

How to use *UltraFine+[®] Next Gen Analytics* – Principal Component Analysis (PCA)

What is Principal Component Analysis?

Principal Component Analysis (PCA) is a statistical technique that we use to reduce the variables (also called dimensions) of your geochemical dataset. In essence, we are transforming the 52 dimensions (elements) into a new coordinate system to describe the data with fewer dimensions (only five in this case) while preserving as much of the information (variation) in the dataset. This makes the data easier to visualise and to work with and can be used to identify patterns in your geochemical dataset using multielement trends. The PCA outputs allow for a rapid, first-pass identification of element association and potential exploration indices within the dataset.

How are the principal components calculated?

Principal Component Analysis (PCA) is performed on centred-log-ratio and quantile-normalised transformed geochemical data for each of your soil samples. Each data point (soil sample) is reduced from n dimensions (number of analysed elements) into five principal components (PC0 to PC4). If more than 10 % of analyses for a given element are missing (e.g., element has not been analysed) this element is not included in the PCA calculations. If less than 10 % of analyses of a given element are missing, the element is included in the PCA analysis, but no principal components are calculated for the affected sample that has that data missing. All values below the detection limit are replaced with half the detection limit value and included in the PCA calculations.

Where do I find the PCA for my project site?

You will find your PCA results in your *Data Package* under *3 Landscape Models* (kmeans4, agg8 or agg12) → *5 Principal Component Analysis*. In this folder you will find a summary PNG file with a spider diagram and spatial PCA plots. The PCAs are also available as shapefiles (under *GIS*) and as a CSV file (ending on *_pca_component.csv*).

There are also two additional CSV files (ending on *_pca_eigenvalue.csv* and *_geochem_components.csv*).

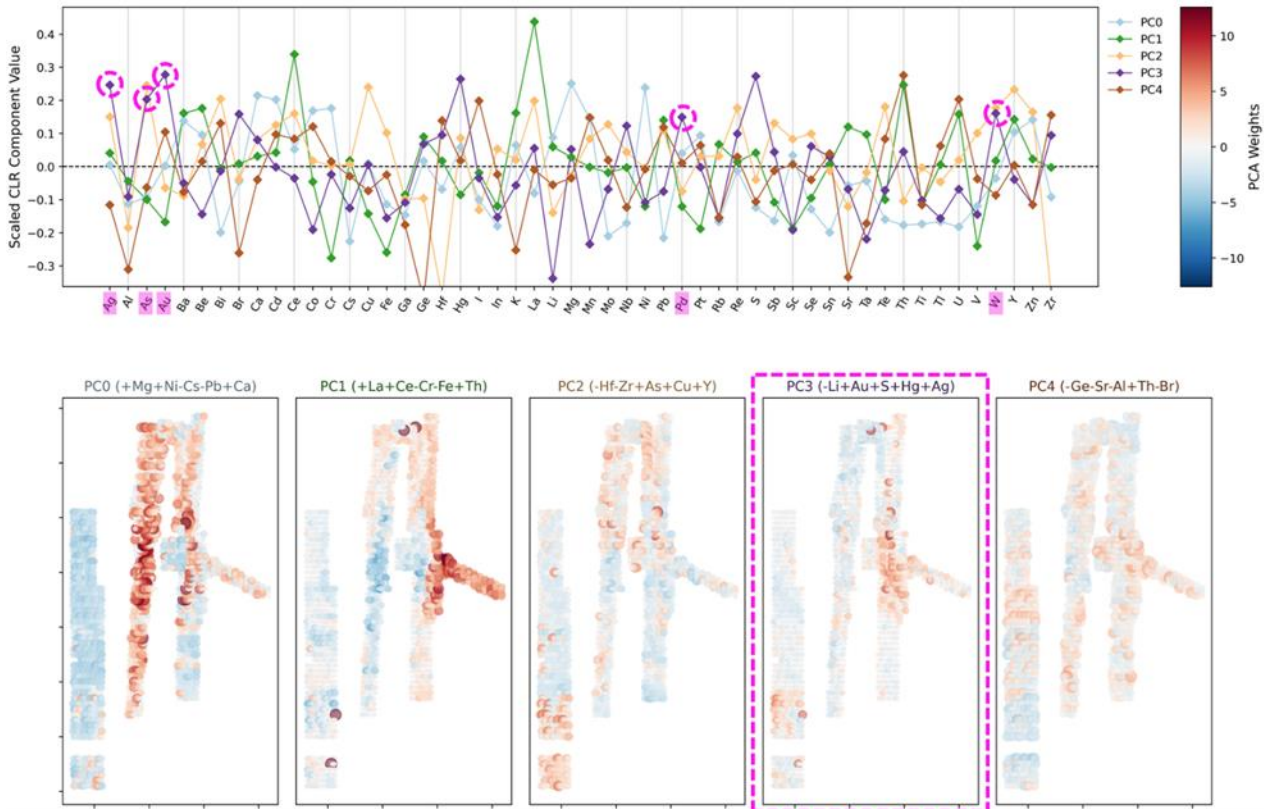
Tip: As a short cut, you can also toggle your PCAs on and off in your Digital Sample Observer (DSO).

How do I interpret my PCA outputs?

The workflow generates five principal components, PC0 through to PC4. The spider diagram and spatial plots in the PNG file will give you a rapid overview of the principal components calculated from your data. The spider diagram shows the loadings (weightings) of each element for each of these five principal components. This illustrates the general geochemical affinity of each principal component. The spatial distribution of each of the principal components by sample is also displayed in the PNG file below the spider diagram. The explained variance of the principal components decreases from PC0 to PC4. Lower principal components will, therefore, often pick out large-scale lithological variation as the major component(s). Principal components with exploration potential usually explain less variance in the data than the major elements related to geological influence. Hence, in many areas, PC2 and PC3 (3rd and 4th components) may represent the more relevant components in the context of mineral exploration. In some cases, signatures relating to mineralisation will constitute relatively low proportions of variance and thus will not be represented by principal components. If you want to know how much of the variance in your data is explained by each principal component, the CSV file ending on *_pca_eigenvalues.csv* contains this information as a percentage for each principal component. Once you have identified principal components of interest, you can drag and drop the shapefiles into your preferred GIS software for in-depth interrogation.

How do I read the PCA spider diagrams?

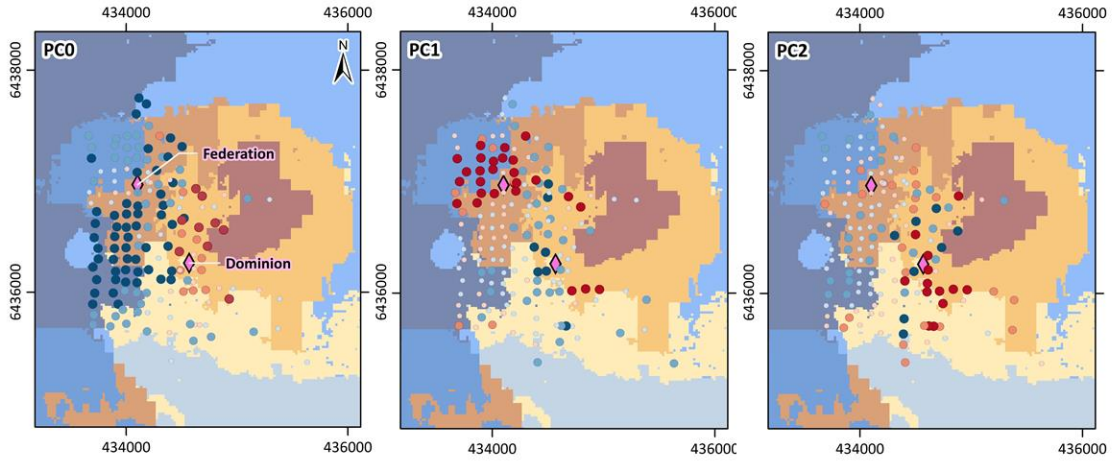
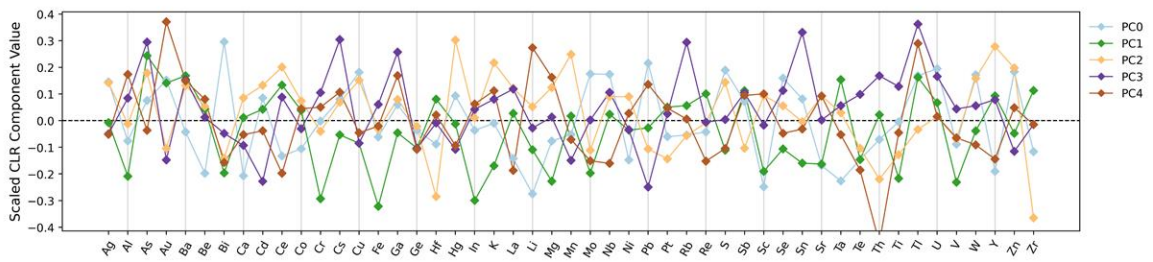
Your spider diagram displays loadings for each principal component by a coloured diamond for each element. These are relative and show whether the principal component is positively or negatively associated with this element. All elements included in your PCA analysis are sorted alphabetically from left to right and we have added visual guiding lines (in grey). The further a diamond plots away from the 0 line (dashed black line), the greater the loading for (influence on) the specific principal component. In the example below an interesting principal component with relatively high loadings for Ag, As, Au, Pd and W (indicated in pink circles and shading) is PC3 (purple diamonds connected by a purple line).



The boxes below the spider diagram show the spatial distribution of each of the five principal components weighted by both colour and symbol size. The top five elemental loadings (the greatest influence on this principal component, whether positive or negative) for each principal component are indicated in the headings above these boxes. The colour red indicates a positive component weight (association); the colour blue indicates a negative component weight (association). The larger the symbols the stronger the association.

Example – Federation and Dominion Prospects

In the example below of the first three PCs from the Federation Project site (read the full report [here](#)), the first principal component (PC0) explains 46 % of the variability within the dataset and, unsurprisingly, highlights shallow geology and climatic influence (slope aspect or elevation and matching the landscape proxy types/colours in the background). PC1 explains 18 % of the variability within the dataset and is associated with As, Au, Ta and Tl. PC2 only explains 15 % of the variability within the dataset, but also shows exploration potential with association of Ag, As, Cd, Cu, Hg, Ni, S, W and Zn. It therefore has positive loadings with some target and pathfinder elements such as Ag, Cu and Zn, but not with Au or Pb. PC2 also is strongly influenced by Mn in the samples. In retrospect, it appears that PC1 highlights the Federation mineralisation (downslope from and immediately to the northwest of the Federation prospect) and PC2 highlights the Dominion prospect (positive PC2 scores coincide broadly with the Dominion prospect and are positive, but are more subdued around Federation).



PC0

- 8.3 - 14.7
- 3.9 - 8.2
- 1.2 - 3.8
- -0.6 - 1.1
- -1.9 - -0.7
- -3.5 - -2.0

PC1

- 2.1 - 7.8
- 1.6 - 2.0
- 0.1 - 1.5
- -1.4 - 0.0
- -2.9 - -1.5
- -5.9 - -3.0

PC2

- 3.1 - 6.8
- 1.6 - 3.0
- 0.1 - 1.5
- -1.4 - 0.0
- -2.9 - -1.5
- -4.6 - -3.0

Proxy regolith types (agg8)

- 1 Shallow residual soil, minor outcrop
- 2 Weathered shallow residual soil
- 3 Sides lope - colluvial
- 4 Shallow sheetwash
- 5 Deeper sheetwash
- 6 Lowland residual soil/Sheetwash
- 7 Deeper transported cover 1
- 8 Deeper transported cover 2

◆ Prospect

