SCALING RATIOS PRODUCE MISLEADING RESULTS IN D&D PROJECTS

K. S. Redus
Redus and Associates

ABSTRACT

Scaling ratios are frequently used in operational nuclear environments to estimate the concentration of one radionuclide using information from an “indicator” radionuclide. The general form of a scaling equation is Y = R x X, where Y is the radionuclide being estimated, X is the indicator radionuclide, and R is the scaling ratio that describes the relationship between X and Y. Often, the measure of interest is the concentration of the radionuclides. When scaling ratios are used in Decontamination and Decommissioning (D&D) projects (and, many times, Remedial Action (RA) projects), misleading and potentially costly results can occur. The reason for this is not a violation of the assumption that some relationship between two radionuclides exists. Rather, it is twofold. First, a D&D or RA situation is not under control like the environment at a operational nuclear facility. A typical performance measure for control is variation about an average. Averages in operational environments exhibit variation of 10% or less of the average value, but the variation in ER environments commonly exceeds 100% of the average. The second reason for concern is that scaling ratios in ER environments are usually calculated by the simple ratio of the arithmetic average of Y to the arithmetic average of X without regard to the underlying linear or non-linear relationship of the data or the probability distributions of X and Y. We have developed a new approach that should replace the current scaling ratio approach. Our new approach – P-Scale® – tackles variation and uses simple statistics to deal with characterization problems in D&D and RA environments. It is easily coupled with the Multi-Agency Radiation Survey and Site Investigation Manual and the Data Quality Objective Process to decrease number of samples, Finally, P-Scale® may be used to provide increase accuracy in transportation documentation and individual shipment manifests.

INTRODUCTION

Estimation of the statistical properties (e.g., the average or the range) of one radionuclide based on measures and known physical relationships with another radionuclide is a typical approach used in operational nuclear environments. This is called “scaling.” A common property of interest is the concentration of the radionuclides. In an operational nuclear facility, the physical relationships between radionuclides and the processes are well known and controlled. It is reasonable to use indicator radionuclides such as Cs-137 or Sr-90 to estimate concentrations of other, difficult to measure, radionuclides. This approach works because an operational environment is in “control.” Control means that variation of an indicator radionuclide is small, i.e., $\pm$ 10% of its nominal value. The scaling method is simple and easy to accomplish, cost is minimized, and safe operations at the facility are maintained.

When the scaling method is applied during Decontamination and Decommissioning (D&D) projects and Remedial Action (RA) projects, misleading and potentially costly results can – and do – occur. In a D&D or RA environment, the interactions of the radionuclides are not well understood and the variation is exceeds 100%.

We have developed a robust approach called P-Scale® that should replace the naïve scaling method approach to deal with the high variation present in D&D or RA projects and the misleading calculation approach that is currently used.

Overview of Scaling Approaches

The Usual Approach

The usual scaling approach employs a scaling ratio, R, that relates Y (the radionuclide that needs to be estimated) to X (the indicator radionuclide used to estimate Y). The usual approach is quite simple and easy to implement. When all assumptions are met, it is a cost-effective way to collect information and estimate new values of Y.

There are three assumptions to the usual approach: (1) an indicator variable, X can estimate another variable, Y; (2) the process is in control, namely, the variability in X or Y is around 10% at 1-sigma; and, (3) a linear relationship exists between X and Y that can be described as Y = R x X. Historical data serves as the input to calculate R, the “scaling ratio.” R is calculated as the average of each of the Y/X values and is called “the average scaling ratio.” Once R has been determined, new X data is obtained. Any transformations to desired units are performed on the X data. The assumptions are verified. Finally, the new values of X are used to calculate Y using R, the scaling ratio determined from the historical data.

Consider the hypothetical data as illustrated in Figure 1. Visual examination of the data indicates some linear relationship exists between X and Y. For this data, the average of X is 65.7, the standard deviation of is 9.3, and the relative variation of X (the standard deviation divided by the average and expressed as a percentage) is around 14%. The average of Y is 177.1, the standard deviation of is 80.1, and the relative variation of Y is 45%. Since the relative variation is small, we would conclude the process is in reasonable control. Using the hypothetical data, R is 2.6.

As can be seen in Figure 1, the average Pu-240/Cs-137 ratio of 2.6 yields a simple linear equation as Pu-240 = 2.6 x Cs-137 which both overestimates or underestimates Pu-240 values for input values of Cs-137. If new data is obtained, say, Cs-137 = 75 pCi/g, then Pu-240 = 195 pCi/g. If the maximum Pu-240/Cs-137 ratio is used, the value of R is 4.5. All Pu-240 values are overestimated further demonstrating the usual approach is misleading and unrealistic.

The Regression Approach

Statistical regression, commonly called “curve fitting,” is technique that minimizes the sum of squared distances between the observed Y and the estimated Y for all X, and this technique is known as least squares linear regression. In scaling jargon, R should not be a non-linear function of X and Y. A regression model helps to understand how well R captures the linear relationship between X and Y. Two common measures are (a) the fit of the model and (b) the coefficient of determination. The former measure is expressed as a probability that at least one of the coefficients in the regression model is zero. For scaling method linear regression models of the form Y = a + R x X (where a is the intercept and R is the scaling ratio), the probability should be less than 5%. This would then indicate there is a 95% “confidence” that either the intercept or the scaling ratio is not zero. The coefficient of determination, denoted as r^2, is the proportion of the total sample variability around the average of Y that can be explained by the linear relationship between Y and X. This value is bounded by 0 and 1, and, in general, the closer r^2 is to one, the more variability around the average of Y is explained by the linear relationship between Y and X. Considerably more detail on linear regression can be found in any elementary statistics text.

Using the hypothetical data, the linear regression of Y and X can be determined. This is presented in Figure 2, and it looks like the form Y = a + R x X. The solid line represents the linear regression fit of X and Y. The dotted lines represent the 95% confidence interval for the regression line as defined for the range of X values. The horizontal line is the average of Y. Using the hypothetical data, the scaling model can be written as Y = -355.8 + 8.1 x X.

The plot and the scaling model validate what was observed by visual inspection – there is a linear relationship between Y and X. The model fit is statistically significant, namely, at least one of the coefficients of the regression model is non-zero. The r^2 value is 0.89 which is a reasonable value to explain the variability around the average of Y as explained by the linear relationship between Y and X. By examining the leverage[a] of the regression, we see that $P < 0.0001$, and this indicates the significance that the scaling ratio is statistically significant different than zero.

The most immediate conclusion to be drawn from Figure 2 is that Y = -355.8 + 8.1 x X (as determined by the regression approach) looks quite different than Y = 2.6 x X (as determined by the usual approach). The regression approach seems to have more adequately captured the variability of the data than the usual approach. The next conclusion to be drawn is that even when process is in control, the usual approach consistently overestimates or underestimates the values of Y. There is no ability to create confidence bounds under the usual approach. Regression can provide confidence bounds on regression line (Pu-240 = -355.8 + 8.1 x Cs-137) and the results are defensible and repeatable. Regression is a viable alternative, but it requires understanding of statistical tool and its limitations. Regression can only be used for forecasts of Pu-240 when new Cs-137 values are defined in same domain as original Cs-137 values that were used for the regression model.

A Case Study

Sampling and analysis was performed on the Gunite and Associated Tanks (GAAT) at the Oak Ridge National Laboratory (1, 2) to support a RA project that emptied the tanks. When the project was completed, contaminated equipment used for this purpose was placed in twenty-two B-25 boxes at a temporary storage location prior to disposal at a suitable location. While a considerable amount of data existed to describe the tank contents, there was no information available that characterized the contaminated equipment.

An effort was undertaken to determine how contaminated the equipment actually was. Data from 42 samples collected during the tank sampling and analysis campaign was used. Scaling ratios were calculated with Cs-137 as the indicator variable, X, for numerous other radionuclides (Am-241, Sr-90, Pu-239, Pu-240, etc.) that were found in the tank contents. Non-Destructive Assay (NDA) was then employed on each of the B-25 boxes to measure dose. When dose measurements were completed, Cs-137 activity was modeled using MicroShield™ software. After suitable correction for the mass and volume of the contaminated equipment, concentrations of the other radionuclides in the contaminated equipment were performed using the previously determined scaling ratios. The ultimate goal was to dispose the contaminated equipment. The estimated concentrations of the other radionuclides were compared to the waste acceptance criteria for the proposed disposition location, and the contaminated equipment was successfully disposed as planned.

To demonstrate the usual approach and the regression approach, we look at how the scaling ratio for Pu-240 is determined when Cs-137 is the indicator radionuclide. As can be seen in Figure 3, the relationship between Cs-137 and Pu-240 looks random – and certainly not linear. The average Cs-137 concentration is 7.88E+06 pCi/g, the standard deviation is 9.27e+06 pCi/g, and the relative variation is 118%. The average Pu-240 concentration is 1.38E+03 pCi/g, the standard deviation is 1.87E+03, and the relative variation is 135%.

A quadratic model is fit to the ln of the concentrations which results in:

$$\ln \text{Pu-240} = 3.064 + 0.323 * \ln \text{Cs-137 [pCi/g]} - 0.22 * (\ln \text{Cs-137} - 13.09)^2 \qquad \text{Eqn. 2}$$

The leverage is significant ($P < 0.0005$). We conclude a quadratic model is a reasonable fit of the GAAT data.

The most immediate conclusion is when we compare the usual approach model of ln Pu-240 = 0.448 x ln Cs-137 to the regression model. We see that usual approach does a poor job in describing the relationship between Cs-137 and Pu-240 for the GAAT tank data. This data describes a process that is not in control with a relative variation almost 275% at 1-sigma further substantiating the process is not in control. Finally, we see that there is really not a linear relationship between Cs-137 and Pu-240 that can be described as Y = R x X. All assumptions associated with the usual approach have been violated.

Observations

There are two key problems that need to be addressed: The first is that D&D projects are not in control. Most D&D projects exhibit variation exceeding 100% at one-sigma. In an operational environment, this could never be tolerated. We need to use variability rather than wish it away. The second problem is that as long as regression assumptions hold, then regression is a viable alternative. The results are repeatable and confidence bounds can be created. As long as newly obtained Cs-137 data falls within bounds of regression model, then estimates of Pu-240 make sense.

Violation of assumptions almost always occurs in D&D projects. The variation of Y and X is too large (much greater than 100%) to develop meaningful regression models. Y may be dependent on variables other than X (e.g., $Y = b_0 + b_1X + b_2W + b_3Z + b_4WX$).

P-Scale® - The New Approach

In this section, we describe P-Scale®, an innovative approach that is remarkably simple and reduces the deficiencies of each of the aforementioned approaches.

Recall the notation that R is the scaling ratio, X is the indicator radionuclide, and Y is the estimated radionuclide. We demonstrate our new approach using the GAAT data for Pu-240 as Y and Cs-137 as X. P-Scale® applies the following logical steps which are illustrated in Figure 4.

1. Use historical data to calculate the simple scaling ratio, R = Y/X, for each X and the corresponding Y
2. Determine F(R), the probability distribution for R, and associated summary statistics
3. Obtain supplemental data for X and calculate F(X), the probability distribution for the indicator radionuclide X
4. Convolve F(R) and F(X) to determine F(Y), the probability distribution for Y; and, determine the expected value Y and the upper 95th percent confidence bound for expected value of Y

Step 1. The simple scaling ratio, R, is calculated by dividing each Y (Pu-240) concentration by each X (Cs-137) concentration. The P-Scale® result for the GAAT data results in 42 values of R ranging from 3.66E-06 to 0.063.

Step 2. We determine F(R), the probability distribution of R, using a statistical goodness of fit test. We conclude it follows a Weibull distribution with scale parameter of 0.002 and shape parameter of 0.549. The Cramer-von Mises W test statistic is 0.04591. The P-value is greater than 0.25 which is sufficient to conclude the data fits the distribution at 95% confidence.

The P-Scale® result for the expected value of the scaling ratio under this Weibull distribution is 0.003. We now have an understanding of the variability of the scaling ratio, specifically, we can say with 95% confidence the scaling ratio will not exceed 0.014.

Step 3. Supplemental data for X, the indicator radionuclide (Cs-137), is used to determine the probability distribution F(X) when using P-Scale®. The data was obtained by NDA measurements from twenty-two B-25 boxes and modeled using MicroShield™. Dose measurements were converted to concentrations using the volume and mass of the containers.

We determine F(X) of the Cs-137 concentration using a statistical goodness of fit test. We conclude it follows a Weibull distribution with scale parameter of 1.11E+04 and shape parameter of 0.43. The Cramer-von Mises W test statistic is 0.11123. The P-value is 0.0677 which is sufficient to conclude the data fits the distribution at 95% confidence. The arithmetic average of the C-137 in the B-25 boxes is 2.23E+04 pCi/g. The standard deviation is 2.81E+04 pCi/g and the relative variation is 126%.

Step 4. We now convolve f(R) and f(X) to generate the probability distribution, F(Y), for Y when applying the P-Scale® approach. F(Y) is the basis for determining the expected value Y and the upper 95th percent confidence limit for the expected value of Y. The variability of R and X are explicitly modeling to determine Y. The following results are summarized:

ID	Parameter	Units	Probability Distribution	Expected Value	Upper 95th percent confidence limit
R	Scaling Ratio	None	Weibull (0.549, 0.0019)	0.003	0.014
X	Cs-137 in B-25 Boxes	pCi/g	Weibull (0.43, .11E+04)	3.07E+04	1.43E+05

Convolution of F(R) and F(Y) is accomplished using by P-Scale® mathematical simulation. One hundred simulations of 100 samples from each of R and X were used to calculate the convolution, F(R) * F(X), where R and X are defined by their unique Weibull distributions. The mathematical simulation is equivalent to (a) assaying 100 tanks samples to calculate R and (b) taking 100 NDA measurements from the B-25 boxes to estimate Cs-137 concentration.

Consider what the Pu-240 estimates in the contaminated equipment would be if the usual scaling approach were applied instead of the P-Scale® approach. The arithmetic average scaling ratio value is 0.003, and the arithmetic average of Cs-137 concentration in the B-25 boxes is 2.23E+04 pCi/g. Using Y = R x X results in a Pu-240 concentration in the contaminated equipment of 0.003 x 2.23E+04 = 6.69E+01 pCi/g. Often times, a D&D project will say “we will the maximum values as a conservative measure to perform scaling and subsequent Y estimation” regardless of the fact the data may be highly skewed to the right and exhibit long tails. The maximum scaling ratio is 0.063 and the maximum Cs-137 concentration obtained by NDA and MicroShield™ is 9.85E+04 pCi/g. If these maximum – or conservative – values are used, then the Pu-240 concentration in the contaminated equipment of 0.063 x 9.85E+04 = 6.21E+03 pCi/g. While this value is certainly conservative, we will see it is highly improbable.

P-Scale® examines F(Y) to determine the expected value of Y and the upper 95th percent confidence limit for the expected value of Y. Examine Figure 5. Application of the P-Scale® approach results in the following observations:

- The expected Pu-240 concentration for the contaminated equipment in the B-25 boxes is 9.06E+01 pCi/g and the standard deviation is 3.54E+01 pCi/g. The relative variation is 39%.

- The Pu-240 concentration for the contaminated equipment in the B-25 boxes follows a Gamma distribution with parameters alpha = 8.29 and beta = 10.93. The Cramer-von Mises W test statistic is 0.250948 and the P-value is 0.1941 indicating at least 95% confidence the fit is reasonable.

- The upper 95th percent confidence limit for the Pu-240 concentration for the contaminated equipment in the B-25 boxes is 1.48E+02 pCi/g.

- A “conservative” estimate for the Pu-240 concentration for the contaminated equipment in the B-25 boxes of 6.21E+01 pCi/g will occur with probability less than 0.0001. Thus, the degree of conservatism is 99.99%.

CONCLUSION

We have developed and demonstrated P-Scale®, a robust mathematical simulation approach that enhances the naïve characterization calculation and use of scaling ratios in D&D projects. Our new approach is straightforward and simple to apply. The benefits of P-Scale® are:

- P-Scale® explicitly accounts for variability of indicator variable, X, and the variable to be estimated, Y, to determine the scaling ratio, R

- P-Scale® allows for the propagation of the variability from supplemental X data to estimate Y when using R

- P-Scale® quantifies "conservative" estimates with a precise probability of occurrence

- P-Scale® determines robust upper confidence bounds for calculated Y given supplemental X data

P-Scale® should be used instead of the usual scaling approach since scaling ratios assumptions are almost always violated. P-Scale® is a quantitatively rigorous and robust statistical tool that may be used in the D&D and RA community for more cost-effective project planning and execution. It is easily coupled with the Multi-Agency Radiation Survey and Site Investigation Manual and the Data Quality Objective Process to decrease number of samples, Finally, P-Scale® may be used to provide increase accuracy in transportation documentation and individual shipment manifests.

P-Scale® technique is proprietary to Redus and Associates. Such an approach has been used in several projects at the Oak Ridge Reservation, and we welcome the opportunity to support D&D projects requiring such quantitative rigor and defensibility. The author may be contacted at kredus@icx.net.

FOOTNOTES

[a] The term leverage is used because a point exerts more influence on the fit if it is farther away from the middle of the plot in the horizontal direction. At the extremes, the differences of the residuals before and after being constrained by the hypothesis are greater and contribute a larger part of the sums of squares for that effect's hypothesis test.

REFERENCES

1. Lockheed Martin Energy Systems, Results of Fall 1994 Sampling of Gunite and Associated Tanks at the Oak Ridge National Laboratory, Oak Ridge, Tennessee, ORNL/ER/Sub-87-99053/74, 1995.
2. Lockheed Martin Energy Systems, Results of 1995 Characterization of Gunite and Associated Tanks at Oak Ridge National Laboratory, Oak Ridge, Tennessee, ORNL/ER/Sub-87-99053/79, 1996.

FIGURES AND TABLES

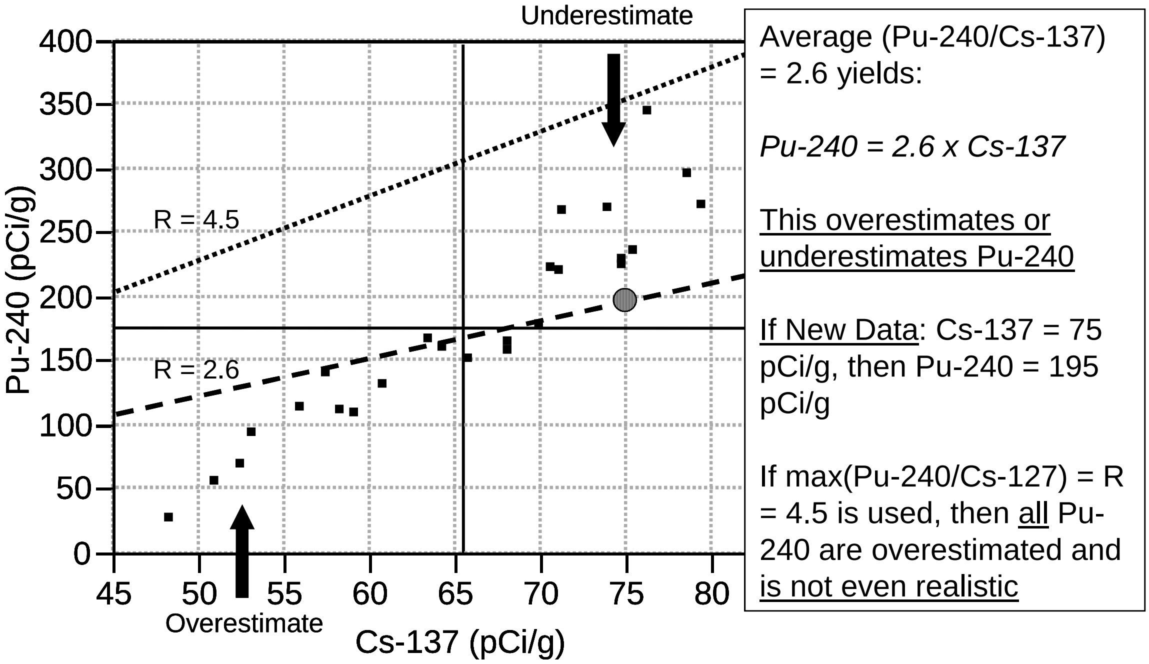

Figure 1. Illustration of the usual approach to scaling ratios.

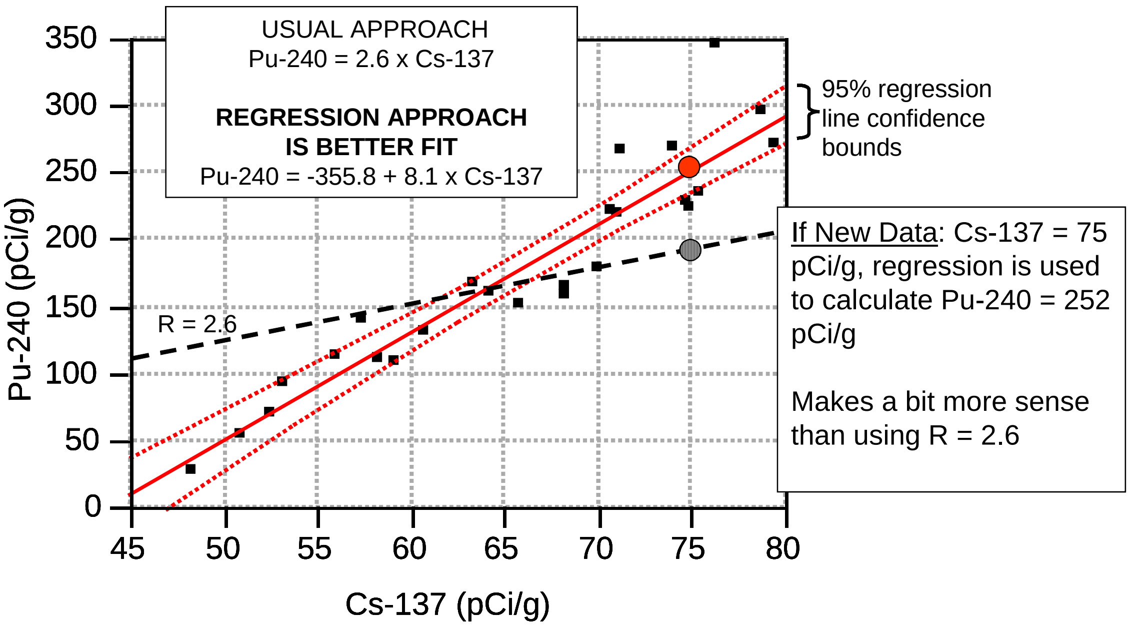

Figure 2. Illustration of regression applied to the estimating a scaling ratio.

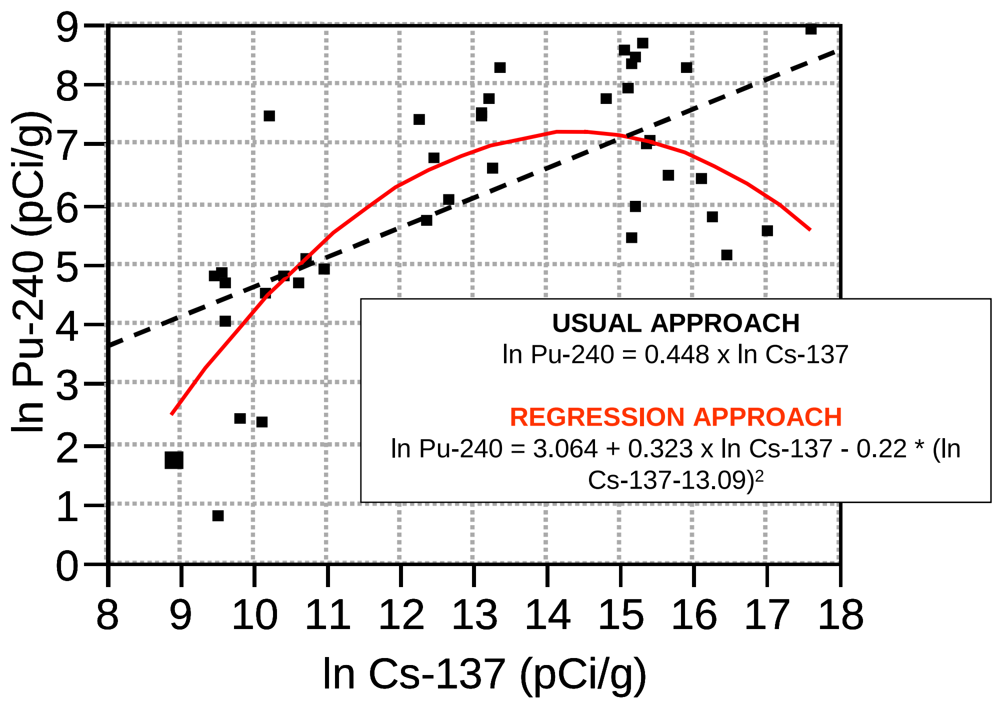

Figure 3. Gunite and Associated Tanks scatter plot for Cs-137 and Pu-240 (log scale).

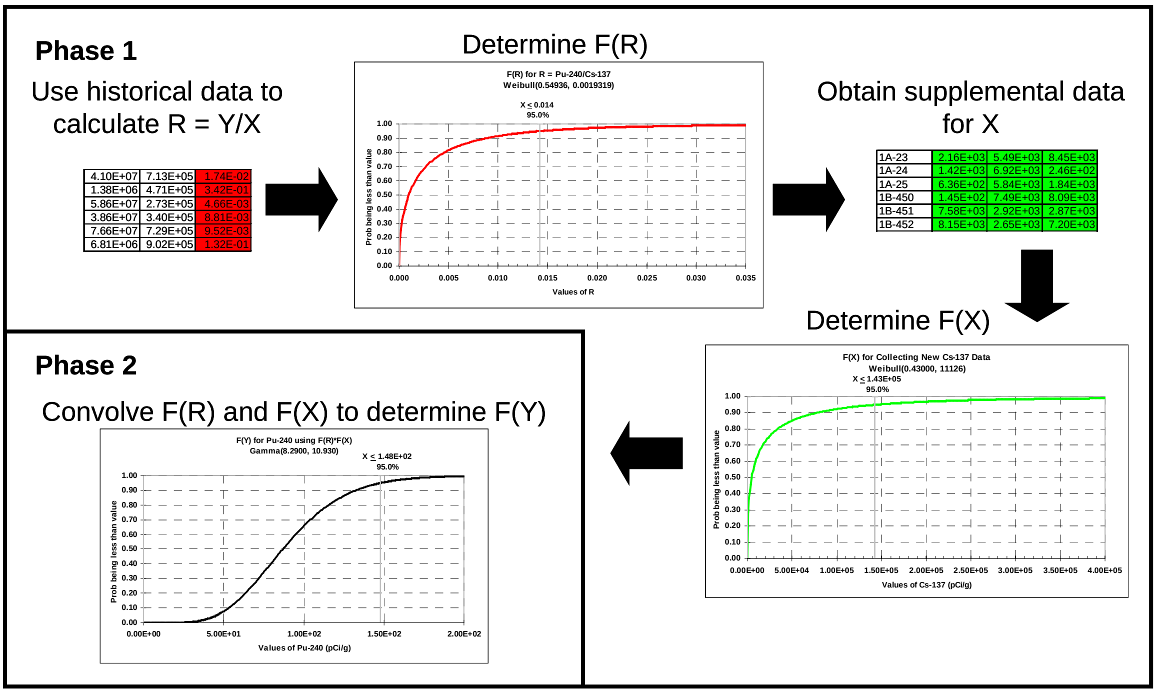

Figure 4. P-Scale® Approach.

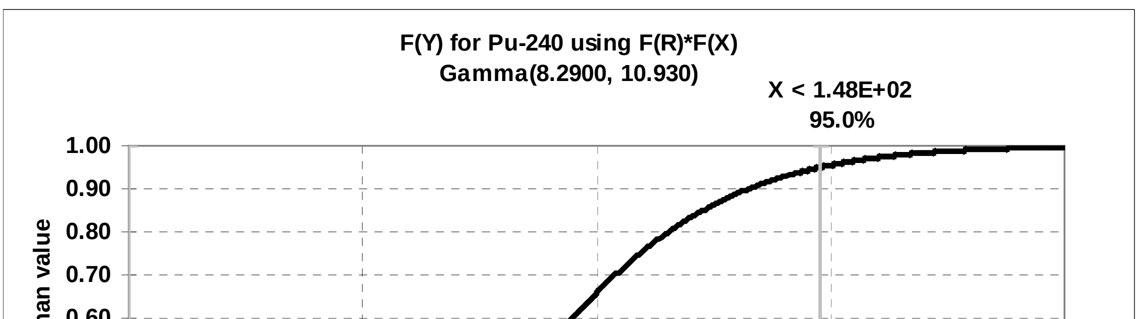

Figure 5. Results of P-Scale® for Pu-240 concentration for the contaminated equipment in GAAT B-25 boxes.