Improved Evaluation of Environmental Radiochemical Inorganic Solid Matrix Replicate Precision:

Normalized Range Analysis Revisited

Robert E. Gladd
James W. Dillard, Ph.D.

I.T. Corporation
OAK RIDGE LABORATORY

ABSTRACT: The Normalized Range Statistic, as defined in EPA-600/4-81, provides the radiological laboratory analyst with a generally robust index of analytical precision in unknown replicate analyses. Inorganic solids, however, present a statistical assessement problem; replicate analyses of low level environmental samples with these necessarily small (usually <20 g.) aliquots frequently return spurious QC "outliers" owing to the relative magitudes of the calculated and "expected" sigmas, the "expected" value being based on a 1 kg. aliquot and the factor employed by the traditional Range Analysis method. This paper examines the mathematical nature of the statistical deficiency and proposes alternative solutions for improvement in the evaluation of environmental radiochemical inorganic solid matrix replicate precision.

The monitoring of laboratory accuracy and precision at the IT Oak Ridge Laboratory is guided by the statistical procedures detailed in EPA-600/4-81-004, methods which provide generally practical empirical point estimators of analytical performance. Accuracy evaluation is accomplished through the use of the "Normalized Deviation" statistic, in which the analytical result of a spiked sample test is "normalized" to the "known" value and "expected laboratory 1-sigma" precision; in traditional statistical parlance, a "Z-transformation." Normalized Deviation statistics (NDEVs) are computed and plotted on control charts with a mean of zero, warning limits at +/-2.0 and control limits of +/- 3.0.

Similarly, the analytical precision of unknown replicates (i.e., samples where a "known" value from a reference standard is not present) is assessed via the Normalized Range (NRANGE) estimator, in which the numerical difference between replicate results is evaluated in the context of both an "expected 1-sigma" precision level and an "expected range" factor. NRANGE statistics are computed and plotted on control charts containing an X-Y origin of zero, an "expected range" of 1.0, and warning and control limits at +3.0 and +4.0 respectively (i.e., mean, or "expected" range plus 2 and 3 sigma). NRANGE points lying above the +4.0 Control Limit mandate an investigation of the analytical data for the replicates in question to ascertain the causes of the excessive divergence.

These statistical tools assume the presence of a liter or kilogram sample aliquot, the latter necessitated by the EPA food matrix crosscheck. Analytical results are adjusted to their respective activities at a liter or kilogram before the statistics are computed. Where the samples are constituted of low-level, low-volume inorganic solid matrices, the generation of spurious "outlier" statistics is a recurring phenomenon, owing principally to the relative magnitudes of the analytically determined 1-sigma and the "expected 1-sigma" used by the NRANGE computation. This a-priori 1-sigma, while empirically appropriate for 1 kg. samples, imposes an unrealistic constraint on small aliquot inorganic replicates. Clearly, in such cases a method of incorporating the analytically determined 1-sigmas must be employed to make the NRANGE statistic reflect the true precision level of the replicates; to the extent that these "outliers" are invalid they con-

tribute to a misleading impression of laboratory precision capabilities and result in unwarranted technical review of the replicate sample data, an examination required of all "out-of-control" QC results.

Since every quantitative sample result is essentially a point-estimate of a "mean" value which approximates a "true" activity or concentration level, it would be tempting to dismiss any "expected sigma" constraints on replicate results, particularly of the types under discussion here, and apply a sort of "t-test" on our experimental "means" incorporating only the analytical 1-sigmas in determining the acceptability of the range between replicate values. Support for such a method would derive from the fact that the relative standard deviation (i.e., the "percent sigma," or coefficient of variation) accompanying each production analysis is not routinely evaluated against an "expected" sigma; it is generally accepted that, as activities and/or aliquots are lesser sigmas will tend to be proportionally greater, frequently approaching or even exceeding the magnitudes of the quantitated activities themselves. An "expected sigma," while serviceable in the main as an objective standard of analytical variability, is frequently inappropriate in light of the component measurement particulars of individual cases such as those under review in this presentation.

$$Z_{alt1} = \frac{|R_1 - R_2|}{\frac{\sigma_1+\sigma_2}{2}} \quad [1]$$

$$Z_{alt2} = \frac{|R_1 - R_2|}{\sqrt{\sigma_1^2+\sigma_2^2}} \quad [2]$$

$$NR_{adj} = NR_{epa} \times \frac{\sigma_{epa}}{\frac{\sigma_1+\sigma_2}{2}} \quad [3]$$

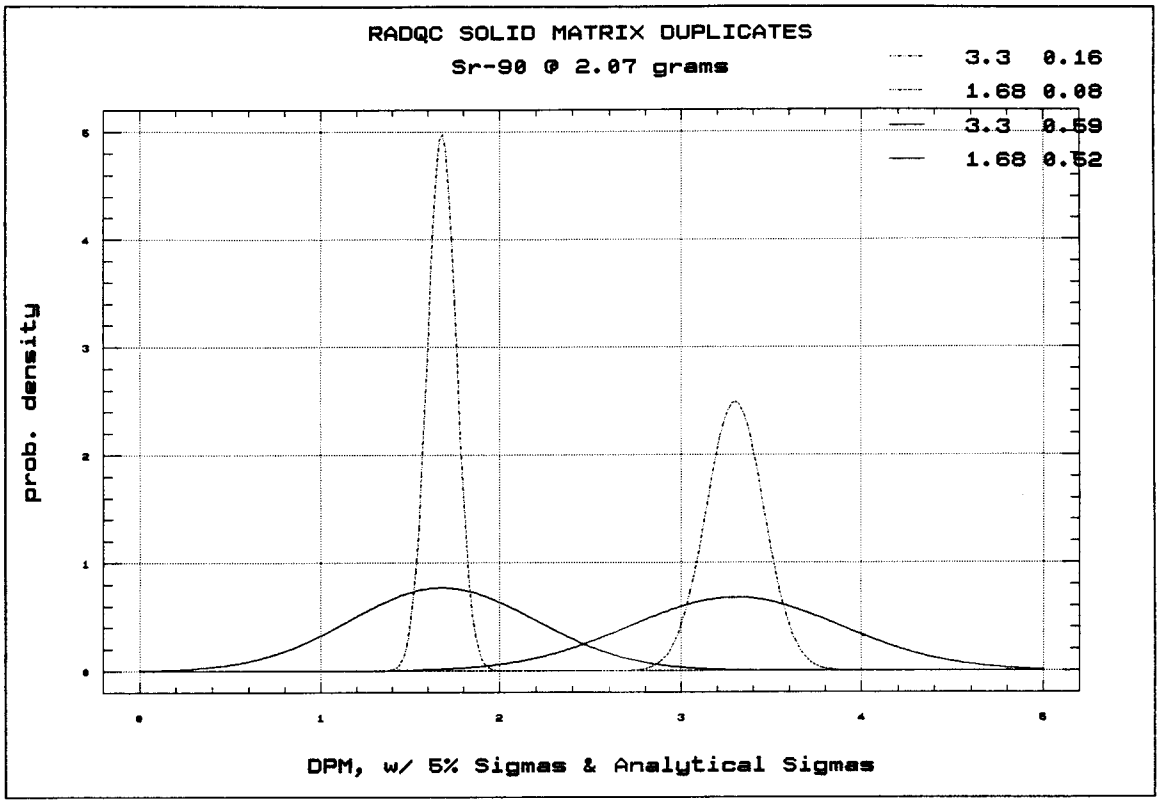
The conventional t-test cannot be directly applied to replicate radioanalytical results owing to the fact that "N=1" for each sample dataset, leaving us without "degrees of freedom" to employ in the derivation of t-values. Given this problem, should we wish to retain the mathematical simplicity of the NRANGE statistic, we could simply replace the a-priori "expected" sigma in the formula with the mean of the analytical percent sigmas. Such a replacement would yield NRANGE statistics derived totally in the context of the error terms of the lab results themselves, removing any empirically "objective" variability standard in favor of the case-specific uncertainty estimates. The virtue of such an approach would be to remove any potential argument over whether an "NRANGE>4" calculated in such a fashion in fact represented an "out-of-control replicate set. Replicate results so divergent as to normalize out to "NR > 4" even after taking their own individual error terms into account would indisputably be indicative of unacceptable precision and would indeed merit technical review to ascertain the causes of the disparity.

Alternatively, eschewing the NRANGE formula entirely, we might statistically examine our replicates via one of two variations on a "Z-test" formulation, employing either the mean of the analytical sigmas or the square root of the sum of the variances as divisors of the replicate range, as shown in the box below.

Under either "z-score" approach (fig.1 & 2 in box) our "outliers" would be those resulting in a z-statistic > 3.0 absolute. This approach makes intuitive sense, but is a bit bothersome in that any utilization of a reference value such as the "expected sigma" is again precluded, a tactic that contravenes an implicit assessment principle of the EPA-600 method: the application of empirical guideposts to laboratory precision capability accounting. A simple adjustment to the NRANGE statistic is therefore proposed, one that incorporates both the a-priori EPA sigma factor ***and*** the mean of the analytical sigmas into the NRANGE calculation, as shown by fig. 3 above: a ratio of "expected" over "found."

It should be readily apparent that where the mean lab sigma nearly equals the "expected" sigma the NRANGE statistic will be quite close to the value returned by the standard method. Where the lab error coefficient is greater than the expected, the NRANGE will be attenuated by the ratio of the two. Further, where the mean lab sigma coefficient is ***smaller*** than the expected, the adjustment factor wil be > 1.0, thereby ***expanding*** the NRANGE value. In this manner the sigma-ratio adjustment factor is a double-edged sword; if individual error terms are better than the expected, the results had better be minimally divergent to avoid being pushed into "out-of-control" status. Such a condition makes methodological sense; analytical sigmas are mathematical expressions of our confidence in our quantitative estimates. Replicates returning better-than-expected error terms ***and*** grossly disparate "means" are indicative of a condition warranting quality control review.

A graphical example effectively illustrates the problem posed by the application of an expected sigma to a low-level small aliquot solid matrix duplicate result set. Two Sr-90 results, at 3.30 and 1.68 DPM respectively, are graphed as normal distributions (see box, below right), first with an assumption of a 5% CV, then overlaid using sigmas derived in the analyses (0.59 and 0.52 1-sigmas, respectively). While the results are displayed as narrow, peaked distributions whose tails are quite far apart under a 5% CV assumption, when viewed in the distributional context of the analytically derived sigmas, quite another picture emerges; the tails of the distributions overlap substantially. The traditional NRANGE statistic for this set came in at NR=14.73, while the "corrected" NR=3.01, and this adjusted Normalized Range value seems appropriate; our replicates diverge, perhaps more than we would prefer, but certainly not to the extent indicated by a Normalized Range of 14.73. Were these replicate results those of a full kilogram vegetation matrix emanating hundreds of DPM, we would perhaps have cause for concern at the disparate replicate values returned by the lab. In the instance of 2.07 gram inorganic aliquots evincing a few DPM, however, the range between R1 and R2 is not all that severe, certainly not to the point implied by an NRANGE statistic of 14.73. The relative standard deviations (the CVs) for R1 and R2 were, respectively, 17.86% and 30.91%.



EPA-600 expected 1-sigma % precision guidelines are grouped by type of analysis into four levels: 5%, 10%, 15%, and 25%. An examination of 116 inorganic solid matrix replicate results from our RADQC™ laboratory QC database is revealing. As might be expected, most of the NRANGE difficulty lies with the analyses classified in the "5% precision" group, as the following table illustrates:

N	EPA 1-sigma	LAB 1-sigma
68	0.05	0.125
4	0.10	0.121
38	0.15	0.159
6	0.25	0.232

The column on the right tabulates the average CVs found in actual practice, grouped by the EPA sigma classifications. It is evident that a 5% sigma represents an unrealistic level of precision where small aliquot solids are concerned. The application of the NRANGE sigma ratio adjustment factor factor to the 116 samples investigated for this research effort reduced the "NR>4.0" outliers by 75%, and the attenuated "adjusted NR" statistics were overwhelming of a magnitude consistent with the type of graphical evidence obtained by plotting the gaussian distributions in the manner of the above example.

Further statistical support for the use of the NR adjustment is seen by a correlation matrix comprised of the values obtained for this QC data under the formulas displayed in figs. 1, 2, and 3:

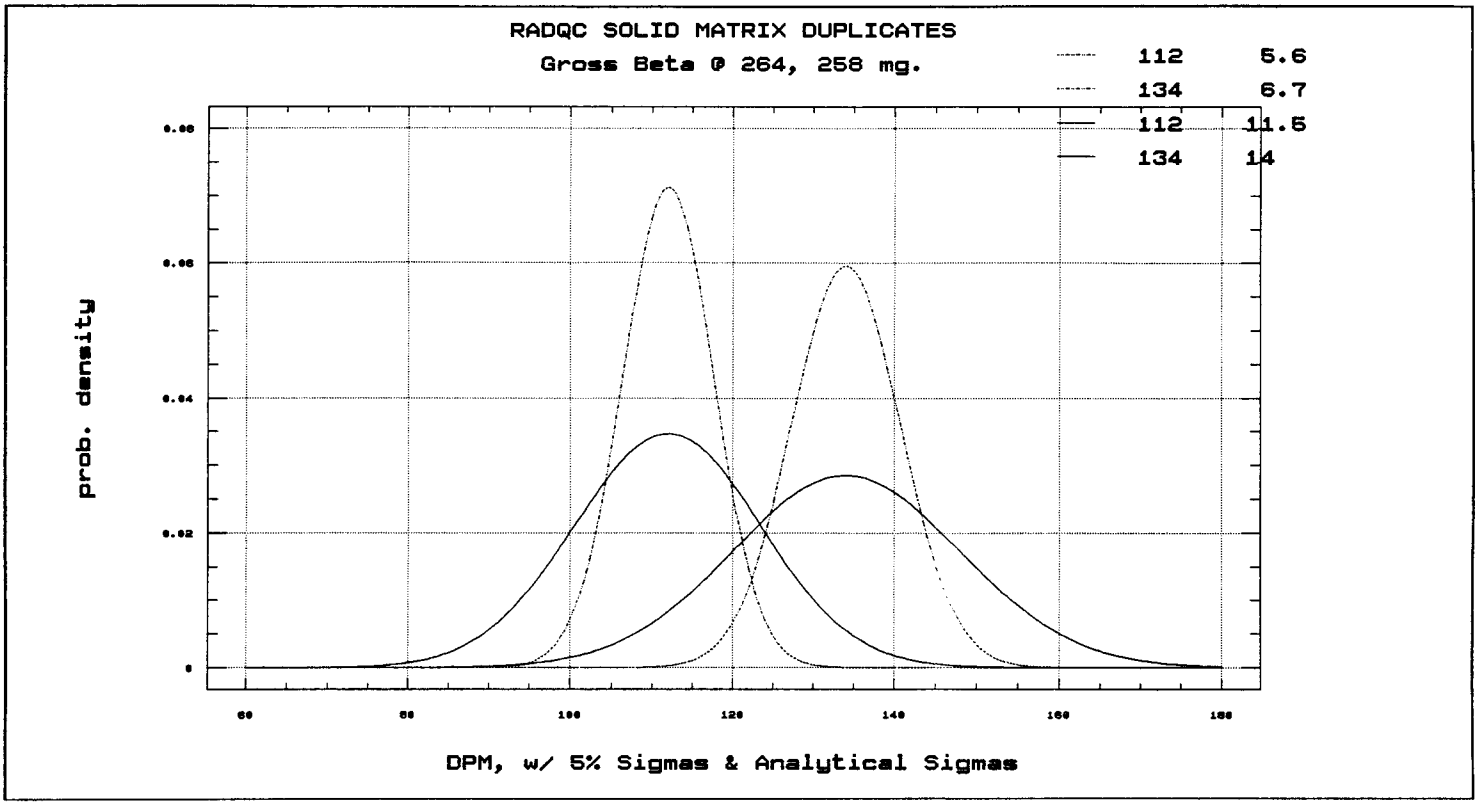
	NR_adj	Z_alt1	Z_alt2
NR_adj	1.000	0.903	0.925
Z_alt1	0.903	1.000	0.956
Z_alt2	0.925	0.956	1.000

The high Pearson-R correlations among the three methods indicate a significant agreement between the two Z-score variations and the adjusted NRANGE method; we are measuring the same phenomenon, irrespective of algebraic method. Any of these formulations serve to reduce the quantity of spurious QC outliers, improving the assessment of inorganic solid matrix precision.

A second graphic plot example is provided below to further demonstrate the utility of the NRANGE adjustment method. This replicate set consisted of Gross Beta analyses performed on aliquot volumes of 264 and 258 mg. The DPM results were calculated to be 112 +/- 23 and 134 +/- 28 (2-sigmas). The unadjusted NR=4.38, just slightly over into the outlier realm. The "corrected" NR=2.11, and again, the distribution plots seem to provide visual agreement with the numerical statistic. The mean CV was 10.36%.

ROBERT E. GLADD is Vice President of CAM3 Associates of Knoxville, TN. He has served since 1986 as a statistical analyst and computer applications consultant to the IT Oak Ridge Laboratory.

JAMES W. DILLARD, Ph.D. is the Technical Director of the Oak Ridge Radioanalytical Laboratory of IT Corporation.



A quick look at some of the aggregate univariate and correlation statistics for the dataset of solid matrix replicates used in this effort is useful. The median uncorrected NRANGE was 2.14, with a mean of 3.74, while the median adjusted NRANGE was 1.48, with a mean of 2.02. The smallest aliquot was 4.8 mg., and the largest was 841.9 grams, with a median of 2.07 grams and mean of approximately 69 grams. The lab sigma precision was, as we might expect, inversely correlated with aliquot volume (R= -.25, statistically "significant" with p=.007) The median lab sigma precision was 12.17% with a mean of 14.04%. One lesson flowing from these data is that perhaps those types of QC analyses currently classified as requiring "5% expected 1-sigma precision" be instead calculated using a 15% expected precision statistic where the samples are those of inorganic solid matrices. The NRANGE sigma ratio adjustment factor employed here is essentially performing roughly that very sort of task in attenuating the Normalized Range where the sigmas are closer to 15% in the laboratory production environment.

REFERENCES

Kanipe, Larry G. (1977), *Handbook for Analytical Quality Control in Radioanalytical Laboratories,* TVA/USEPA Interagency Energy-Environment Research & Development Program (EPA 600/7-77-088)

Walpole, Ronald E. & Myers, Raymond H. (1989), *Probability and Statistics for Engineers and Scientists, Fourth Edition,* New York, MacMillan

EPA 600/4-81-004, USEPA, EMSL Office of Research & Development, Las Vegas, NV.