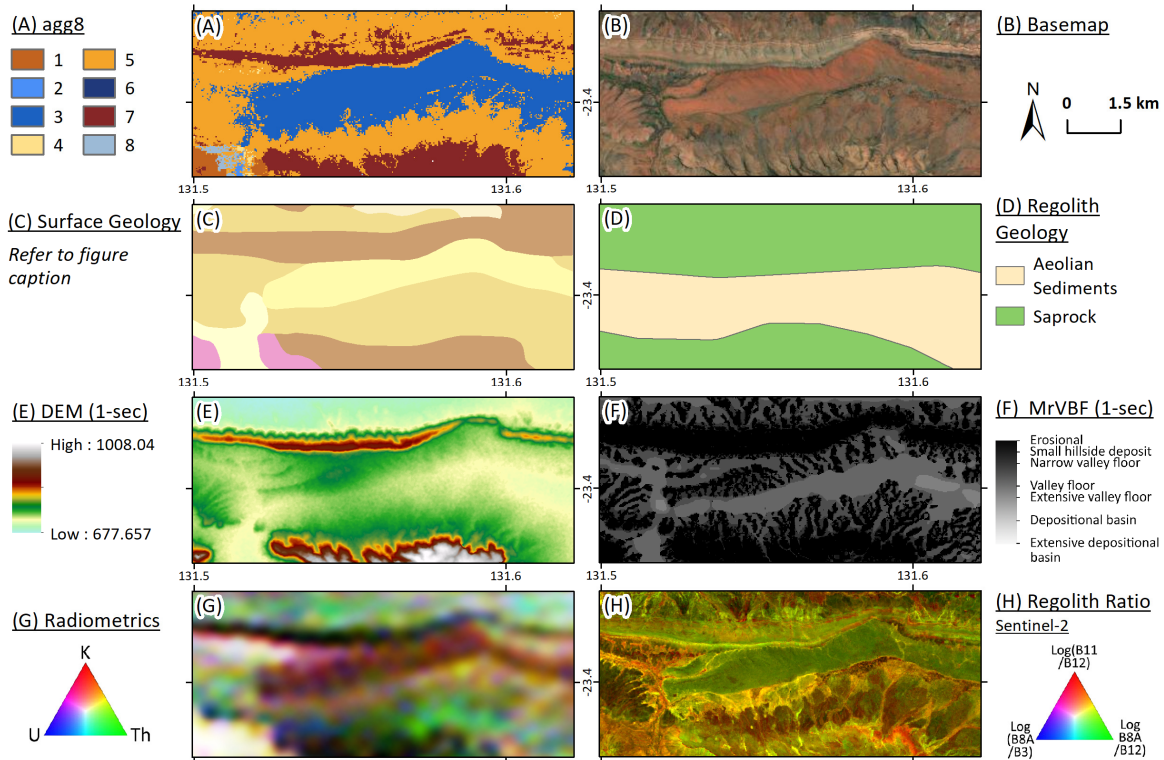


Figure 11: Assessment of landscape clusters as proxy for aeolian cover produced by agglomerative clustering (agg8) in (A) on the example of a sand dune, compared to satellite imagery in (B) and traditional map outputs (surface geology in (C) and regolith geology in (D)). Landscape clusters were produced from the input layers DEM (E), MrVBF (F), radiometrics (G) and regolith ratio (H).

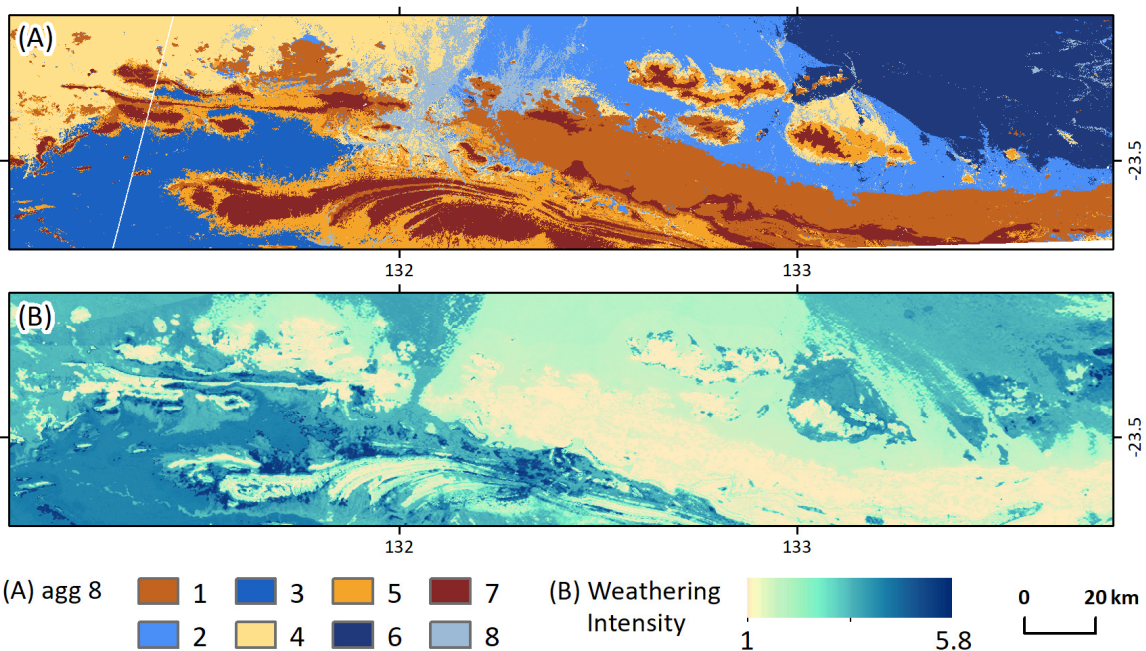


Figure 12: Comparison of landscape clusters 2, 4 and 6 with weathering intensity over the MacDonnell Ranges project area. (A) Landscape cluster types derived via the clustering algorithm agg8 plotted in spatial context. (B) Weathering intensity (Wilford and Roberts 2019) derived from radiometrics and DEM.

3.2 Comparison of UltraFine+® to traditional soil survey analyses

The UltraFine+® soil analysis method is based on separating and analysing only the ultrafine fraction (< 2 µm) of a given sample as most metals in transported cover tend to attenuate preferentially to clay particles and other “scavenging” phases with large surface areas. By removing the bulk of the coarse-grained, “barren” portion of the sample, the signal to background ratio is increased and nugget effects are removed. In addition, the analysis requires smaller sample volumes due to the enhanced sensitivity with up to 100 – 250 % increased concentrations of Au, Cu, Zn observed (Noble et al. 2018).

Here we compare the results of the UltraFine+® analysis with those carried out by the NTGS as part of a stream sediment survey in 2001. The NTGS previously analysed 35 elements using four-acid (HF-HClO₄-HNO₃-HCl) digest via ICPMS and OES, and Au using a 24 hour cyanide leach via ICPMS at Genalysis Laboratory Services in Perth (see Dunster and Mügge 2001 for detailed methods). The UltraFine+® analysis includes 18 additional elements not previously analysed by the NTGS survey (Ag, Al, Br, Cs, Ga, Ge, Hf, Hg, I, In, La, Pt, Re, S, Sc, Ta, Tl, Zr) providing additional pathfinder elements for mineral exploration. However, previously reported analyses for Na and P are not available via the UltraFine+® method due to the addition of a dispersant containing both elements (Table 3).

Table 3: Comparison between the 2001 analyses and the UltraFine+® method in 2020 with available analyses, detection limits and Pearson’s correlation coefficient *r*. All analyses are reported in ppm, except for Au and Pt which are reported in ppb. Since the data acquisition in 2020, the element Pd has been added to the UltraFine+® element suite.

Element	NTGS Method	NTGS 2001 DL	UFF 2020 DL	<i>r</i>
Ag			0.01	
Al			10	
As	A/MS	1	0.5	0.18
Au	CN.1/MS	0.01	0.5	0.054
Ba	A/OES	2	0.2	0.13
Be	A/MS	0.1	0.2	0.21
Bi	A/MS	0.01	0.1	0.66
Br			1	
Ca	A/OES	10	10	0.68
Cd	A/MS	0.1	0.01	0.028
Ce	A/MS	0.01	0.05	0.23
Co	A/MS	0.1	0.2	0.66
Cr	A/OES	2	2	0.58
Cs			0.1	
Cu	A/OES	1	0.2	0.69
Fe	A/OES	100	100	
Ga			0.05	
Ge			0.05	
Hf			0.02	
Hg			0.01	
I			1	
In			0.01	
K	A/OES	20	10	0.05
La			0.05	
Li	A/MS	0.1	0.5	0.32
Mg	A/OES	20	10	0.72
Mn	A/OES	1	2	0.58
Mo	A/MS	0.1	0.1	0.32
Na	A/OES	20		
Nb	A/MS	0.05	0.01	0.29
Ni	A/OES	1	2	0.8
P	A/OES	20		
Pb	A/MS	2	0.2	0.29
Pt			1	
Rb	A/MS	0.05	0.1	0.38
Re			0.0001	
S			50	
Sb	A/MS	0.05	0.1	0.49
Sc			1	
Se	A/MS	2	0.05	
Sn	A/MS	0.1	0.2	0.14
Sr	A/MS	0.05	0.1	0.072
Ta			0.005	
Te	A/MS	0.1	0.001	0.074
Th	A/MS	0.01	0.02	0.33
Ti	A/OES	5	10	0.3
Tl			0.1	
U	A/MS	0.01	0.02	0.69
V	A/OES	2	2	0.31
W	A/MS	0.1	0.01	0.076
Y	A/MS	0.05	0.05	0.3
Zn	A/OES	1	0.2	0.22
Zr			1	

While detection limits are generally lower in the 2001 data for most elements (Table 3), the improved recovery of elements from the ultrafine fraction results in higher concentrations of elements, effectively reducing the number of results below the detection limit. Hence, for most elements, the UltraFine+® data shows a higher median than the previous survey results, despite the generally lower detection limits in the four-acid digest and 24 hour cyanide leach analyses. The 2001 survey results show overall higher recovery only for Ba, Ca, K, Mg, Nb, Rb, Sr, Ti, and W, although the median of the concentrations remains higher in the UltraFine+® data for Mg and Rb.

Overall, the results indicate that the UltraFine+® method improves the recovery of trace metals as well as the resolution of concentrations near the detection limit, which enables the delineation of subtle geochemical enrichments for elements As, Au, Be, Bi, Ce, Cd, Co, Cr, Cu, Fe, Li, Mn, Mo, Ni, Pb, Sb, Se, Sn, Th, U, V, Y and Zn. The detection of these subtle variations is particularly relevant for exploration through transported cover.

For metals of interest within the MacDonnell Ranges project, such as Au, Cu, Pb and Zn, the UltraFine+® method resulted in higher recovery with the median of the UltraFine+® data above the upper 25 % of the original survey results (Figure 13). This results in overall recoveries between 2 times (Cu) to more than 5 times (Zn) higher in the UltraFine+® dataset (Figure 14). The UltraFine+® data also shows a more representative distribution near the detection limit as well as more anomalous values for Au, Pb and Zn allowing for a more differentiated interpretation of the results (Figure 13).

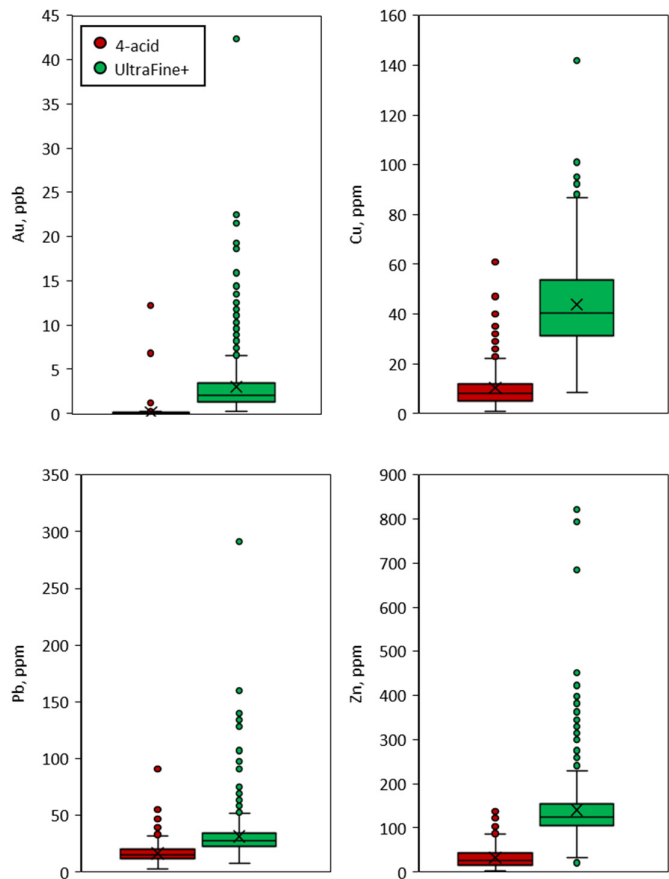


Figure 13: Boxplots of analyses for Au, Cu, Zn and Pb. The 2001 survey results are displayed in red and the 2020 results from re-analysis via UltraFine+® are displayed in green.

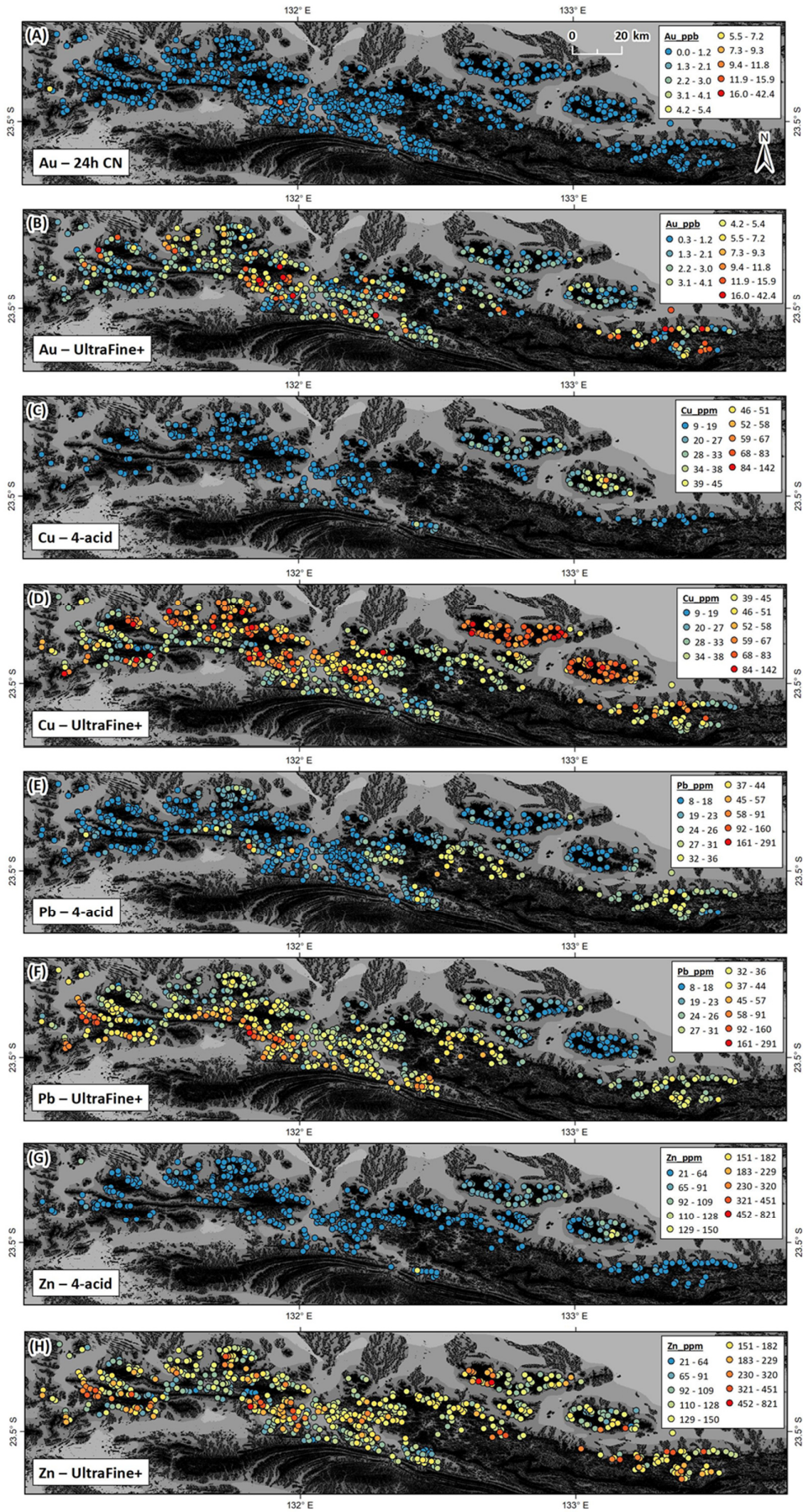
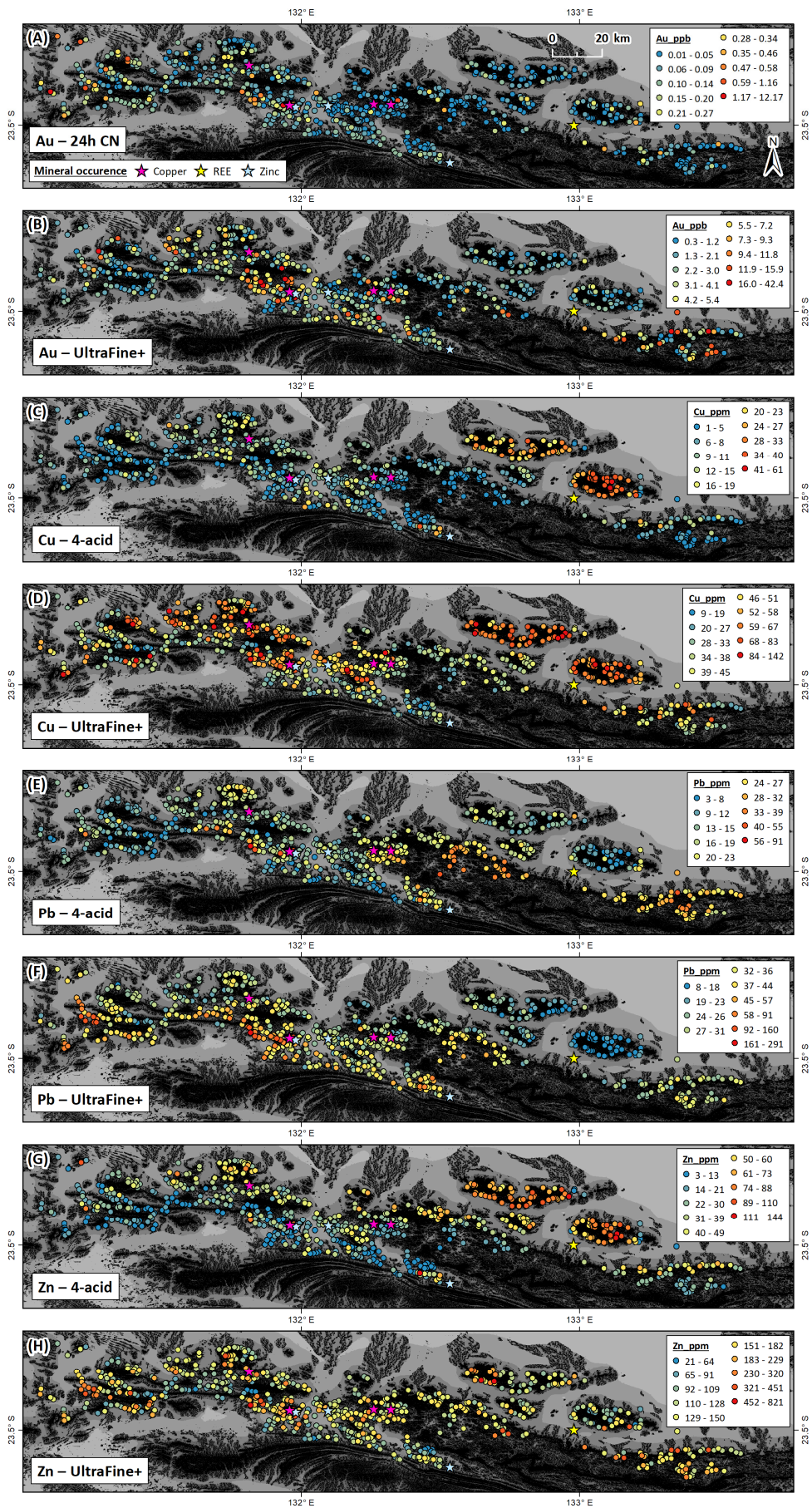


Figure 14 (previous page): Comparison of the NTGS's 2001 survey and the re-analysed results in 2020 displayed on the same scale. (A) Spatial distribution of Au analysed via 24-hour cyanide leach. (B) Spatial distribution of Au analysed via UltraFine+®. (C) Spatial distribution of Cu analysed via four-acid digest. (D) Spatial distribution of Cu analysed via UltraFine+®. (E) Spatial distribution of Pb analysed via four-acid digest. (F) Spatial distribution of Pb analysed via UltraFine+®. (G) Spatial distribution of Zn analysed via four-acid digest. (H) Spatial distribution of Zn analysed via UltraFine+®.

Pearson's correlation coefficient (r) values for Au, Pb and Zn show little to no correlation between the analysis methods, with very low values (0.05, 0.29 and 0.22, respectively; Table 3) possibly reflecting the different approaches to sampling and analysis. UltraFine+®, designed for recovering metals in transported cover, likely captures more of the mobile phases within the stream sediment environment, while more immobile elements, such as Ti, are likely better recovered by four-acid digestion. It is therefore not surprising that the spatial distributions of outliers differ between the analysis methods (Figure 15). An exception for mobile elements of interest in the MacDonnell Ranges project is Cu, which correlates reasonably well between both datasets ($r = 0.69$) indicating that Cu may largely be recovered from similar hosts, such as secondary mineral coatings on larger grains, although recovery with the UltraFine+® method is still up to 2.3 times higher in comparison indicating a significant amount of Cu present within the ultrafine fraction (Table 3, Figure 15).

Figure 15 (next page): Comparison of the NTGS's 2001 survey and the re-analysed results in 2020 displayed on scales with natural breaks within each dataset. (A) Spatial distribution of Au analysed via 24-hour cyanide leach. (B) Spatial distribution of Au analysed via UltraFine+®. Note that Au values below the detection limits have been replaced with half the detection limits (0.005 in the 2001 survey data and 0.25 in the 2020 UltraFine+® data). (C) Spatial distribution of Cu analysed via four-acid digest. (D) Spatial distribution of Cu analysed via UltraFine+®. (E) Spatial distribution of Pb analysed via four-acid digest. (F) Spatial distribution of Pb analysed via UltraFine+®. (G) Spatial distribution of Zn analysed via four-acid digest. (H) Spatial distribution of Zn analysed via UltraFine+®.



3.3 Outliers by landscape type

Outliers in soil geochemical datasets are typically calculated based on all collected samples, regardless of their landscape context. However, this 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 beneath exposed outcrops or shallow residual materials, while the same mineralisation would have a much weaker elemental signal in samples collected over moderately thick depositional landscape types. With the ability to approximate landscape types from spatial data via machine learning, the UltraFine+® Next Gen Analytics workflow can identify outliers within each individual landscape cluster. This provides a basic, first-pass interpretation of geochemical samples by proxy regolith type and the identification of otherwise “overlooked” potential anomalies. The aim is to minimise overlooking mineral deposits in transported cover, which is common when targeting only the highest concentrations with little regard for changes in soil properties and landscape type.

The UltraFine+® Next Gen Analytics workflow generates plots of centred-log ratio transformed results of all analysed elements by percentiles and separates these by regolith type (Figure 16A). Traditional outliers calculated from all data as a single sample population are presented as comparison (white boxplot). However, by grouping the sample population by landscape clusters and calculating outliers for each cluster (coloured boxplots), potential anomalies are highlighted within different landscape settings (triangles below the dashed line). For most of these landscape types, these outliers would have been considered unremarkable if evaluated as part of the whole dataset, demonstrating the benefit of evaluating the geochemistry in a landscape/regolith context.

While the landscapes clusters generated for the MacDonnell Ranges project area are appropriate proxies for the regolith types within the area (see section 3.1), the UltraFine+® Next Gen Analytics workflow was designed for soil samples. The approach to identify outliers by landscape type on stream sediment samples is therefore not as representative, since this material will likely collect element signatures from a broader area than many soil samples. This potentially causes elemental signatures in this dataset to cross landscape proxy boundaries and results should be interpreted with caution. Nevertheless, the MacDonnell Ranges project site was chosen as a first-generation, large-scale trial site with a focus on principal functionality of the UltraFine+® Next Gen Analytics workflow rather than targeting tangible exploration outcomes and we present some example outputs for outliers by landscape types here.

Due to the density of samples compared to the total area modelled, sample populations in clusters 3, 6 and 7 are too small for meaningful interpretation with no to very few outliers present in these landscape types (e.g., for Ni in Figure 16). In general, we consider the minimum number of samples for statistically relevant analyses to be 50 per landscape type and recommend to disregard outliers generated by landscape clusters with less than 15 samples per landscape type as these are not meaningful. As a precaution, outliers in clusters with less than 10 samples do not display on maps or in shapefiles generated in the UltraFine+® Next Gen Analytics workflow. Most samples in the MacDonnell Ranges project area are from residual or erosional landscape settings (441 samples in landscape clusters 1, 5 and 7) and would, therefore, be expected to be well represented by outliers in the overall sample population. However, for the residual landscape cluster 1, when assessed individually rather than with other erosional/residual landscape clusters (i.e., 7, exposed outcrop, and 5, side slope material) the workflow identifies outliers at generally

lower concentrations for many elements (e.g., Ni; Figure 16). This residual landscape cluster (cluster 1) coincides with granitic lithologies (Figure 6A, G) which are generally low in base metals, and geochemical outliers in this landscape cluster might therefore be worth considering as indicative of a different potential mineral system.

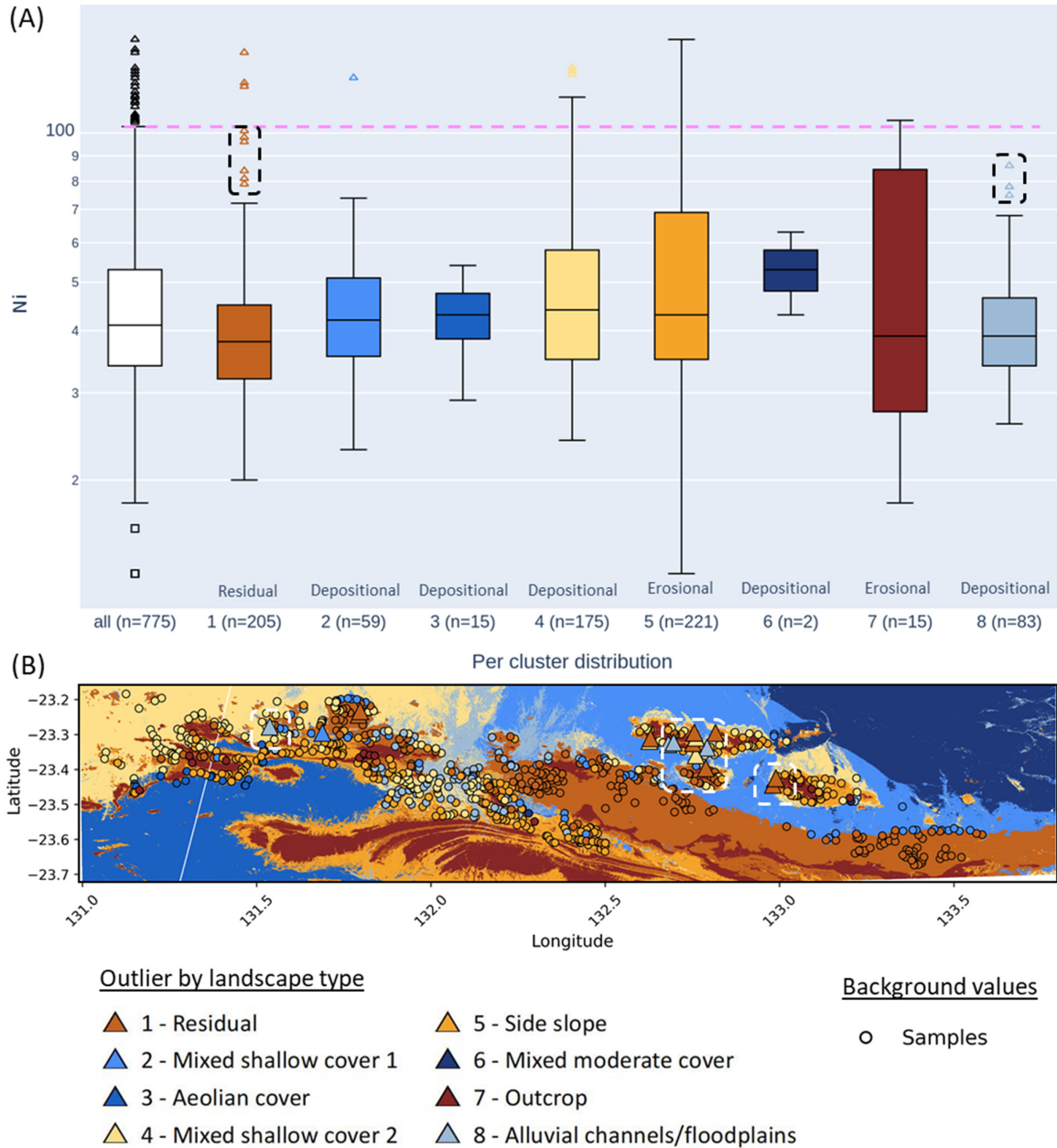


Figure 16: Ni outliers by landscape cluster over the MacDonnell Ranges project area. (A) Boxplots for all data (white box) and by landscape type (coloured boxes). Dashed mauve line indicates the upper 25 % boundary for the whole sample population. Commonly observed soil anomalies are samples above the dashed line (white triangles). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)). (B) Spatial distribution of Ni outliers (triangles) by proxy regolith type. White dashed boxes indicate outliers by landscape type below the dashed mauve line in (A).

3.3.1 Examples of outliers by landscape type - Gold and related pathfinders

Mineral occurrences containing Au in association with Cu and Fe-oxides within the MacDonnell Ranges project area have been observed at Haasts Bluff (Northern Territory Mineral Occurrences; Figure 2). Gold ranges from below detection limit to 42.4 ppb with a mean of 3.0 ppb (Figure 14B). The greatest concentrations were measured around Haasts Bluff and Ulpuruta, and northwest of Haasts Bluff/south of Mount Larrie, as well as in the southeast of the project area (Figure 14B). With very similar median values in all sample populations by landscape type compared to the overall sample population, there are only a limited number of additional outliers by landscape type. However, both overall outliers and outliers by landscape type identify anomalous values in the above areas (Figure 17).

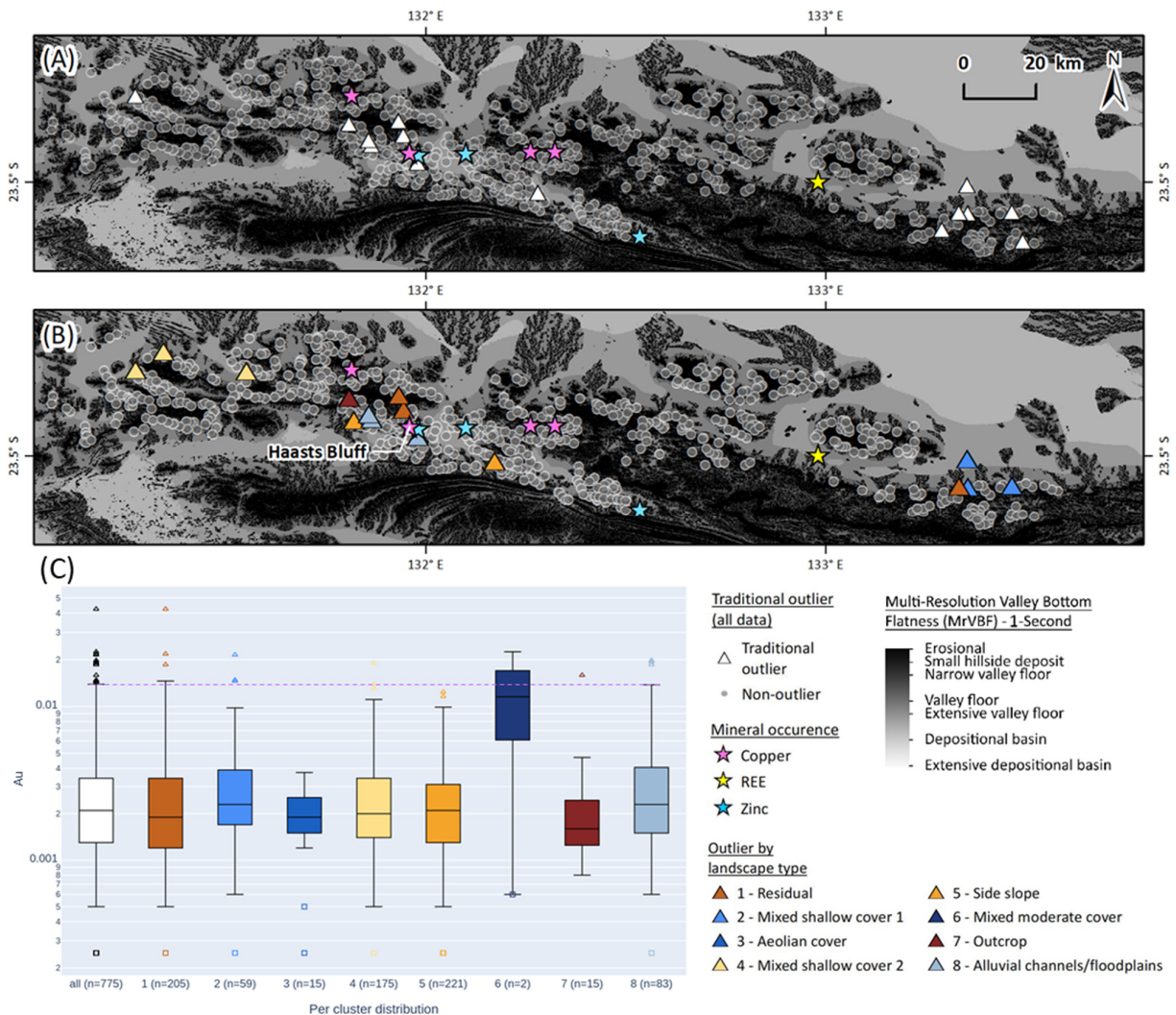


Figure 17: Comparison of Au outliers in the whole sample population to Au outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Au outliers for all data. (B) Spatial distribution of Au outliers by landscape cluster. Background value samples plotted in grey. The known Au-related mineral occurrence is labelled. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). Commonly observed soil anomalies are samples above the dashed horizontal line (white triangles in (A)). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)).

Concentrations for Bi range from 0.1 ppm to 5.4 ppm with a mean of 0.6 ppm (Figure 18C). Outliers for Bi generally highlight the same areas as the data for Au with broader signatures. Grouping sample populations by landscape type also identifies outliers in depositional landscape settings, i.e., landscape clusters 2, 4 and 8, which would have been overlooked. While outliers in landscape cluster 2 (blue, mixed shallow cover 1) would traditionally be considered unremarkable, they indicate elevated Bi concentrations in shallow cover, for example in the proximity of Glen Helen I (a Pb-Zn-Cu mineral occurrence).

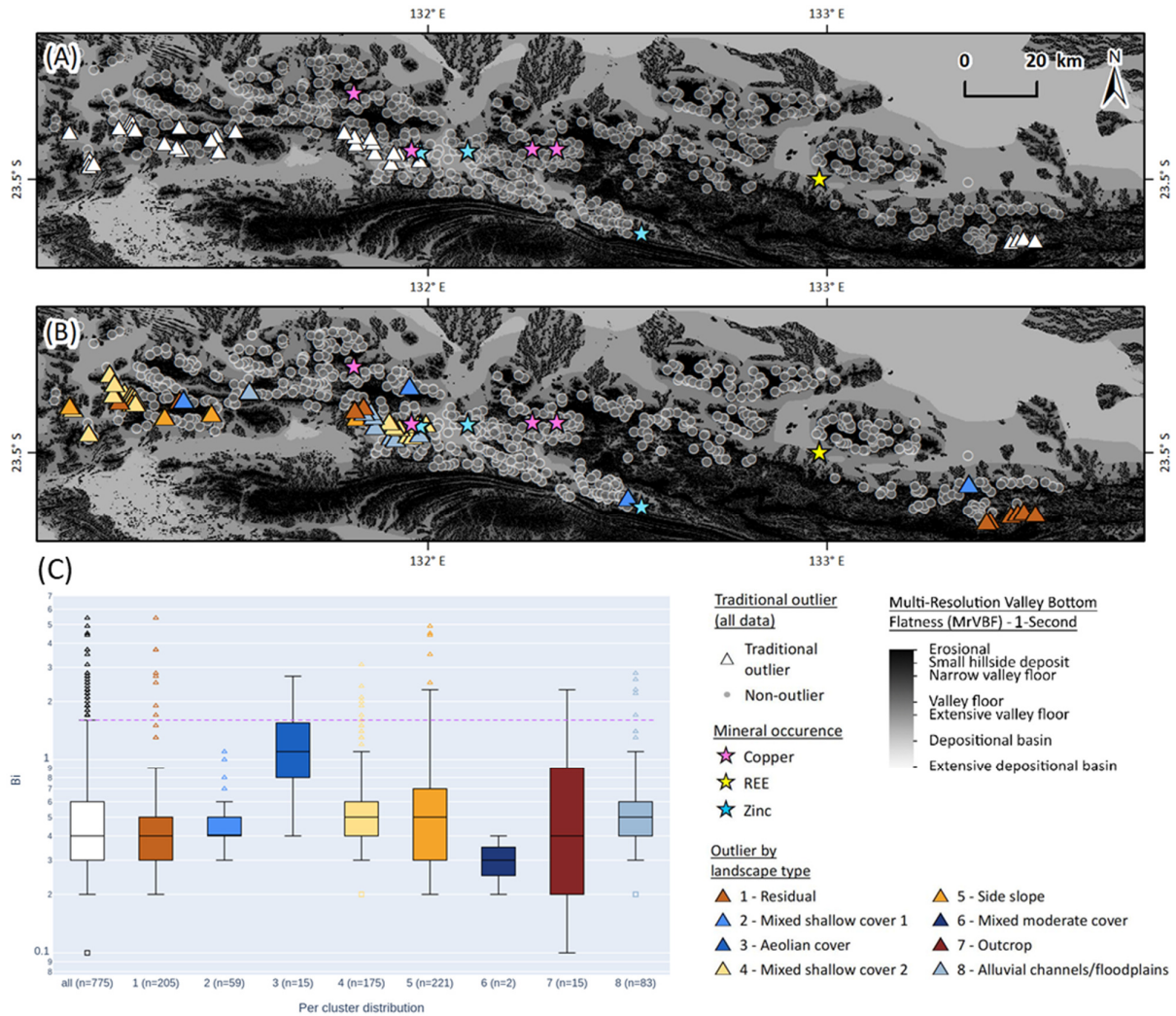


Figure 18: Comparison of Bi outliers in the whole sample population to Bi outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Bi outliers for all data. (B) Spatial distribution of Bi outliers by landscape cluster. Background value samples plotted in grey. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). Commonly observed soil anomalies are samples above the dashed horizontal line (white triangles in (A)). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)).

3.3.2 Examples of outliers by landscape type - Base metals

The base metals Cu, Zn and Pb are the main elements of interest within the MacDonnell Ranges project area, with several mineral occurrences reported at Glen Helen I (carbonate-hosted Pb-Zn-Cu), Glen Helen II and Glen Helen III (vein-hosted Cu-Pb-Zn-Ag), Stokes Yard (pegmatitic Zn-Cu-Pb-

Ag), Mount Larrie (pegmatitic Cu veins), Haasts Bluff (Fe-oxide associated Cu-Au) and Ulpuruta (vein hosted Zn-Cu-Pb; Figure 2). The spatial distribution of highest concentrations of these three elements varies (Figure 14F-H). However, all three elements show elevated signatures to the south of Mount Larrie and in proximity to the Ulpuruta and Haasts Bluff mineral occurrences and west thereof (Figure 15D, F, H).

Zinc concentrations range from 20.7 ppm to 821.0 ppm with a mean of 140.8 ppm (Figure 19A). Separating the sample populations by landscape type identifies additional outliers in sideslope settings near the mineral occurrences Ulpuruta, Haasts Bluff and Mount Larrie, providing a stronger target signature for Zn than outliers calculated from the entire sample population (Figure 19B).

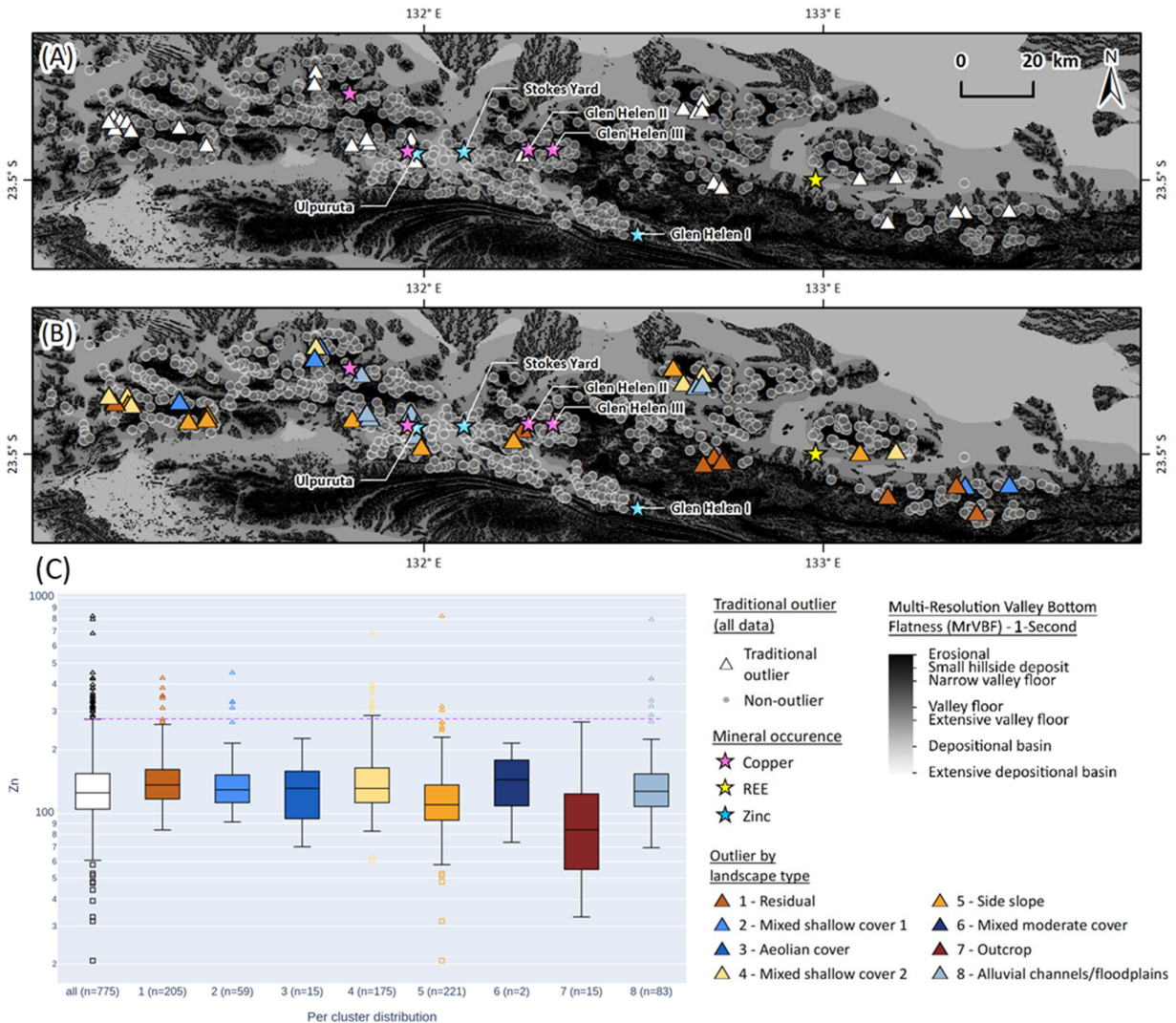


Figure 19: Comparison of Zn outliers in the whole sample population to Zn outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Zn outliers for all data. (B) Spatial distribution of Zn outliers by landscape cluster. Background value samples plotted in grey. The known Zn-related mineral occurrences are labelled. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). Commonly observed soil anomalies are samples above the dashed horizontal line (white triangles in (A)). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)).

Lead concentrations range from 7.6 ppm to 291 ppm with a mean of 31.2 ppm (Figure 20A). The greatest concentrations were measured in the west of the project area and to the west of the Ulpuruta and Haasts Bluff mineral occurrences (Figure 15F). An outlier also occurs near Glen Helen I. Outliers by landscape type identified some additional potential targets in residual and depositional landscape clusters (clusters 1 and 8; Figure 20). Some of the outliers within the whole dataset are no longer classed as outliers when calculating outliers by landscape type (e.g., compare white triangles in the west of the project area in Figure 20A with coloured triangles in Figure 20B).

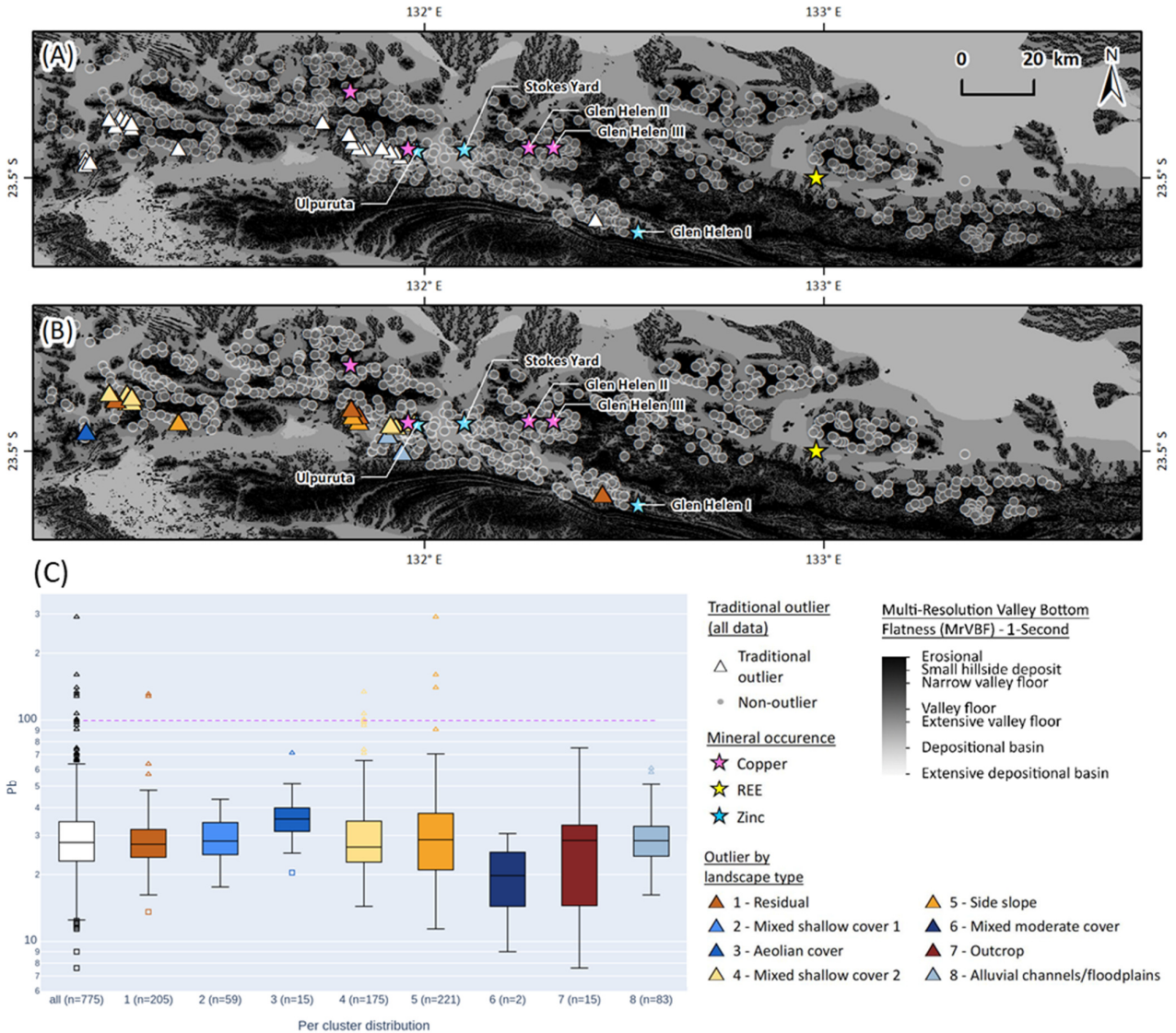


Figure 20: Comparison of Pb outliers in the whole sample population to Pb outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Pb outliers for all data. (B) Spatial distribution of Pb outliers by landscape cluster. Background value samples plotted in grey. The known Pb-related mineral occurrences are labelled. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). Commonly observed soil anomalies are samples above the dashed horizontal line (white triangles in (A)). Those shown below the dashed mauve line would not be easily observed without the landscape context (coloured triangles in (B)).

Copper concentrations range from 8.6 ppm to 142.0 ppm with a mean of 43.7 ppm (Figure 21A). The greatest concentrations were measured in exposed outcrop and side slope settings, especially in the granitic and granulitic units in the northwest of the project area as well as in proximity to Ulapuruta, Haasts Bluff and Mount Larrie (Figure 15D). However, there is only one outlier in the overall dataset, and none when sample populations are separated by landscape type, indicating that the results are representative of background concentrations and there is unlikely to be Cu mineralisation systems near these samples (Figure 21B).

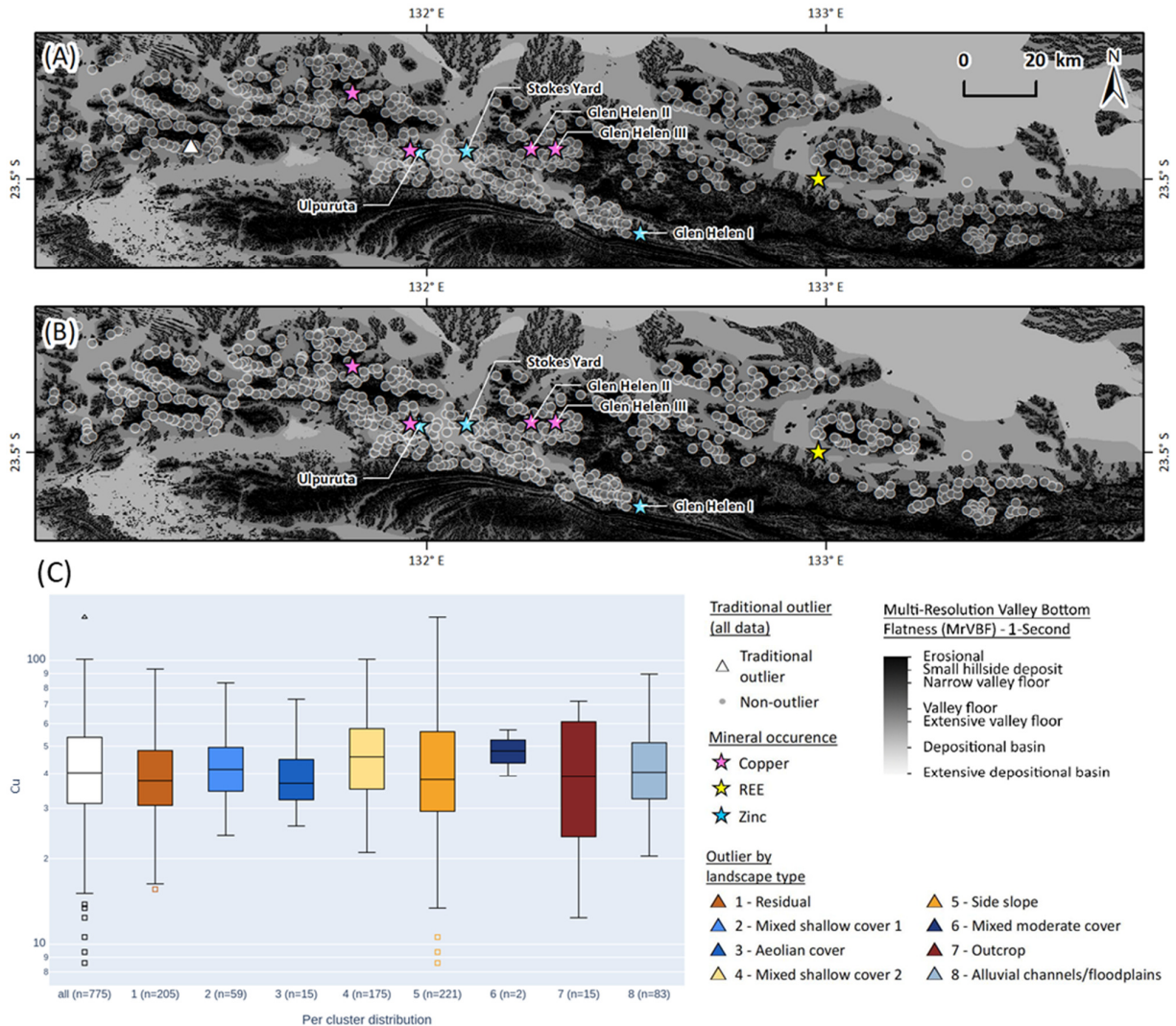


Figure 21: Comparison of Cu outliers in the whole sample population to Cu outliers by landscape cluster over the MacDonnell Ranges project area. (A) Spatial distribution of Cu outliers for all data. (B) Spatial distribution of Cu outliers by landscape cluster. Background value samples plotted in grey. The known Cu-related mineral occurrences are labelled. (C) Boxplots for all data (white box) and by landscape type (coloured boxes). No outliers by landscape type were identified.

3.4 Exploration Indices

Deriving exploration indices via principal component analysis (PCA) is part of the UltraFine+® Next Gen Analytics workflow to automate the identification of patterns in the geochemical dataset. The PCA is performed on quantile-normalised centre-log transformed data of all elements analysed for

each soil sample. The automated output generates five principal components (PC0 to PC4). These principal components capture the maximum variance among the samples and effectively reduce the number of dimensions. The influence of each element on each of these five principal components are automatically output in a spider diagram (Figure 22). The further away an element plots from the 0 line for a given principal component (coloured lines), the greater the loading for (influence on) the specific principal component. This allows for a rapid, first-pass identification of element association and potential exploration indices within the dataset. The spatial distribution of samples coloured by weight of the principal component is another automatic output (Figure 23). Principal component analyses for each sample are available as shapefiles within the MacDonnell Ranges project data package (Appendix B).

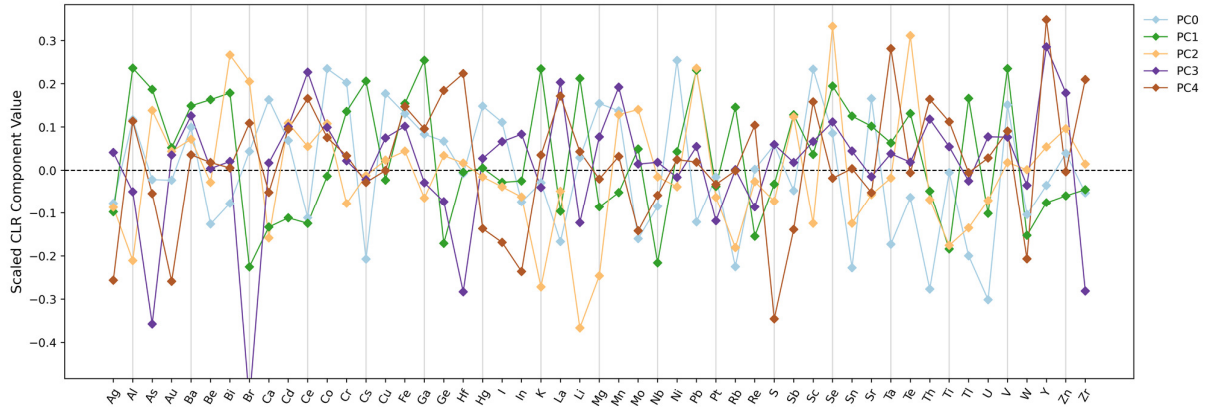


Figure 22: Elemental loadings for each of the first five principal components for the MacDonnell Ranges project. The further away an element plots from the 0 line, the greater the loading for (influence on) the specific principal component.

While the importance (explained variance) of each of the principal components decreases from PC0 to PC4, principal components are heavily influenced by their landscape setting. Earlier principal components will therefore often pick out large-scale lithological variation as the major component(s). However, principal components with exploration potential usually explain less variance in the data than the major elements and geological influence. Hence, in many areas, PC2 and higher may represent the more relevant component in the context of mineral exploration. In some cases, signatures relating to mineralisation will constitute relatively low proportions of variance and thus not be represented by principal components.

In the case of the MacDonnell Ranges project, the first principal component (PC0) explains 46.3 % of the variability within the dataset and readily differentiates felsic (blue) from mafic (red) type rocks (Figure 23A) as indicated by the opposing associations of U and Th, and Ni, Co and Sc. Principal Component 1 explains 20.0 % of the variability within the dataset and emphasises similar, lithology-related variation. Principal component 2 only explains 13.0 % of the variability within the dataset, but shows some exploration potential with association of Pb, As, Bi and Se. The spatial distribution of PC2 coincides broadly with known Au and Zn-Cu-Pb mineral occurrences (Haasts Bluff, Ulparuta, Mount Larrie and Stokes Yard, but not with Glen Helen II and Glen Helen III; Figure 24). Principal component 3 and 4 only explain 11.2 % and 9.5 % of the variability within the dataset.

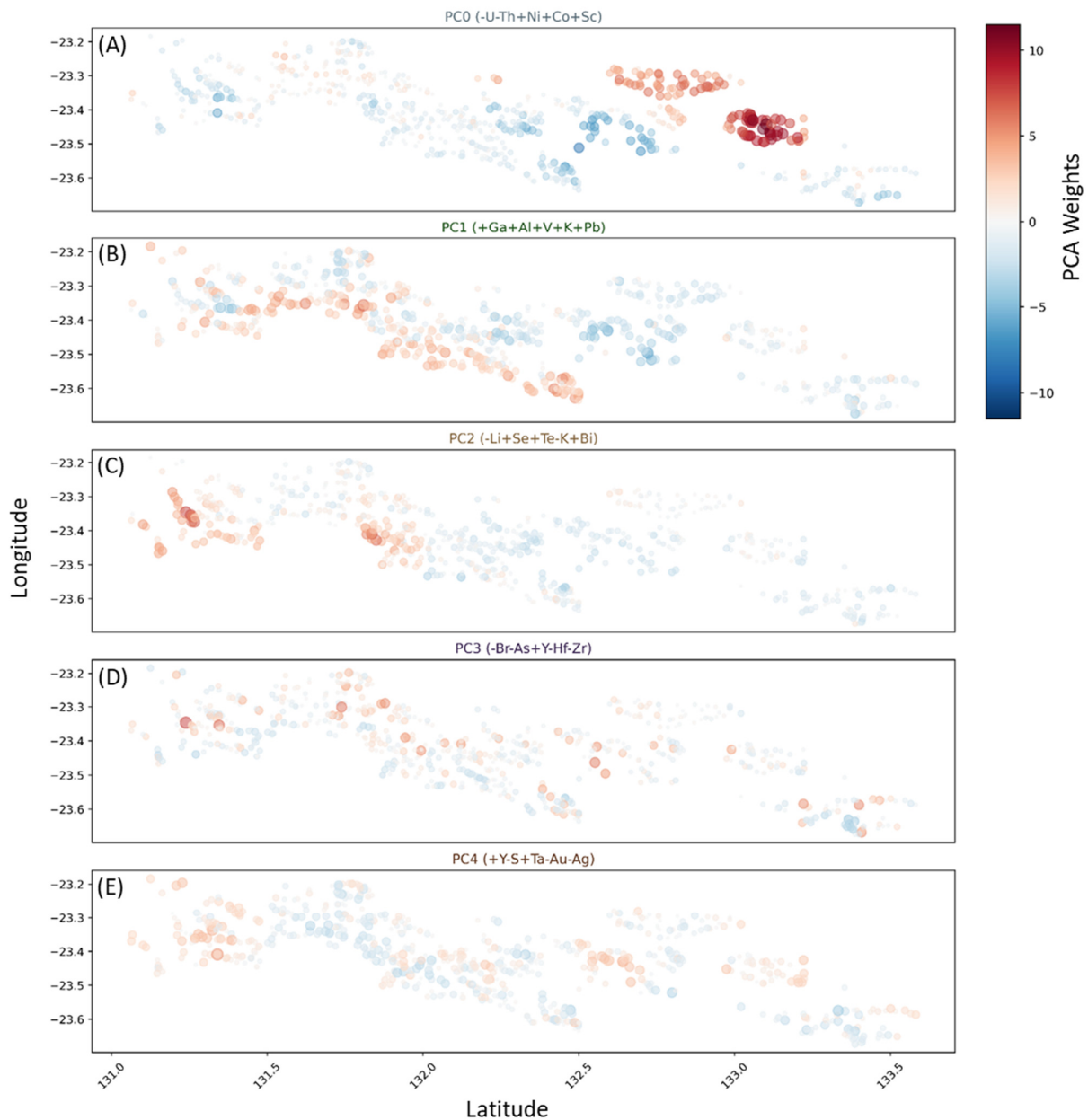


Figure 23: Automated output of the spatial distribution of principal components weighted by both colour and symbol size (absolute magnitude). The top five elemental loadings (highest influence) for each principal component are indicated as headings. The colour red indicates a positive component weight (association); the colour blue indicates a negative component weight. The larger the symbols the stronger the association. (A) Spatial distribution of principle component 0 weightings. (B) Spatial distribution of principle component 1 weightings. (C) Spatial distribution of principle component 2 weightings. (D) Spatial distribution of principle component 3 weightings. (E) Spatial distribution of principle component 4 weightings.

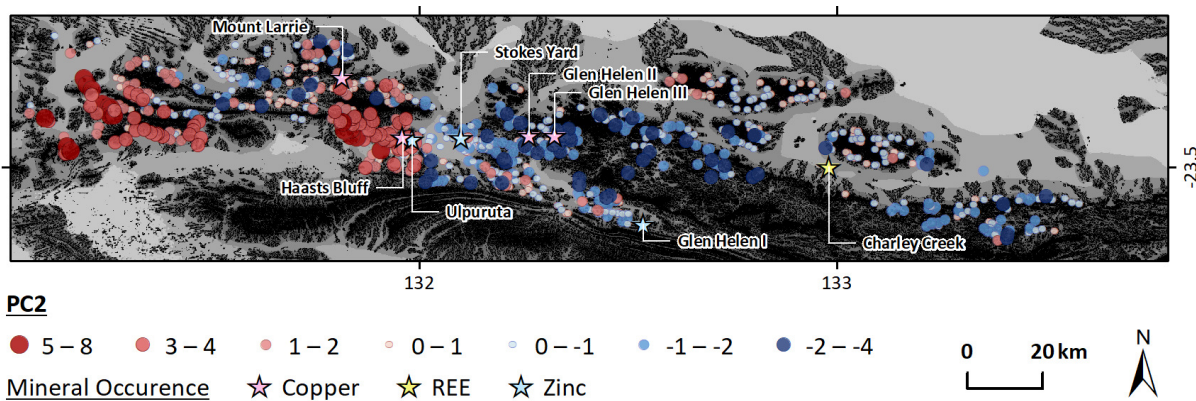


Figure 24: Spatial distribution of principal component 2 coincides broadly with known Au and Zn-Cu-Pb mineral occurrences Haasts Bluff, Ulpuruta, Mount Larrie and Stokes Yard, but not with Glen Helen II and Glen Helen III.

3.5 Other soil properties

In addition to the multi-element geochemistry suite, several other soil parameters that are related to improving the understanding of metal mobility have been added to the UltraFine+® Next Gen Analytics workflow. These include pH, conductivity, particle size distribution, and visible and near infrared and thermal infrared mineral proxies (VNIR and FTIR). This data is also intended to be used for the identification of soil property trends that are related to landscape settings to enable the interpretation of geochemical results within context, especially with respect to the potential of better explaining false positive results. Particle size distribution and FTIR were not available for the MacDonnell Ranges dataset.

3.5.1 pH

The mobility (solubility and precipitation) of many metals, including transition metals, such as the elements of interest in the MacDonnell Ranges project area Zn and Cu, are driven by pH and redox state. The pH of soils is mainly dependent on the mineralogy of the parent material and is, for example, expected to be more acidic around sulphide-rich deposits and more alkaline where parent material has higher buffering capacity, e.g., via silicate or carbonate mediated buffering during weathering and soil formation. Soil pH can therefore be a useful indicator for broad lithological changes, and potentially indicate certain types of mineral occurrences. Hence, it can be a useful tool to confirm landscape types to provide better context for geochemical sample interpretation.

Despite the influence of pH on soil chemistry, especially adsorption and ion exchange capacity of trace metals on clay, no correlation between pH and metal distributions were observed in the MacDonnell Ranges project area. Sample pH in the MacDonnell Ranges project area is circumneutral with very few values below pH 6 or above pH 10 and a mean pH of 8.0. The spatial distribution of pH measurements broadly follows the variation in parent material indicated by PC1 (see Section 3.4; Figure 25) with generally higher pH in outcropping and residual saprock and saprolite and lower pH in colluvial and alluvial transported cover. The patterns in the spatial distribution of pH in these samples is also influenced by the sampled material. Stream sediment samples are not a well-developed sample medium for the UltraFine+® method. Furthermore, at

the time of analysing these samples, the method for recovering consistent pH measurements from samples was still under development and improvements have since been made to the pH measurements to ensure greater consistency.

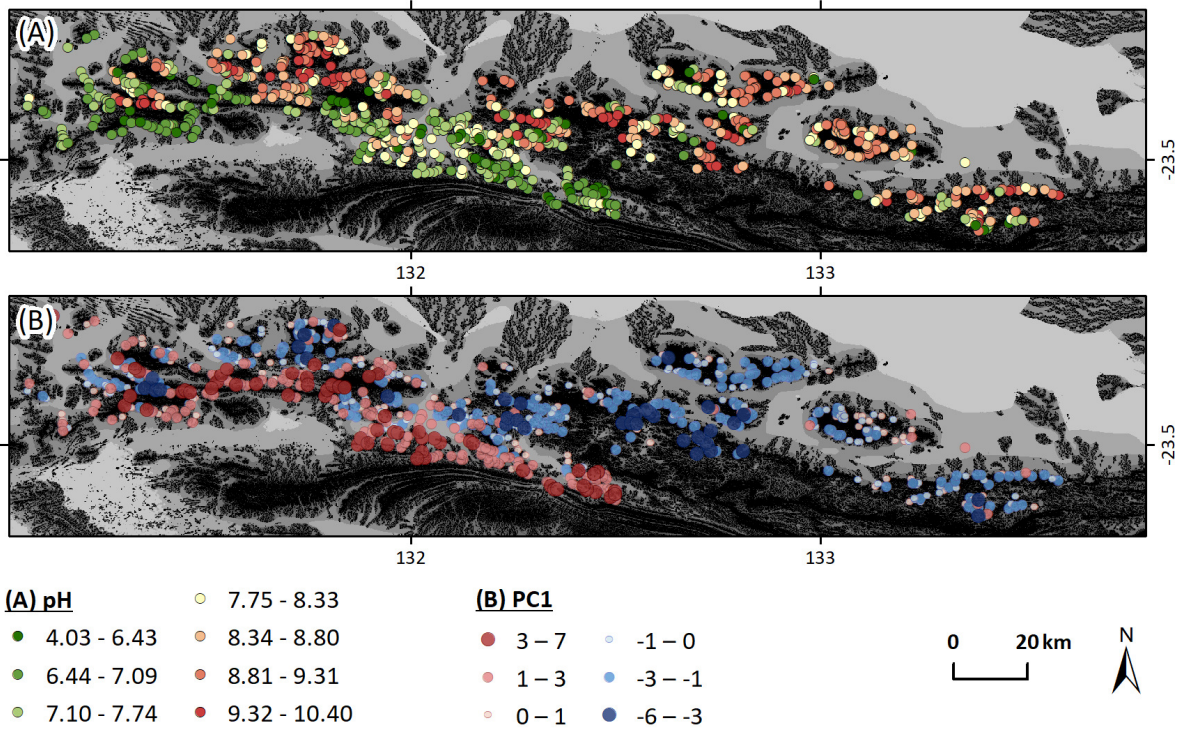


Figure 25: Variation in sample pH (A) appears to broadly correlate with the variation of underlying parent material captured in principal component 1 (B).

3.5.2 VNIR

In addition to pH and redox state, the concentration of many metals in soil samples is also related to the composition of a given sample, such as the abundance of secondary phases including iron oxides and clays. Near infrared spectral measurements, which can indicate key mineral groups and provide abundance estimates for iron oxides and kaolinite abundance are part of the UltraFine+® analysis. While some trends can be observed, such as higher iron oxide abundance in transported cover in the western part of the project area compared to the lower iron oxide abundance in the outcropping and residual landscapes to the east (Figure 26C), this method was not designed for stream sediment samples and interpretations should be regarded with caution.

The main mineral groups (Mineral group 1 and 2) detected in the visible to shortwave infrared region, report only the dominant mineral-group that contributes >51 % to the spectral unmixing algorithm. This does not report individual minerals but mineral groups (i.e., white mica rather than muscovite). The mineral groups reported with the UltraFine+® method are Kaolin, Smectite, White-Mica, Amphibole, Chlorite, Dark-Mica, Al-bearing minerals, Mg-bearing minerals, Carbonate and Sulfate. The stream sediment samples in the MacDonnell Ranges project area are dominated by kaolinite, smectite and white mica in both the primary and secondary mineral groups (Figure 26A, B).

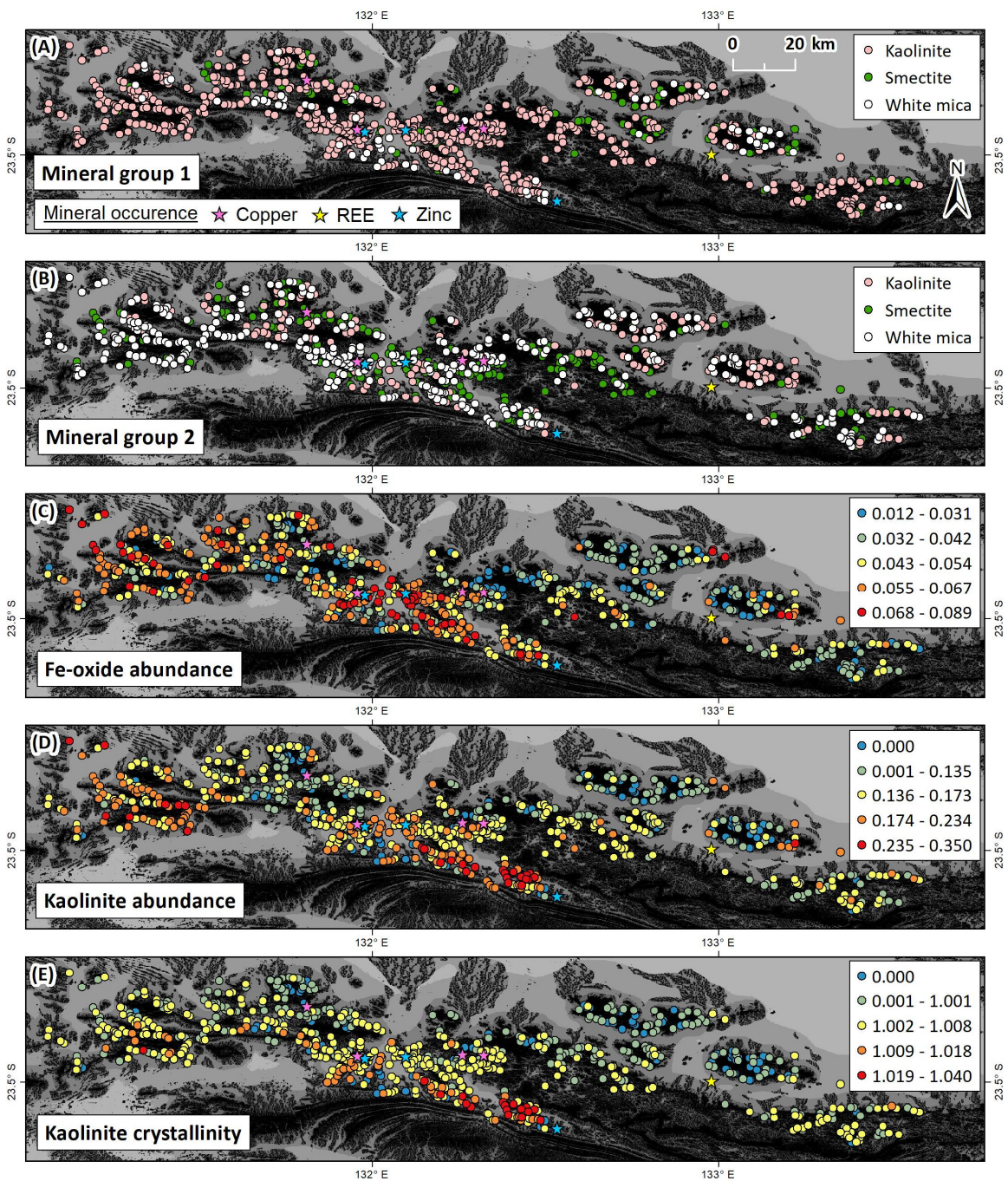


Figure 26: Spatial distribution of VNIR results over the MacDonnell Ranges project area. (A) Primary mineral group (contributes >51 % to the spectral unmixing algorithm). (B) Secondary mineral group. (C) Relative iron oxide abundance. (D) Relative kaolinite abundance. (E) Kaolinite crystallinity.

Cations are commonly loosely adsorbed onto clay interlayer surfaces in soils (e.g., Cu to vermiculite, Furnare et al. 2005) and can be effective metal sinks in soils (e.g., ore grade soils, Nieto et al. 2011). Swelling clays, such as smectites, in particular, have negatively charged interlayer surfaces due to ion substitution, which attracts water and hydrated cations. Due to its different structure, kaolinite (a non-swelling clay) does not attract metals into interlayer spaces (Farrokhpay et al. 2016). It is therefore worth considering whether an anomaly of a metal of interest is present in a swelling or non-swelling clay. With the ability to report main clay mineral

groups, the UltraFine+® method now enables normalisation of geochemical results against clay species.

Many trace metals are also readily adsorbed to clays and iron (oxyhydr)oxides, especially those with a high surface area. For example, divalent cations such as Ni and Zn precipitate by adsorption to iron oxides under circumneutral pH conditions, which are the prevalent conditions within the MacDonnell Ranges project area.

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. An example for the MacDonnell Ranges is presented in Figure 27 as a demonstration of the approach, where Zn concentrations have been normalised with relative iron oxide abundance.

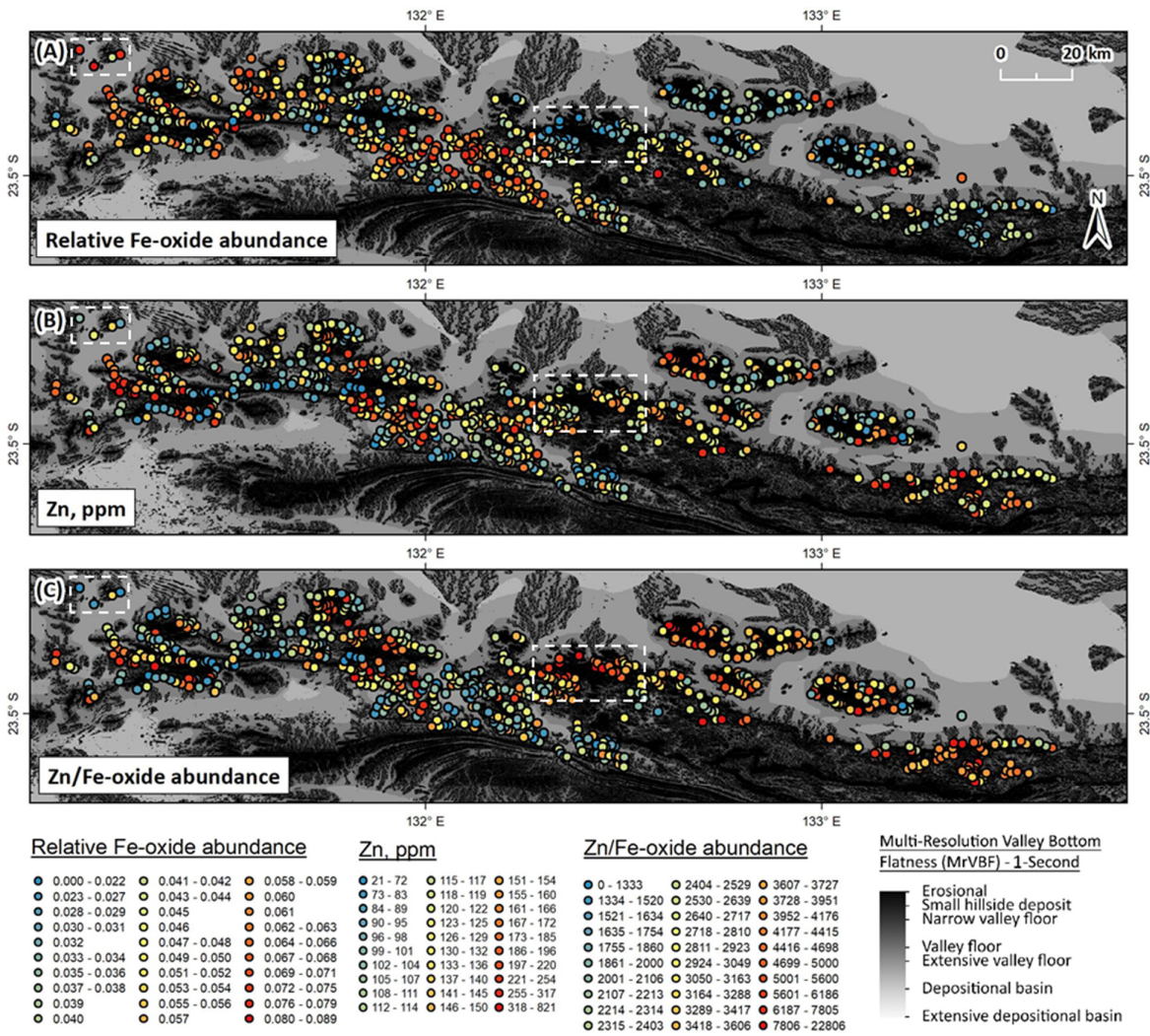


Figure 27: Example of normalising geochemical data with VNIR analyses. (A) Relative iron oxide abundance. (B) Spatial distribution of Zn concentrations. (C) Spatial distribution of Zn normalised to iron oxides. Dashed white boxes indicate areas of interest.

4 Summary

The Northern Territory Geological Survey provided 785 stream sediment samples to the UltraFine+® Next Gen Analytics research project in 2020. These samples were part of a larger background study on historical stream sediments carried out in 2001. The MacDonnell Ranges project site was chosen as a first-generation, large-scale trial site with a focus on principal functionality of the workflow rather than a focused effort to generate new tangible exploration targets and areas.

The samples submitted to this research project were previously analysed via four-acid digest and 24h cyanide leach (Au only) and a comparison between these methods and the UltraFine+® method showed that improved recovery of elements from the ultrafine fraction resulted in higher concentrations of elements, effectively reducing the number of results below the detection limit, which enables the delineation of subtle geochemical enrichments for elements. The detection of these subtle variations is particularly relevant for exploration through transported cover. The UltraFine+® method also provides 18 additional pathfinder elements for mineral exploration. For metals of interest within the MacDonnell Ranges project, such as Au, Cu, Pb and Zn, the UltraFine+® method resulted in higher recovery with overall recoveries between 2 times (Cu) to more than 5 times (Zn) higher in the UltraFine+® dataset. There was little to no correlation between the analysis methods, reflecting element partitioning in size fractions and the different approaches to sampling and analysis.

This report presented some example outputs of the UltraFine+® Next Gen Analytics workflow with the goal of utilising machine learning to integrate spatial data and soil properties to provide landscape context for geochemical data and a basic, first-pass data interpretation. These outputs included proxy regolith landscape clusters, maps and boxplots of elemental outliers by landscape type and exploration indices. The data presented herein was analysed in September 2020. Since then, the components of the UltraFine+® workflow have undergone continuous improvements. Therefore, some components such as analysis results for FTIR, sizing and Pd concentration and related outputs, such as soil texture diagrams, dispersion direction, regolith indices and catchment analyses were not available for this project.

A variety of clustering methods were trialled to generate landscape proxies for the MacDonnell Ranges project and the recommended outputs for the MacDonnell Ranges project are those produced via an agglomerative algorithm with eight landscape clusters (agg8). Due to the large size of the MacDonnell Ranges project area and the resulting complexity of the landscape an even larger number of clusters would have resulted in a more detailed approximation of the complex regolith types. However, the number of landscape clusters are the result of a balanced approach to represent the major landscape types of the area while also enabling meaningful interpretation of geochemical data. Despite these constraints, the resulting output for the MacDonnell Ranges project area provides a more detailed landscape context than publicly available interpreted regolith products.

It is important to note that while the landscape clusters generated for the MacDonnell Ranges project area are appropriate proxies for the regolith types within the area, the UltraFine+® method

and related workflow were not designed for stream sediment samples and outputs should be interpreted with caution. Future outputs of the workflow are likely to have a catchment analysis interpretation that will make the current approach more valuable for stream sediment sampling programs.

While the sample type and the availability of analyses limited the benefit of the UltraFine+® Next Gen Analytics workflow for exploration targeting in this project, the MacDonnell Ranges project site was crucial in developing the approach, workflow and outputs for the UltraFine+® Next Gen Analytics for Discovery research project and the results of this project provide a good background data set for future exploration activities in the region and surrounding areas.

Appendix A

A.1 Surface geology legend for Figures 3, 6, 8 -11

Surface Geology

QUATERNARY

- Alluvium
- Dunes
- Colluvium
- Lake deposits

CENOZOIC

- Conglomerate
- Waite Formation / limestone
- Calcrete
- ferruginous duricrust
- Sand plain
- Silcrete

PALEOZOIC

- Retrograde schist zones

DEVONIAN

- Brewer Conglomerate,
Hermannsburg Sandstone

SILURIAN-DEVONIAN

- Mereenie Sandstone

CAMBRIAN-ORDOVICIAN

- Larapinta Group

CAMBRIAN

- Pertaorrrta Group

PROTEROZOIC

- Dashwood Gabbro Complex

NEOPROTEROZOIC

- Sedimentary siliciclastic +
sedimentary carbonate

MESOPROTEROZOIC-NEOPROTER...

- Old Hamilton Downs Gneiss

MESOPROTEROZOIC

- Teapot Granite Complex

PALEOPROTEROZOIC

- Andrew Young Igneous Complex
- Ellery Granitic Complex
- Narwietooma Metamorphic
Complex
- Mafic metamorphic rocks
- Madderns Yard Metamorphic
Complex
- Yaya Metamorphic Complex

Appendix B

B.1 UltraFine+® Next Gen Analytics NTGS data package 2022

Appendix C

C.1 UltraFine+® Standard – UFF 320.xlsx

References

- Bardwell, N., and Gray, D.J., 2016. Hydrogeochemistry of Northern Territory: Data Release: Accompanying Notes. CSIRO, Australia. EP156405 21p
- Bureau of Meteorology, 2022. Online climate data for Alice Springs Airport station, http://www.bom.gov.au/climate/averages/tables/cw_015590.shtml [last accessed April 2022].
- Craig, M.A., 2006. Regolith map of the Northern Territory, 1:2 500 000 scale. Northern Territory Geological Survey, Darwin. https://data.nt.gov.au/dataset/strike---northern-territory-regolith-landforms-2500k/resource/04f4e399-f446-4309-a1ec-f2233849f71f?inner_span=True
- 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.
- Farrokhpay, S., Ndlovu, B., Bradshaw, D., 2016. Behaviour of swelling clays versus non-swelling clays in flotation. *Minerals Engineering*, Volumes 96–97, 59-66, <https://doi.org/10.1016/j.mineng.2016.04.011>.
- Furnare, L.J., Vailionis, A., Strawn D.G., 2005. Polarized XANES and EXAFS spectroscopic investigation into copper(II) complexes on vermiculite.
- 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. Record 1. Geoscience Australia, Canberra. <http://pid.geoscience.gov.au/dataset/ga/72759>
- Geoscience Australia, 2012. Surface Geology of Australia, 1:1 000 000 scale, 2012 edition. Bioregional Assessment Source Dataset. <http://data.bioregionalassessments.gov.au/dataset/8284767e-b5b1-4d8b-b8e6-b334fa972611>
- Gray, D., Reid, N., Noble, R., Throne, R., Giblin, A., 2019. Hydrogeochemical Mapping of the Australian Continent. CSIRO, Australia. Report EP195905 110 p.
- 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](https://doi.org/10.1016/S0375-6742(97)00046-0).
- Isbell R.F., 2021. National Committee on Soil and Terrain. The Australian Soil Classification. 3rd edn. CSIRO Publishing, Melbourne
- McInnes, L, Healy, J., Melville, J., 2018. "Uniform manifold approximation and projection for dimension reduction". arXiv:1802.03426.
- Nieto, F., Suárez, S., Martín, F.J., Velasco, F., 2011. Serpentine and chlorite as effective Ni-Cu sinks during weathering of the Aguablanca sulphide deposit (SW Spain). TEM evidence for metal-

retention mechanisms in sheet silicates. *EU J. Mineral.*, 23 (2011), pp. 179-196, 10.1127/0935-1221/2011/0023-2084

Noble, R., Lau, I., Anand, R., Pinchand, T., 2018. MRIWA Report No. 462: Multi-scaled near surface exploration using ultrafine soils: Geological Survey of Western Australia, Report 190, 99p.

Noble, R.R.P., Lau, I.C., Anand, R.R. and 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>

Poudjom Djomani, Y., Minty, B.R.S., 2019. Radiometric Grid of Australia (Radmap) v4 2019 filtered pct potassium grid. Geoscience Australia, Canberra. <http://dx.doi.org/10.26186/5dd48d628f4f6>

Poudjom Djomani, Y., Minty, B.R.S., 2019. Radiometric Grid of Australia (Radmap) v4 2019 filtered ppm thorium. Geoscience Australia, Canberra. <http://dx.doi.org/10.26186/5dd48e3eb6367>

Poudjom Djomani, Y., Minty, B.R.S., 2019. Radiometric Grid of Australia (Radmap) v4 2019 filtered ppm uranium. Geoscience Australia, Canberra. <http://dx.doi.org/10.26186/5dd48ee78c980>

Raymond, O.L., Liu, S., Gallagher, R., Highet, L.M., Zhang, W., 2012. Surface Geology of Australia, 1:1 000 000 scale, 2012 edition [Digital Dataset]. Geoscience Australia, Commonwealth of Australia, Canberra. <http://www.ga.gov.au>

Warren, R.G., Shaw, R.D., 1995. HERMANNSBURG, Northern Territory. 1:250 000 geological map series and explanatory notes, SF53-15. Northern Territory Geological Survey.

Wilford, J., Roberts, D., 2019. Weathering Intensity Model of Australia. Geoscience Australia, Canberra. DOI:10.26186/5c6387a429914.

Woolrych, T., Batty, S. D., 2007. A Semi Automated Technique to Regolith-Landform mapping in West Africa. *ASEG Extended Abstracts*, 2007(1), 1–4. <https://doi.org/10.1071/ASEG2007ab167>

As Australia's national science agency and innovation catalyst, CSIRO is solving the greatest challenges through innovative science and technology.

CSIRO. Unlocking a better future for everyone.

Contact us

1300 363 400
+61 3 9545 2176
csiroenquiries@csiro.au
csiro.au

For further information

Mineral Resources
Ryan Noble
+61 8 6436 8684
ryan.noble@csiro.au
csiro.au/MineralResources

Mineral Resources
Anicia Henne
+61 8 6436 8697
anicia.henne@csiro.au
csiro.au/MineralResources