UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

Technical Letter
Saudi Arabian Mineral
Exploration - 20
Prepared February 27, 1965
Issued August 30, 1965

Dr. Fadil K. Kabbani
Deputy Minister for Mineral Resources
Directorate General for Mineral Resources
Ministry of Petroleum and Mineral Resources
Jiddah, Saudi Arabia

Dear Dr. Kabbani:

Transmitted herewith are 10 copies of:

TECHNICAL LETTER NUMBER 20
THE USE OF ANALYTICAL DATA

by

Paul K. Theobald, Jr.*

Sincerely,

Glen F. Brown

Glen F. Brown, Chief
Saudi Arabian Mineral Exploration Project

* U. S. Geological Survey, Jiddah, Saudi Arabia

UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

Technical Letter
Saudi Arabian Mineral
Exploration - 20
Prepared February 27, 1965
Issued August 30, 1965

THE USE OF ANALYTICAL DATA

by

Paul K. Theobald, Jr.*

Some comment on the nature of analytical data that will come from the trace analysis laboratory, how the data may be used, the nature of samples, and the use of numbers in general is given below to aid geologists in the use of these data. A common and innocent practice exists among geologists, particularly those in mine evaluation, of following certain long established rules that are either not applicable or incorrect. In the interpretation of trace analysis data particularly these "rules" are both not applicable and incorrect.

The accuracy of an analysis of a natural material is not known and cannot be known. The precision of an analysis is established by repeating the determination, by comparing the analysis with those of other analysts, or by comparing results obtained by the use of several analytical techniques. Though this seems a semantic argument, the distinction is necessary. For example, in an interpretation of analytical data for one element, a geochemist felt obliged to reduce all his analytical data by about 30 percent because he could not reproduce the accepted value for that element in one of the international standard samples. Within months of his own publication, a new accepted value was published that agreed precisely with the original result obtained by this geochemist. The distinction is basic to interpretation: precision, the correct usage, is the variation expected about an accepted value; accuracy, the unknown quantity, is the single, correct value.

Reports of analysis from an analytical laboratory almost always give one figure that is not truely significant. For example, 64.79 percent silica reported

* U. S. Geological Survey, Jiddah, Saudi Arabia

in a standard rock analysis should be read "nearly 64.8 percent." This distinction becomes particularly critical for semi-quantitative analysis, where a single figure may be reported and may not be truly significant. As an example, 70 ppm (parts per million) usually means "the most likely value lies between 50 and 100 ppm;" and 15 ppm usually means "the most likely value lies between 10 and 20 ppm." In trace analysis the analyst usually controls the interval to be reported by the expected precision. In most instances the step between one number and the next is approximately the standard deviation of precision. Unfortunately this interval is not uniform in the chemical analyses, but it varies from element to element and from batch to batch of samples for a given element. In general, when adequate numbers of samples from a given area have been analysed, the geologist can get a fair idea of the report interval. It is then best to group the data combining three, or where data are abundant two, report intervals for interpretation.

The precision of sampling is invariably poor in comparison to the analytical precision. Taking a sample lifts the geologist above the level of speculation about an area and places him in a "game of chance." Generally, the more samples involved, the better the interpretation, not because this improves the precision of any sample, but because it gives improved control over the magnitude of variation.

No amount of "numerical adjustment" will improve the precision of the samples, the analyses, or the interpretation. Statistical analysis is a powerful tool for determining the limits of precision; that is, to establish the reliability of an interpretation. An example from an available report will illustrate what is meant:

From 33 samples one is dropped as not representative on a basis of the material sampled, an acceptable and necessary procedure even though subjective. The remaining 32 samples range in metal content from 1.45 to 13.36 units and average 7.42 units. In the words of the author of the report "a more accurate weighted average of 7.41" units was obtained by weighting each sample according to

the length sampled. This procedure is incorrect because: 1) Accuracy is not involved, the accurate value could be higher or lower by a factor of 2 or even 10 and there is no way of knowing; 2) there is no real difference between 7.41 and 7.42 so neighter is better nor worse than the other; 3) though the data were available, no estimate of the precision of the conclusion was made. The first two of these objections should be self explanatory. The third objection leads to the following conclusions: The average of 7.42 has a Standard Deviation of 3.3 units, that is, about 60 percent of the values are between 4.5 and 10.5 units; The Confidence Interval about the average leads to the statement "There is one chance in ten that the average is not between 6.5 and 8.5." As stated, the conclusion of this author is incorrect. The average metal content of these samples is 7.5 $\pm$ 1 unit.

If the precision of analyses in the above example is excellent it will be $\pm$ 0.05 unit; if it is poor, $\pm$ 0.5 unit; if it is terrible, $\pm$ 1 unit. That is, if the precision of the analyses quoted above is terrible, it will not effect the result! If, as geologists, we wish to object to the quality of a chemist's work, we had better establish first that our samples justify the quality of analysis we are receiving.

The results of many of the trace analyses provided by the laboratory will be $\pm$ 50 percent of the reported number, and these will be as good or better than the samples. The techniques are designed to produce maximum information at minimum cost rather than maximum precision at any cost. Hopefully, by reducing analytical cost, we can overcome the inadequacy of our samples by increasing their numbers. Please use the data wisely.