REMOVING BIAS FROM GEOCHEMICAL INTERPRETATION: THE APPLICATION OF HYPERGEOMETRIC STATISTICS

Ryan R.P. Noble

EIGG/CRC LEME, Curtin University of Technology, Brodie Hall Building, 1 Turner Ave., Technology Park, Bentley, WA, 6102

As geochemical exploration extends into the 21st century mineral deposits are becoming more difficult to locate. New or rehashed geochemical techniques are being used to aid in the discovery of ore bodies beneath cover. These techniques often claim some sort of advantage over the competition, shown in orientation studies, but often it is difficult to assess the quality of the technique. The technique developer's interpretation of geochemical responses is often biased as they already understand where they expect or hope to get positive responses from their chosen technique. The use of hypergeometric statistics, developed by Stanley (2003), is a simple method to remove some of the interpretational bias and provide an empirical assessment of the techniques employed.

Hypergeometric statistics requires an orientation survey using prior knowledge of the underlying geology to predict sites of anomalous response. The following hypergeometric combinatorial formula relates the probability of the anomalous points and false positives to the successful detection of mineralisation (Stanley 2003), while an orientation survey and examples of the information to use the above formula is also presented (Figure 1).

$$P(x) = \frac{\binom{a}{x}\binom{t-a}{k-x}}{\binom{t}{k}} \times 100\%$$

a = number of predicted anomalous points
x = number of correct anomalous points
t = total sample points
k = number of anomalous points (true and false)

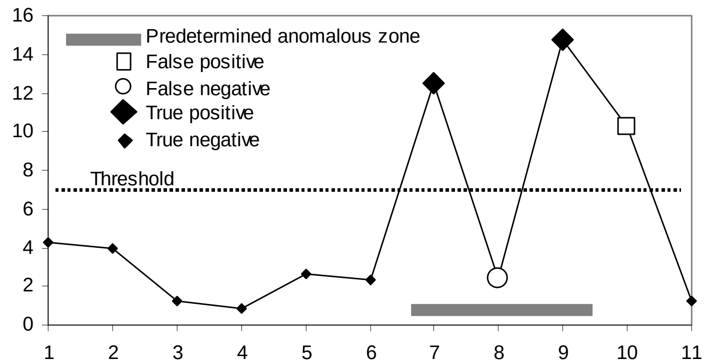


Figure 1: An example traverse with the key pieces of information required to use hypergeometric statistics including the threshold, predetermined anomalous zone, and types of responses.

Hypergeometric statistics allow a rigorous comparison of conventional and new exploration techniques—the lower the probability the more successfully the mineralisation is predicted (Stanley 2003). An assessment of anomaly expression was conducted on various published and hypothetical data using hypergeometric statistics. The aim of this presentation is to explore the use of this statistical method to show quantitative analysis of various traverses across mineralisation.

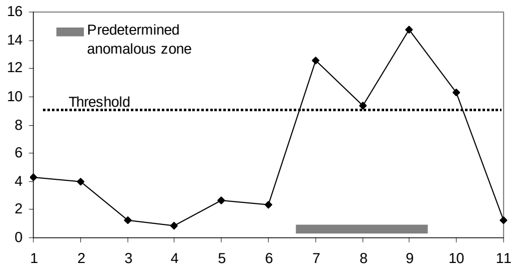
Some studies indicate that certain techniques have been successful or offer an advantage in the identification of mineralisation, but to many readers the interpretation of success is difficult to quantify. Other studies have conducted comparisons of techniques investigating the application of such analyses as sequential versus selective leaches, as well as some new proprietry technologies like Enzyme Leach (Activation Laboratories Ltd.), Mobile Metal Ion (®Wamtech Pty. Ltd) and Bacterial Leach, now called Locatore (®Sologic Pty Ltd). The studies often provide graphical representations of the analysis and the underlying zone of mineralisation, but the results of different techniques are difficult to compare.

In: Roach I.C. ed. 2004. *Regolith 2004.* CRC LEME, pp. 256-259.

Eight traverses (Figures 2-9) are used to provide examples of the use of hypergeometric statistics. The values of each traverse and its success were calculated, with success P(x) < 5% (Table 1). The success rate is a general standard indicating that to obtain this result at random, odds would be less than 1 in 20. The predetermined anomalous zone is assumed to be over a single area of mineralisation. Traverse A was a successfully predicted traverse, even though there was one false positive, although this could theoretically be the influence of a larger than initially thought dispersion halo (Figure 2). The same traverse with a single false negative put into the survey significantly changes the hypergeometric probability to approximately 15%, and even though the traverse visually looks promising, random number generation could replicate this response or better approximately 1 out of 7 times (Table 1, Figure 3).

Table 1: The results of hypergeometric statistics applied to some published and hypothetical orientations traverses

Traverse	Sample points	True positive	False positive	False negative	hypergeometric probability P(x)	Success
A	11	3	1	0	2.4	Yes
B	11	2	1	1	15.2	Unlikely
C	21	2	0	0	1.4	Yes
D	21	2	6	0	13.3	No, threshold or background influence
E	8	1	0	0	12.5	No, inconclusive
F	24	1	0	0	4.2	Yes, influence of sample points
G	24	3	0	8	8.15	No, influence of poorly defined anomalous zone
H	24	3	0	1	0.2	Yes



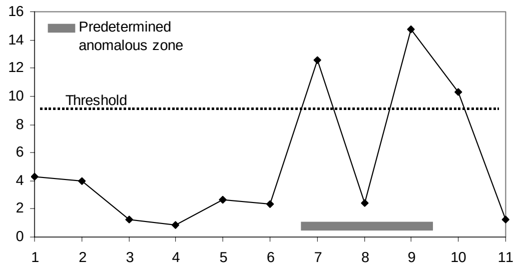
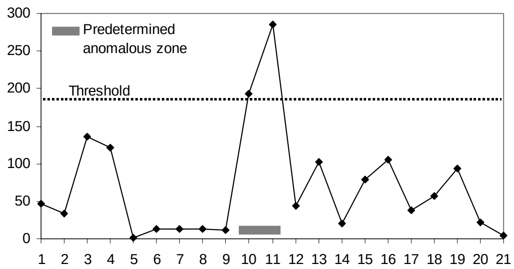


Figure 2 (left) and 3 (right): Traverse A and B (left and right, respectively), with B showing the change in the evaluation due to a false negative over the anomalous zone.

Thresholds are another important consideration and should be set statistically, not by human judgement. A threshold that is too low will probably result in too many false positives, as can be seen in the difference between success and failure for traverses C and D, respectively (Table 1, Figures 4 and 5). A similar response will also occur if the background variation is too large, compared to true positive peaks.



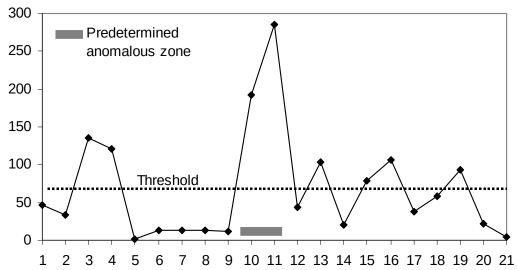
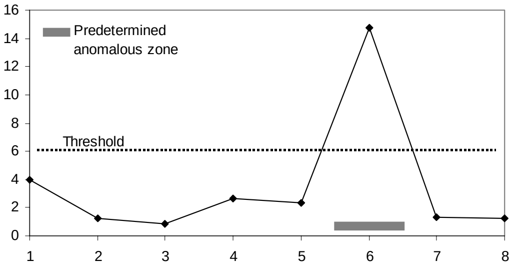


Figure 4 (left) and 5 (right): Traverse C and D (left and right, respectively), with D showing the change in the evaluation due to a lower threshold with more false positive responses.

The number of sample points can also greatly influence the effectiveness of a hypergeometric evaluation. Even with a single point true positive response over mineralisation and no other false positives, the number of samples needs to be at least 21to achieve success. Traverse E response could be generated randomly 1 out of 8 times or 12.5%, and is not a good enough response to warrant drilling (Table 1, Figure 6). The success is greatly improved when more samples are taken from the background (Table 1, Figure 7).



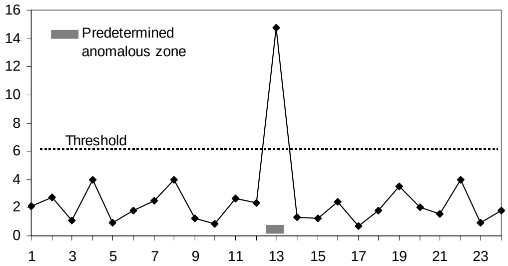
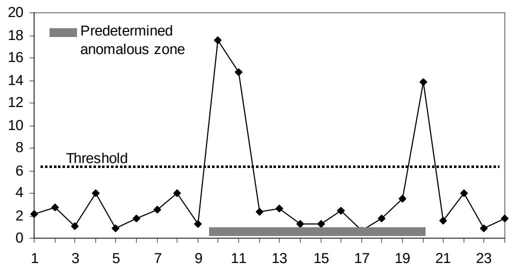
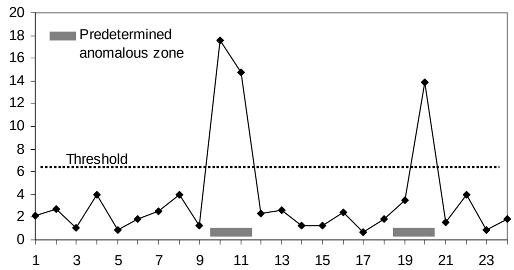


Figure 6 (left) and 7 (right): Traverse E and F (left and right, respectively), with F showing the change in the evaluation due to an increase in sampling points.

Understanding of the dispersion patterns above mineralisation is a critical and ongoing study involving many research organisations. The understanding of the dispersion above the mineralisation in the orientation study can also improve the effectiveness of the hypergeometric evaluation. The 'rabbit ear' anomaly pattern has been described and detailed in numerous studies (Govett 1976, Hamilton 1998). In a theoretical example below, understanding this pattern and limiting the predetermined anomalous zone would greatly improve the success of the evaluation (Table 1, Figures 8 and 9). Traverse G assumes a uniform dispersion above a single mineralised zone, however, if it were known that the 'rabbit ear' dispersion pattern was common in the area, the success of the technique could be improved by limiting the predetermined anomalous zone. Traverse H adjusts the anomalous zone, and even with a false negative, the hypergeometric probability is very low and the technique is successful (Table 1, Figures 8 and 9). Like most tools employed in exploration, the knowledge of the user is critical in obtaining the best results.





Figures 8 (left) and 9 (right): Traverse G and H (left and right, respectively), with H showing the change in the evaluation due to increased understanding by limiting the size of the anomalous zone.

Hypergeometric statstics were applied to a set of analytical techniques over five actual traverses in the Stawell region of Victoria. From the values it is evident that the majority of orientation survey traverses failed to detect the underlying mineralisation (Table 2). The Bacterial Leach (Locatore®) was successful in two of the five traverses, while none of the other techniques predicted the underlying mineralisation. The Bacterial Leach technique quantitatively has some advantage over the others, but whether or not that is a 'successful' method in this region is very debatable, considering it failed on the other three traverses, yielding only a 20% success rate (Table 2). Another important consideration is the depth of cover, which ranged from 25 to 75m in this study site. Perhaps at a shallower site the results of many of the techniques may have been improved.

Another aspect of bias in assessing the analytical method is that often only the successful methods are shown. Of course, some of these proclaimed successes had failures that are not reported, and in turn mislead the reader regarding the evaluation of the technique. It is important when assessing new exploration techniques and methods that the failures be acknowledged, even if they are not a significant part of the paper.

The example results indicate that although quite a few of these traverses looked like 'successes', many were not significant and could be replicated at random. Other bias considerations such as target zone, background levels and dispersal mechanisms are important in the application of hypergeometric statistics, given the user previously defines these parameters. The application of hypergeometric statistics to other data is subject to bias from the interpretation of thresholds and regions of expected anomalous responses. Although this cannot be controlled, by drawing attention to this the user can understand and minimise their bias applications. A blind study is one option that can be used to truly evaluate methods in orientation surveys which is detailed in research by Stanley (2003).

Table 2: A comparison of leaching techniques using a hypergeometric evaluation in the Wildwood prospect, Victoria.

Technique*	Successful analysis P(x) < 5	Number of orientation surveys	Technique success rate
Total Dissolution	0	5	0
Bacterial Leach	2	5	20
0.1M HA	0	5	0
0.25M HA	0	5	0
Ammonium chloride	0	5	0
Ammonium acetate	0	5	0
EC	0	5	0
pH	0	5	0

*Values were from element suites previously used (Noble 2004) or actual value for pH and EC. Individual elements were also analysed with no successful predictions for all techniques.

The results indicate that the examples of data and graphical representation used in orientation surveys can be misleading and are not clearly identifying the mineralisation. Sometimes geochemical exploration is not successful, but the natural bias of the viewer will interpret the response as a success. For this reason it is important to be able to make successful claims with some quantitative analysis, and hypergeometric statistics provide a simple approach to conduct a comparative assessment of various exploration techniques.

Acknowledgements: The support of CRC LEME, MPI mines, Stawell Gold Mine and the Departments of Applied Chemistry and Applied Geology at Curtin University is appreciated.

REFERENCES

GOVETT G.J.S. 1976. Detection of deeply buried and blind sulphide deposits by measurement of H+ and conductivity of closely spaced soil samples. *Journal of Geochemical Exploration* **6**, 359-382.

HAMILTON S.M. 1998. Electrochemical mass-transport in overburden; a new model to account for the formation of selective leach geochemical anomalies in glacial terrain. *Journal of Geochemical Exploration* **63**,155-172.

NOBLE R.R.P. 2004. Bacterial leaching and other techniques to identify gold undercover (N.W. Victoria). *In: AMEC and CRC LEME, Minerals Exploration Seminar* **Abstracts**, Perth, WA, pp. 29-31.

STANLEY C.R. 2003. Statistical evaluation of anomaly recognition performance. *Geochemistry: Exploration, Environment, Analysis* **3**, 3-12.