Applications of Geochemistry in Targeting

With Emphasis on Large Stream and Lake Sediment Data Compilations

Paul D. Agnew[1]

[1]*Rio Tinto Exploration Pty. Limited.*

Geochemistry is a powerful, empirical exploration tool used by the vast majority of exploration programs in one form or another. For decades exploration companies as well as State and Federal Geological Survey organizations from all over the world have collected and assayed stream sediments, lake sediments, tills, soils, lags and rocks on a local/prospect scale and on a regional scale. These analytical data are rarely destroyed or lost; they accumulate over the years, their geographic distribution gradually expanding, and the volume of data increasing exponentially with the proliferation of commercial laboratories offering rapid, inexpensive multi-element analysis. Once closely guarded, the myriad of geochemical databases which house these geochemical assays are increasingly being shared amongst exploration companies and Geological Surveys are releasing vast quantities of geochemical data at little or no cost. In many parts of the world it is now possible to compile geochemical databases which cover entire geological provinces and with nothing more than a personal laptop computer begin to interpret and visualize geochemical information on a very large scale, often with surprising implications for regional target generation. We have entered a new era in geochemical exploration where many exploration companies have, or at least have access to, the same geochemical data. It is the manner in which these data are interpreted and visualized that increasingly provides the competitive advantage and the exploration opportunities.

Geochemical Data and the Information Revolution

Anybody who has scanned the web sites of Geological Surveys in many parts of the world will be aware that with the click of a button they can download regional geochemical databases at no cost, or browse survey metadata and purchase geochemical data at a tiny fraction of the original collection and analysis cost. Easy access to data via the internet is nothing new but the storage space and processing power of ordinary personal computers is only recently beginning to match the volume of data available. Entire corporate geochemical databases can now be stored on a personal computer and merged with databases from other organizations without overloading the hard disk. The processing power of modern PC's outperforms the supercomputers of just a decade ago. We have database, GIS and statistical analysis software which can handle hundreds of thousands of records with ease. All of the tools are available to compile and interpret geochemical data on a massive scale; the question is will it be worth the effort? What role do large geochemical data compilations play in identifying exploration targets and what tools and techniques are required to identify and prioritize the best exploration opportunities?

Limitations of Historical Geochemical Data

It is reasonable to assume that the geochemical data in any large compilation are "consistently inconsistent". As each new data set is added, the number of inconsistencies in the compilation increase as each of the following parameters varies between surveys and sometimes within individual surveys;
– Sample media, collection technique and sampling density
– Size fraction analysed
– Sample preparation and digestion
– Analytical techniques and the sophistication of analytical instrumentation
– Quality control protocols and calibration of analytical instrumentation
– Analytical suite

Variations in these critical operational parameters are commonly not recorded with the geochemical data and as time goes by, it becomes increasingly difficult to find out exactly what was collected and how it was prepared and analysed. In addition to this there is rarely any consistency in the type and format of field observations recorded at sample sites and stored with the geochemical data. Coded observations are common, but the associated look-up tables for the codes are disappointingly rare.

The basic principles of geochemistry tell us that different sample media, sampling protocols and analytical techniques result in different background levels and different responses to mineralisation and lithological variation. It certainly does not make sense to combine different sample media and in many cases, combining geochemical data from the same media from different surveys is also, understandably, thought not to be a particularly useful or reliable approach to geochemical target identification. Therefore, common practice is to either interpret geochemical data one survey at a time or to first identify an area of interest based on geological and/or geophysical criteria and then to interpret available geochemical data over a much smaller area where survey parameters are likely to be more consistent. However, for those explorers who want to understand the "geochemical landscape" on a much larger scale, combining different sources of geochemical data from the same sample medium is a necessity. The purpose of this paper is to outline some of the basic tools available to assist with the leveling, interpretation and visualisation of large geochemical data compilations, to provide some insights into the types of geochemical patterns which might emerge when data are visualized on a sub-continental scale and to propose how these patterns might play an important role in target generation for a range of commodities.

## Tools for the Interpretation of Large Geochemical Data Compilations

Geochemical Images and Data Leveling
Geochemical images are an invaluable tool for both initial identification of significant base level shifts between and within surveys and for recognition of patterns of anomalism, enrichment and depletion at all scales. On a regional scale, the gridding of stream and lake sediment data provides a "workable" interpolation of metal concentrations between data points. However, at a smaller scale, particularly with stream sediments, the geochemical images should not be considered an interpolation of metal concentrations between individual sample sites but a visualisation tool to highlight clusters/groups of samples or individual samples with anomalous or elevated results. The source regions for these samples need to be interpreted from the drainage pattern and topography, not from the image itself, using GIS software to overlay sample sites, drainage and other topographical information on the geochemical image.

A geochemical image produced from an unleveled geochemical data compilation often looks like a patchwork quilt displaying patterns that are clearly not geological. Thematic maps displaying various data source and survey parameters can be overlaid on such images to identify which are causing base level shifts for which elements. In many cases, simply overlaying map sheet boundaries will clearly identify individual surveys or parts of surveys where sampling, or more commonly, analytical parameters, have varied resulting in base level shifts relative to adjacent map sheets. In such cases, the map sheet name can be used as a grouping parameter for quantification of the base level shift and leveling. Thematic maps which should be reconciled against the raw geochemical images include data source, the year/date samples were collected, size fraction analysed, digestion and analytical techniques and analytical batch number. In all cases, the objective is to correlate non-geological patterns on the images with a survey parameter which can then be used for grouping and leveling of the data. As a last resort, where operational parameters are not systematically recorded in the compiled database, it may be necessary to manually flag groups of samples with dramatically different background levels for each element and then use the flags for grouping and leveling although this can be a complex and time consuming exercise.

Once survey parameters which relate to significant base level shifts have been identified, the parameters can be use to split the geochemical data into subsets. Side by side box plots of each of these subsets provide a useful visually based approach for quantifying the degree to which 50th, 75th and 90th percentile levels vary between each group. It is not uncommon to see median levels from one map sheet or survey, approach or even exceed 90th percentile levels from other surveys, (see figure 1). In some instances, the differences observed between groups may actually relate to geological background, so it is emphasized that comparison of data subsets must not rely on percentile shifts alone, it is important to

relate the shifts to patterns that are not geologically controlled using geochemical images and geological maps.
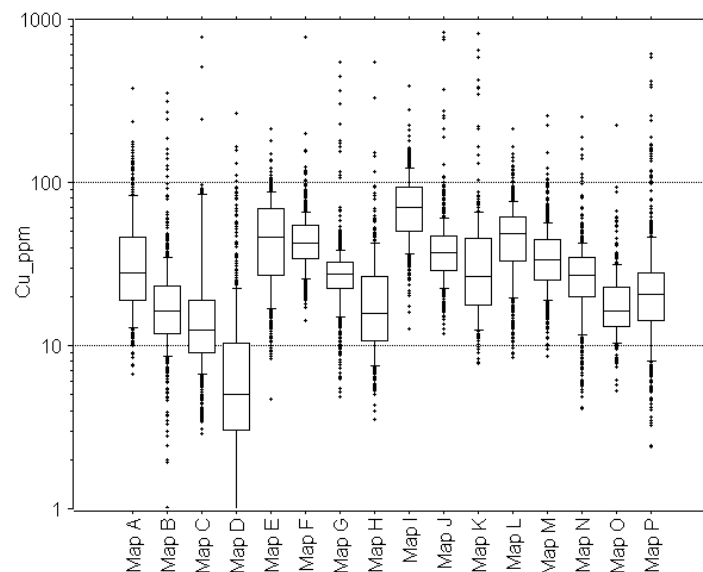
Figure 1 – Comparison of box plots of Cu in stream sediments from sixteen adjacent 1:250,000 scale map sheets, (n = 8,951 samples). The samples from each map sheet were collected at different times and some were prepared and analysed using different techniques, resulting in dramatic shifts in median levels (central line in each box) between some sheets. Note how the 90th percentile (top "whisker") for some sheets is below median levels on other sheets. Such dramatic background level variation is unlikely to be geologically controlled.

There are numerous approaches to the leveling of geochemical data, none of which can be effective if the source of the base level shift has not been identified and the database consistently coded such that it can be subset into coherent groups which can then be leveled relative to each other. Standard scores (z-scores) are a simple but useful tool for geochemical data leveling. They are calculated by subtracting the mean of each data subset from the values in that subset, then dividing by the standard deviation of the subset. As such, data from each subset are represented as standard deviations from a mean of zero. This would usually be calculated on log-transformed data to ensure subset data distributions are approximately gaussian before mean and standard deviation values are calculated.

In some instances multiple phases of leveling may be necessary to remove the non-geological patterns evident on the raw geochemical images of a regional data compilation. For example, it may be necessary to first level by the source of the data (eg. Geological Surveys from states A, B, C and D), then by individual map sheet areas, each of which may constitute a phase of a sampling program which could have run for many years (with changes in analytical instrumentation and quality control generating the base level shifts), and finally by the size fraction analysed which may have varied within and between individual surveys. Note that if there is no significant difference between subsets at any phase of leveling, the leveling will have no effect. It is also noted that some of the more subtle differences between subsets will invariably be a result of natural geological background variation and that the leveling process tends to subdue or even remove such differences. This is an unavoidable limitation of geochemical data leveling which can adversely impact the geochemical mapping of regional-scale lithochemical patterns. However, leveling is unlikely to significantly affect geochemical responses associated with mineralisation since mineralisation is a rare and generally spatially isolated geological occurrence which is unlikely to influence geochemical background levels over large areas.

Geologically and Geographically Controlled Data Selection

One of the key advantages of compiling diverse sets of regional stream and lake sediment geochemical data covering large areas is the capacity to rapidly select using GIS, geologically or geographically controlled subsets of the data for more detailed interpretation. Once the data have been leveled, the raw concentrations are less important than the relative magnitude of the z-scores compared to the other samples in the subset. As such, the basis on which a data subset is selected can have a significant influence on the patterns of enrichment and anomalism observed when visualizing the data. With the proliferation of geological, structural and geographic data now available in GIS format, with relatively simple buffering and query commands it is possible to select geochemical data from areas which match any

number of geological and geographical settings favorable for mineralisation. For example, consider the capacity to rapidly select all available geochemical data within;

– A particular geological province of part of a province
– A prospective geological unit or lithotype (or within a specified distance of such a unit)
– A specified distance from major structures with a specified orientation range
– A specified distance from coastlines and/or key infrastructure
– Land which is accessible to exploration and mining

Such data selection criteria frequently go beyond any one geochemical survey and having all available data together makes it possible to select data matching such criteria on a large scale.

Conversely, by deliberately not selecting data according to geological criteria which are inextricably tied up with ore genesis models which tend to "lead all of the sheep into the same paddock", the explorer who is prepared to include geochemical data from all terrains, including those deemed not prospective by competitors, is the explorer most likely to discover a new style of mineralisation or a previously unrecognized geological setting favorable for mineralisation.

"Signature-Based" Geochemical Data Interpretation

Regardless of the sources of data in a large scale geochemical data compilation, there is a high probability that many workers will have interpreted the data in the past. A properly leveled data compilation can be interpreted using traditional univariate techniques based on identification of anomalous threshold levels for target commodity elements and mapping of anomalous samples. Such an approach will invariably lead you to the high magnitude results and clearly has merit if you are searching for an unusual commodity which may not have been considered in the past. Alternatively, the leveling process may enhance results which in their raw form were not particularly outstanding as a result of an artificially subdued background, but in their leveled form represent genuine anomalies relative to the surrounding samples. However, when targeting for more mainstream commodities, it is generally prudent to assume that you will need to go further than simply locating the high results if you want to identify an exploration target which has not already been followed up by previous workers. At the very least you will need to drop the threshold levels in both target commodity elements and key pathfinder elements to identify more subtle responses which may have been overlooked in the past. Inevitably this leads to much larger numbers of potential targets, so an interpretation approach is required to identify the best opportunities amongst hundreds or even thousands of candidates, depending on the size of the area under investigation and the number of samples involved.

The basis of "signature-based" geochemical data interpretation lies in the empirical observation that large mineralised systems with a surface geochemical expression have a strong tendency to display anomalism in a range of elements and that the combination of anomalous elements is broadly correlated with the style of mineralisation. Key features of the "signature-based" approach (Agnew, 1999) include;

– Recognition of the dominant geochemical associations in a dataset.
– Relating these associations or "signatures" to specific geological, regolith and mineralisation controls.
– Selecting exploration targets based on coincident multi-element anomalism in prospective combinations of elements.
– Prioritizing targets in a GIS environment to identify those geochemical targets which are associated with the most favorable geological, structural and geophysical settings for mineralisation.

Factor analysis (Davis, 1973) is a powerful tool for identifying and quantifying the elemental associations in regional geochemical data. It is important that distributions for elements entered into the analysis are reasonably well defined and approximately gaussian. Leveled data will already have been log transformed prior to leveling. Log transformation is generally adequate for elements which have been analysed with detection limits sufficiently low enough to define a complete data distribution. Unfortunately, many low crustal abundance elements (Au, Ag, Cd, Bi, Mo, Pb, Sb and U amongst the more important elements for exploration) tend to have very poorly defined data distributions with the vast majority of results reporting below detection. In most cases such elements are not suitable for inclusion in factor analysis if a realistic result is desired. Results also tend to be more realistic when elements which are not correlating with other elements (indicated by low "squared multiple correlation" or loading heavily on one factor with no associated elements) are removed from the analysis.

Images generated from factor scores are a powerful tool for highlighting regional lithochemical and regolith controls on stream and lake sediment geochemistry and sometimes may also highlight minerali-

sation related geochemical associations. Unfortunately, when working with large data compilations, inconsistent analytical suites and inappropriate detection limits for many elements often dramatically reduce the number of samples where factor analysis can effectively be applied. Other techniques are required to highlight the critical geochemical associations which relate to mineralisation in the area of interest.

Mineral occurrence databases in GIS are the key to identifying local and regional mineralisation signatures. The process involves first selecting the types of occurrences of interest, with emphasis on larger deposits and operating mines and mapping these onto geochemical images of all available elements in the data compilation. Record a list of elements returning anomalous or elevated responses in the vicinity of each of the occurrences of interest. As will be discussed later, it is not critical that the responses are directly draining the deposit as the geochemical "footprint" of large mineralizing systems is often considerably larger than an individual deposit. In most cases, the sampling density will be low and sampling may not have directly targeted or tested many of the occurrences. However, over large areas with many mineral occurrences, it is generally possible to build a workable list of elements which are frequently associated with mineralisation of a particular style or styles.

In the absence of comprehensive mineral occurrence information it may still be possible to extract a mineralisation signature from the data by simply sorting a spreadsheet of all elements with the top 10% of results in each column made bold. By scanning across the top of the spreadsheet after sorting by each element in descending order, a reasonable indication of the geochemical associations at the "top end" of each element will often be apparent. The highly anomalous and outlier results for ore and pathfinder elements which come to the top of a spreadsheet when sorted in descending order often relate to samples directly affected by mineralisation and/or contamination from mining and old workings. As such, other elements in the top 10% which also come to the top of the spreadsheet are likely to provide some indication of the geochemical signatures associated with mineralisation in the area of interest. The technique is simplistic in the extreme but the results can be surprisingly effective. The key is to look for solid blocks of bold (top 10%) results from other elements when sorting by target commodity and associated pathfinder elements.

Data Visualisation for the Selection of Multi-element Anomalies
Once the key element combinations for target mineralisation styles have been established there are a number of data visualisation techniques which can assist in rapidly selecting potential target areas with coincident anomalism in the key element combinations. If the results of factor analysis have yielded a geochemical association which is clearly associated with mineralisation, then a geochemical image of the factor scores is a logical starting point. Such factors generally feature low in the factor list as there are relatively few samples affected by mineralisation and hence they can only explain a small proportion of the total variance in the data. These factors are also likely to be the least reliable statistically and should not be the only basis on which preliminary target areas are selected.

An additive index is a simple tool for combining elements associated with mineralisation into a single multi-element variable which can be made into a geochemical image to highlight potential target areas. If the data have already been leveled, all elements will be in standardized form (mean of zero, units in standard deviations) and can simply be added together in any combination desired. A well considered combination of elements known to correlate at the "top-end" of the data and empirically observed to be broadly associated with known mineralisation in the area of interest can be a powerful tool for selecting potential target areas. An important feature of additive indices is that more subtle responses which may not stand out in raw data will be significantly enhanced if the geochemical signature matches that of the index. Monitoring the response of an additive index in the vicinity of known mineralisation of interest is an important step in validating the combination of elements being used. A good index will significantly smooth, enhance and enlarge the anomalism associated with mineralisation. If similar responses are noted in areas where mineralisation is not recorded, there are good grounds for the selection of a promising exploration target.

One limitation of an additive index is that not all elements in the index may be contributing to an anomalous index response. Indeed, a single highly anomalous element may generate an index anomaly with little or no contribution from the other elements in the index. The RGB (Red-Green-Blue) anomaly map is a useful complementary visualisation tool which addresses this limitation. Greyscale geochemical images are produced for 3 key elements associated with mineralisation. A threshold is then applied to

each image such that any portion of the image above the threshold is white and the remainder of the image is black. A 95th percentile threshold is a good starting point but different threshold levels can be applied to each image to highlight anomalous patterns as required. The three threshold images are then combined into a single RGB image with the red channel assigned to one element, the green to another and blue to the third. Where the threshold portions of any two elements are coincident, the image area will be colored cyan, yellow or magenta depending on the combination of elements. If all three elements are coincident, the area will be white. For example, if searching for porphyry related Cu-(Au/Mo) mineralisation a Au-Cu-Mo RGB Anomaly Map would highlight the following element combinations; Au (red), Cu (green), Mo (blue), Au-Cu (yellow), Au-Mo (magenta), Cu-Mo (cyan), Au-Cu-Mo (white).

RGB anomaly maps are a powerful visualisation tool for rapidly identifying individual samples or groups of samples with coincident anomalism is any three elements. Element combinations can be selected for specific styles of mineralisation as well as to enhance patterns of zonation. For example a Cu-Mo-Mn combination may be useful for identifying the Cu-Mo core of a porphyry copper system with a peripheral Mn halo. By setting threshold levels lower, it is possible to highlight more subtle responses in key element combinations which may not have been previously recognized using univariate techniques which focus on the highest results.

Patterns of Anomalism in Very Large Stream and Lake Sediment Data Compilations

A single sample from a drainage containing outcropping mineralisation will probably stand out clearly if that sample is visualized with a relatively small number of samples from an otherwise barren local area. As more data are visualized together, covering increasingly larger geographic areas, there is an increasing likelihood of encountering higher and higher results as the diversity of geological and regolith environments encountered contribute more and more opportunities for unusually high concentrations of each element. If the same sample is included in a sub-continental scale data compilation, it will have to "compete" against ten of thousands of other samples to "be seen". Geochemical responses directly associated with mineralisation may literally disappear into the background. On one level, this is a strong disincentive to visualize geochemistry on a very large scale, and hence the importance of large data compilations to allow rapid geologically and geographically controlled selection of data for interpretation and visualisation. On another level, visualizing geochemical data on a large scale provides the exploration geochemist with a unique opportunity to discover the patterns which emerge from those samples which are able to "stand out" amongst the tens of thousands. Three broad types of geochemical patterns are commonly observed in geographically large stream sediment and lake sediment geochemical data compilations;

– Regional scale lithochemical responses which generally map geochemically distinctive lithologies such as granitoids, mafic and ultramafic lithologies, carbonates and some sedimentary units such metalliferous black shales or metamorphic equivalents.
– Highly anomalous individual samples or small clusters of samples with tend to be associated with known mineral occurrences and/or mine workings and the associated contamination. Those with no recorded associated mineralisation of course constitute potential exploration targets. In other instances, the mineral occurrences highlighted by such anomalies may in themselves constitute an exploration target.
– Very large, discrete geochemical anomalies in key ore and pathfinder elements made up from numerous adjacent samples, typically highlighting areas in the order of 25-100km in diameter. These features tend to have remarkably sharp edges and may display broad structural control on a sub-continental scale. Strong geological controls are generally not evident at the surface, with the anomalous areas transgressing local geological boundaries and structures.

It is the last of these examples which is particularly intriguing to the exploration geochemist. They are clearly too large to be associated with a single orebody. They will often, but not always, highlight clusters of mineral occurrences or even entire mining districts with most of the major deposits falling within the anomalous area. The "geochemical signature" of these features ranges from simple (one element) to complex and may or may not broadly reflect the dominant mineralisation styles within the anomaly. In rare instances, broad zonation patterns more typical of individual deposits may be observed, even though the patterns are somewhat "blurred" in stream sediment data and occur on a much larger scale.

The source of such large, discrete geochemical features is not clear. In many instances there is an undeniable contribution from clusters of mineral occurrences and the contamination associated with old

workings, mines and smelters. But it would be a mistake to attribute their source solely to contamination from known mineralisation, as this does not account for the thousands of mineral occurrences clusters, mines and smelters which are not associated with such large geochemical anomalies. Do these very large geochemical anomalies represent the surface expression of deep-seated fertile plutons or perhaps the remnants of giant caldera structures or mantle-tapping structural corridors? Regardless of their origin(s) they are undoubtedly of considerable significance for geochemical targeting. It is postulated that while these features are generally too large to be associated with a single orebody, they are potentially highly indicative of permissive environments for the production of very large mineralised systems, particularly for porphyry related Cu-(Au/Mo) and epithermal Au. For example, approximately 35% of known, significant porphyry-related Cu-(Mo/Au) deposits on the North American continent lie within very large, "continental-scale" Cu anomalies. But these deposits account for more than 60% of the total Cu endowment of all North American porphyry related systems. Clearly the chances for big deposits are considerably better within the very large Cu anomalies, suggesting these features provide an important targeting tool for prioritizing brownfields districts and for identification of new greenfields areas where large anomalies are present outside of established mining districts.

Although very large geochemical anomalies are not common, finding them in amongst the hundreds of thousand of samples available to anybody prepared to make the effort to compile, level and visualize geochemical data on a very large scale, is a powerful incentive for those given the critical task of identifying the best areas to explore.

Acknowledgements

References

Agnew P.D., 1999, Interpretation and Visualisation of Soil and Rockchip Geochemistry, Sepon Project, Laos, in *Exploration Geochemistry for the New Millennium, AIG Bulletin 3*, 113-123.

Davis J.C., 1973, *Statistical and Data Analysis in Geology*, 473-519, John Wiley & Sons.