APPENDIX H
STATISTICAL ANALYSIS AND RESULTS OF SOLUBILITY DATA

M

THIS PAGE INTENTIONALLY LEFT BLANK

(M)

STATISTICAL ANALYSIS AND RESULTS OF SOLUBILITY DATA

1.0 INTRODUCTION

Engineered Alternatives (EA) that increase the pH of the brine in the repository by two to three pH units results in improved performance of the repository because of lower actinide solubilities. Statistical analysis of existing relevant solubility data was performed to quantify the effects of such a pH increase on radionuclide solubilities.

Linear segments of the data sets displayed in Appendix E (Figures E-1 through E-12) were evaluated with a simple linear regression model to establish estimates of actinide solubilities at pH values of 6.1 and 8.3. Generally, the solubility of the radionuclides followed a linear trend between these pH values. The pH value of 6.1 corresponds to an indigenous Salado Formation brine (Deal et al., 1995) evaluated in the baseline case. The pH value of 8.3 corresponds to the approximate pH established in Salado brine by the brucite buffer when a limited amount of lime (CaO) is added to the backfill (Appendix G). A simple linear regression analysis was carried out on the linear or near linear data segments using Statgraphics Plus Version 7.0 software. The results of the regression analyses are shown in Figures H-1 through H-5. The estimated actinide solubilities and ranges are summarized in Table H-1 and additional statistical parameters are given in Table H-2.

2.0 REGRESSION ANALYSIS METHODOLOGY

Figures H-1 through H-5 show the data points used for regression analysis as well as the bold regression line. The solubility data were regressed using a least squares linear regression. Curved lines directly above and below the regression line (Figures H-1 through H-5) indicate the 95 percent confidence on the regression line. The true population regression line has a 95 percent chance of occurring between these two curves. The two curved lines farthest from the regression line in Figures H-1 through H-5 are the upper and lower 95 percent prediction intervals. These curves indicate the confidence interval for predicting a single value of the concentration at any pH. Upper and lower values reported for the 95 percent prediction interval represent the estimated range of solubility for the given actinide at the pH values of 6.1 and 8.3. The range takes into account the experimental uncertainties in the data and the uncertainties introduced by predicting a solubility at a pH value of interest. The estimated actinide solubilities and ranges determined from these curves are summarized in Table H-1.

In addition to the regression line and the confidence and prediction intervals around the regression line, several other regression analysis statistics were also calculated. These additional statistics are, the intercept and slope of the regression line, the correlation coefficient, the coefficient of determination, the standard error of the estimate, a T-test for significance of slope and intercept, and an analysis of variance (F-test) for significance of slope. These statistics are summarized in Table H-2.

The correlation coefficient (r), the coefficient of determination (r^2), and the standard error of the estimate are statistical parameters which describe how well the data points fit the linear regression line. The T-test and the analysis of variance test determine whether the concentration

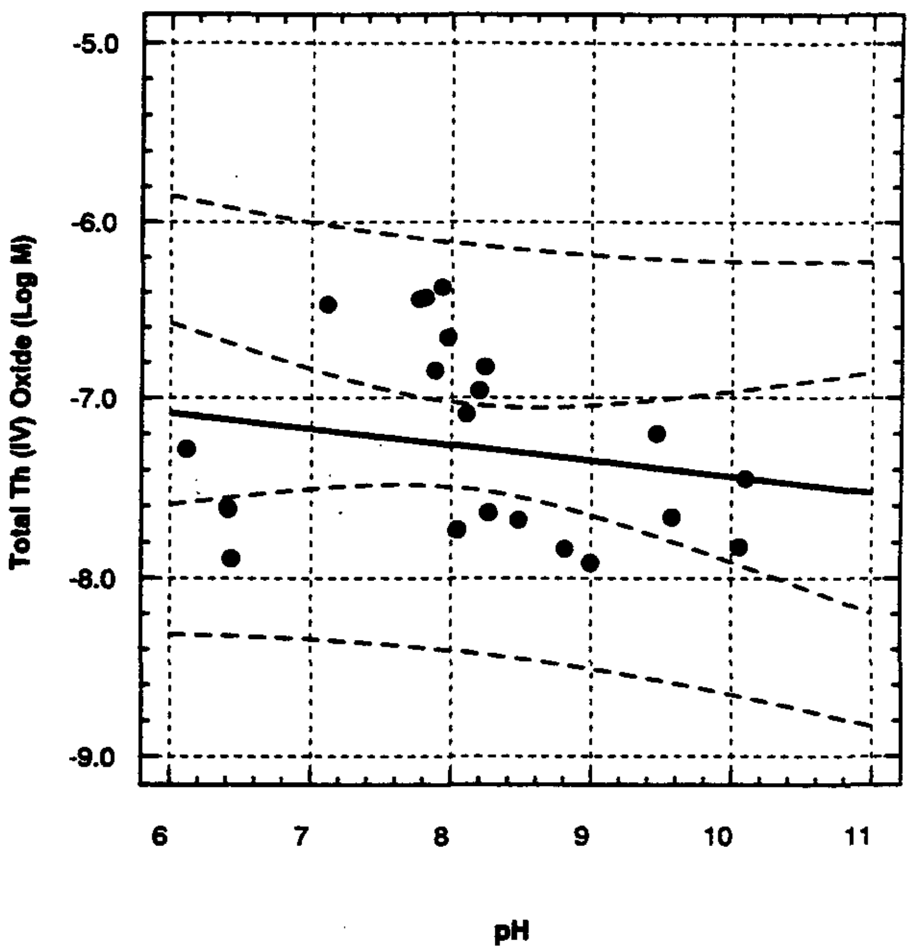


Figure H-1
Linear Regression for Hydrous Th (IV) oxide Solubility as a Function of pH for 3 M Solutions of NaCl; Data from Felmy et al. (1991)

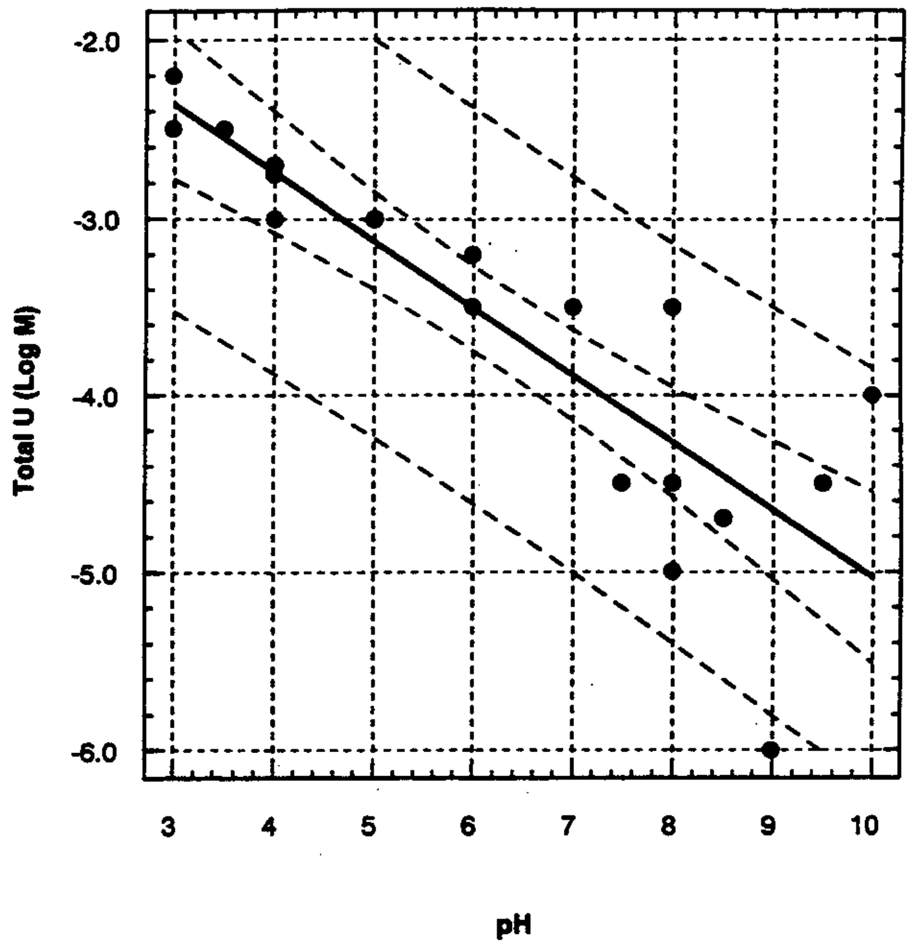


Figure H-2
Linear Regression for UO_3:$2H_2O$ Solubility as a Function of pH for Dilute Solutions; Data from Krupka et al. (1985)

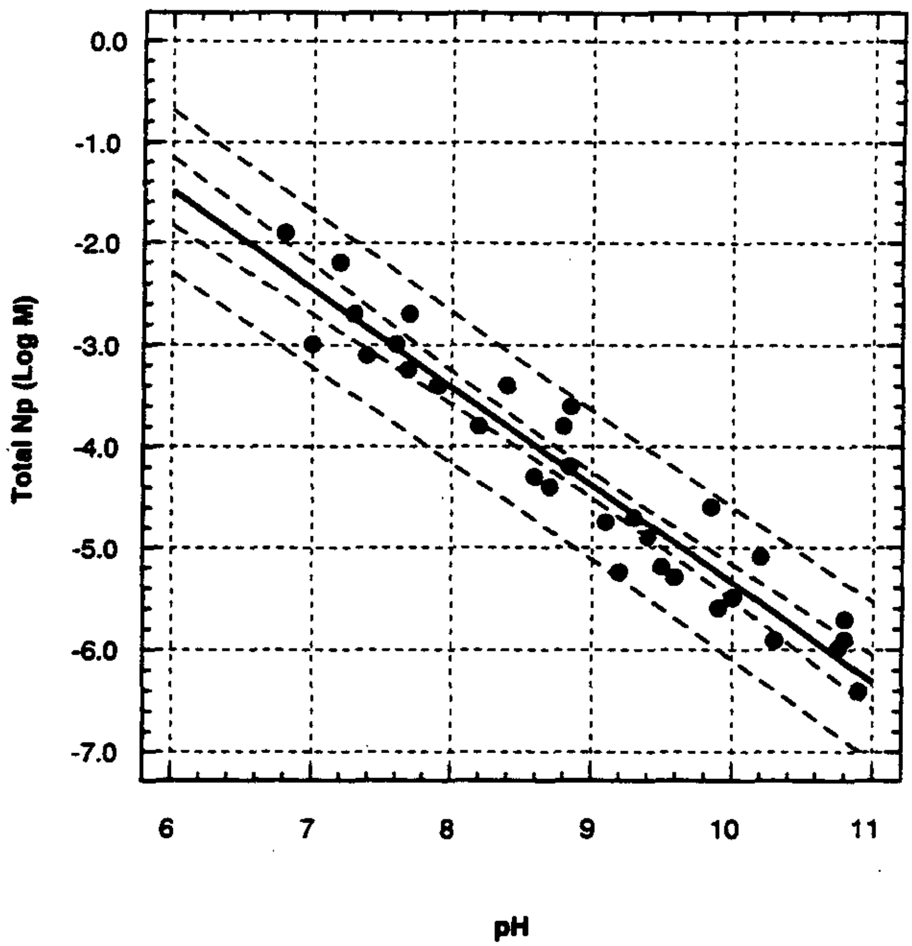
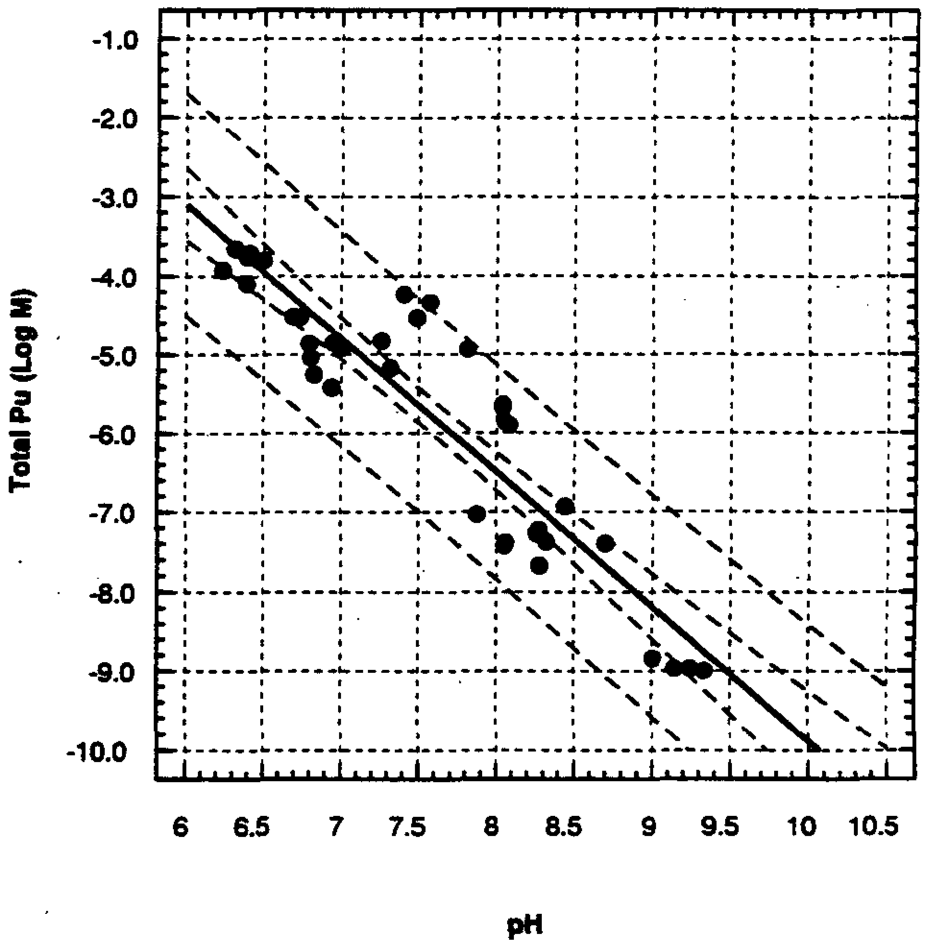


Figure H-3
Linear Regression for $NpO_2(OH)$ Solubility as a Function of pH for 1 M and 3 M Solutions of $NaClO_4$; Data from Kim et al. (1985b) and Neck et al. (1992).



M

Figure H-4
Linear Regression for Amorphous $^{239}Pu(OH)_3$ Solubility as a Function of pH for PBB1 and PBB3 Brine Solutions; Data from Rai et al. (1987)

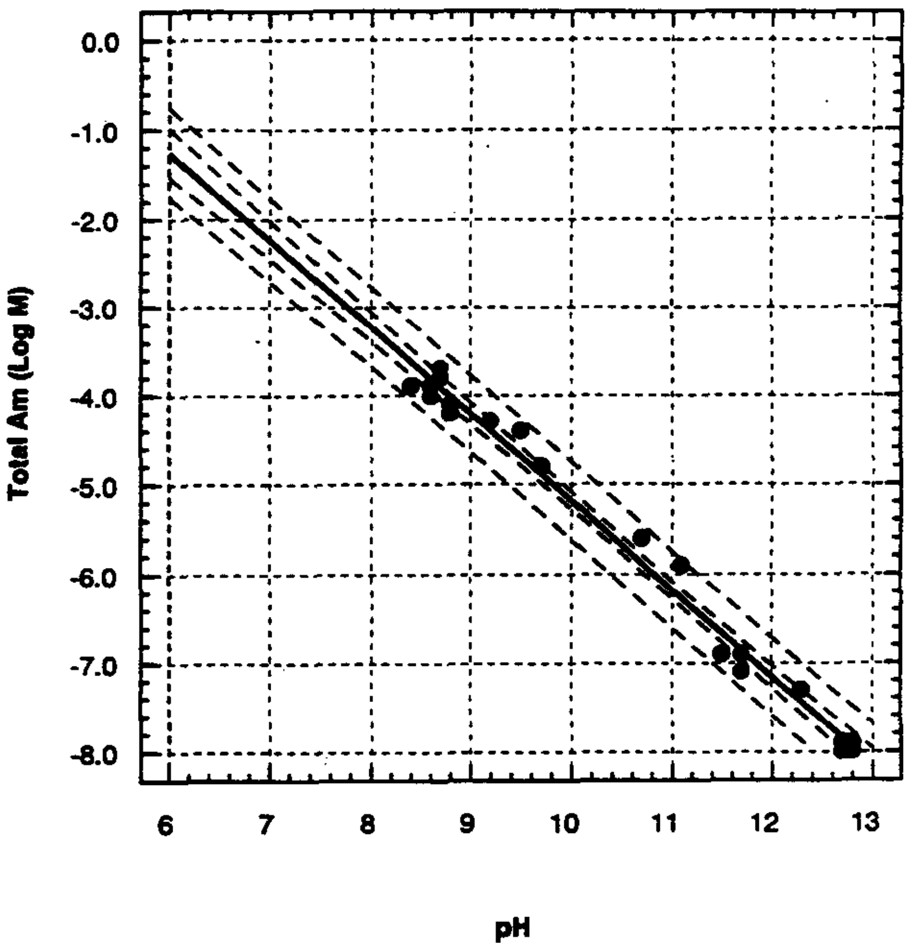


Figure H-5
Linear Regression for AmO2(OH) Solubility as a Function of pH for 5 M Solutions of NaCl; Data from Kim et al. (1985b)

and pH values are linearly correlated by examining the slope and intercept of the regression line. Each of these statistical parameters and procedures are described below.

The coefficient of determination (r^2) is the square of the correlation coefficient (r). The coefficient of determination is defined as the ratio of the sum of squares of the error in the regression line (explained variation) to the sum of squares of the total error (Kennedy and Neville, 1986). This ratio must lie between zero and unity. When $r^2 = 1$, all the variation is explained and the data points have a perfect fit with the regression line. The closer r^2 is to unity the less unexplained variation exists and the better the data points fit the regression line.

The standard error of the estimate is a measure of the variability of the dependent variable estimated by the regression line. The standard error of the estimate is defined as the square root of the sum of the squares of the residuals divided by the number of degrees of freedom (Kennedy and Neville, 1986). The residuals or deviations are the amount of difference between the modeled regression line and the measured value of concentration. Thus, if the sum of the squares of the residuals are small, the data points all fall close to or on the regression line, and the value of The Standard Error of the Estimate is low. The lower the Standard Error of the Estimate, the greater the confidence that the linear regression fits the data.

The T-test for significance of slope and intercept is used to determine if the slope or intercept of the regression line are significantly different than zero (Kennedy and Neville, 1986). Likewise, the analysis of variance (F-test), tests for the significance of slope (Kennedy and Neville, 1986). A slope and/or intercept of zero may indicate no correlation or dependence of concentration on pH. These tests are performed under the hypothesis that the slope or the intercept are equal to zero (null hypothesis), and then testing this hypothesis against the alternative hypothesis that the slope or the intercept are not equal to zero. The test results are indicated by the probability level. The probability level is the probability of a Type I error or the probability of rejecting the null hypothesis when it is true (Kennedy and Neville, 1986). Thus, if the probability level is greater than 0.05, there is no reasonable doubt to reject the null hypothesis at the 95th percent confidence level and it is concluded that the slope or intercept is equal to zero. If the probability level is less than 0.05, the null hypothesis is rejected at the 95th percent confidence level and it is concluded that the slope or the intercept are not equal to zero.

These statistical parameters and tests were used to determine how well the linear model fits the solubility data over the pH range of interest. If the solubility data follow a linear trend the coefficient of determination (r^2) will be close to unity, and the standard error of the estimate will be low. If there is a strong correlation between concentration of the radionuclides and pH, the regression line will also have a non-zero slope. If the linear model fits the solubility data and the regression line is significant, then there is a high degree of confidence in the predicted values of solubility at pH values of 6.1 and 8.3.

3.0 RESULTS OF REGRESSION ANALYSES

A discussion of the solubility of Th, U, Np, Pu and Am with respect to Ph is presented below. Results from the regression analysis and predicted solubilities at Ph values of 6.1 and 8.3 are discussed.

3.1 Thorium

The Th solubility data of Felmy et al. (1991) for amorphous $ThO_2 \cdot xH_2O$ in 3M NaCl solutions (Figure H-1) are the closest approximation to saturated NaCl conditions imposed by Salado brine. Data points falling in the pH interval of 6 to 10 were used in the linear regression model and results are displayed on Figure H-1, with Table H-1 and Table H-2 summarizing results of the regression analysis. The coefficient of determination value (r^2 = 3 percent) given on Figure H-1 indicate the data are a very poor fit to the linear model. Additionally, results from the T-test for significance of slope and results from the analysis of variance (Table H-2) indicate that the slope is equal to zero at the 95th percent confidence limit (probability level = 0.4).

Felmy et al. (1991) noted that most of the measured Th concentrations above a pH of 6 are near the analytical detection limit of about -8.5 log M, and greater error in analytical precision probably accounts for the large scatter in the data set. However, the solubility of ThO_2 is known to be independent of pH above near neutral pH values (Felmy et al., 1991), and theoretical calculations of the solubility of crystalline ThO_2 indicate a linear trend for Th concentration above a pH of 6, corresponding to an intercept of -14 log M and a slope of zero (Langmuir and Herman, 1980). Therefore, the solubility estimates determined from linear regression (Table H-1) were used to estimate $ThO_2 \cdot xH_2O$ solubility because analytical measurement of Th concentration is difficult above a pH of 7, due to the insoluble nature of ThO_2, and a linear relationship is known to exist once ThO_2 solubility becomes independent of pH.

3.2 Uranium

Uranium solubility data (Figures H-2 and H-3) indicate that schoepite ($UO_3 \cdot 2H_2O$) is more soluble than uraninite (UO_2) and, as a conservative measure, the most soluble U phase is selected for this analysis. The solubility data reported by Krupka et al. (1985) for amorphous schoepite were selected over that of Bruno and Sandino (1989) because of the greater pH range studied. Similar U concentrations are reported for amorphous schoepite when the two studies overlap in pH space (Figure H-3). Figure H-2 displays results for the linear model over the pH range of 3 to 10, Table H-1 and Table H-2 summarize the results from the regression analysis. The coefficient of determination on Figure H-2 indicate that the data of Krupka et al. (1985) are a fair fit to the model (r^2 = 75 percent). Additionally, results from the T-test for significance and the analysis of variance indicate that the slope and intercept are not equal to zero at the 95th percent confidence level (probability level = 0.0).

Solubility data for schoepite in saturated NaCl brine was not found during this study. Examining Figure H-2, it appears that the solubility of uraninite is increased by about 2 orders of magnitude in the presence of a saturated NaCl brine. Therefore, results summarized in Table H-1 for shoepite solubility in dilute solutions are recommended to be increased by two orders of magnitude to account for the expected enhancement of schoepite solubility in Salado brine.

3.3 Neptunium

Solubility data on Np solids (Figures H-4 and H-5) indicate that NpO_2OH is more soluble than NpO_2 and, as a conservative measure, the most soluble Np phase is selected for this analysis. The 1M and 3M $NaClO_4$ data sets of Neck et al. (1992) and Kim et al. (1985b) are in good agreement and the closest approximation to saturated NaCl conditions imposed by Salado brine. Therefore, these data sets were selected for regression analysis over the pH interval of 6 to 11.

Figure H-3 summarizes the regression analysis for the data sets. The r^2 value of 92 percent, indicates that the data sets are a good fit to the linear model and the T-test and analysis of variance indicate a non zero slope at the 95^{th} percent confidence limit (probability level = 0.0). Table H-1 summarizes the model results for the pH values of interest and Table H-2 summarizes additional regression statistics.

3.4 Plutonium

Figures H-6 through H-10 summarize data on the solubility of Pu solids. The most soluble phase over the pH interval of interest (6 to 9) is amorphous $Pu(OH)_3$ (Figure H-6) and, as a conservative measure, the most soluble Pu phase is selected for this analysis. Rai et al. (1987) studied the solubility of amorphous $Pu(OH)_3$ in Permian Basin brines (PBB), and these data sets are chosen for regression analysis because the saturated NaCl PBB are good analog for Salado brine. Regression results presented on Figure H-4 indicate the linear model is a fair to good fit to the data sets [r^2 = 85 percent, slope and intercept are not equal to zero (probability level = 0.0)]. Table H-1 summarizes the model results for the pH values of interest and Table H-2 summarizes additional regression statistics.

3.5 Americium

Americium solubility data are presented on Figures H-11 and A-12, and the most soluble phase indicted on these plots is AmO_2OH (Figure H-12). The solubility data of Kim et al. (1985b) for $^{241}AmO_2OH$ in 5M NaCl solution is selected for this analysis, because these data represent the closest approximation to saturated NaCl conditions imposed by Salado brine. Figure H-5 summarizes the regression analysis and indicates that the linear model is an excellent fit to the data over the pH interval of 8 to 13 [r^2 = 99 percent, slope and intercept are not equal to zero (probability level = 0.0)]. Note that the regression analysis has been extrapolated to a pH of 6 to allow a calculation of the solubility estimate at a pH of 6.1 (baseline case for Salado brine). Table H-1 summarizes the model results for the pH values of interest and Table H-2 summarizes additional regression statistics.

TABLE H-1

RESULTS OF REGRESSION ANALYSIS ON ACTINIDE SOLUBILITY

Actinide	Most Soluble Oxide/Hydroxide Phase	Solution Composition	pH	Estimated Solubility	Estimated Range of Solubility
				log (mole/liter)	
Thorium	$ThO_2 \cdot xH_2O$	3M NaCl	6.1	-7.1	-5.9 to -8.3
			8.3	-7.3	-6.1 to -8.4
Uranium	$UO_3 \cdot 2H_2O(a)$	dilute	6.1	-3.5[a]	-2.4 to -4.7[a]
			8.3	-4.4[a]	-3.2 to -5.5[a]
Neptunium	$NpO_2OH(a)$	1M and 3M $NaClO_4$	6.1	-1.6	-0.8 to -2.4
			8.3	-3.7	-3.0 to -4.4
Plutonium	$Pu(OH)_3(a)$	Permian Basin Brines	6.1	-3.3	-1.9 to -4.7
			8.3	-7.0	-5.6 to -8.4
Americium	$AmO_2OH(a)$	5M NaCl	6.1	-1.4[b]	-0.9 to -1.8[b]
			8.3	-3.5	-3.1 to -4.0

[a]It is recommended that values be increased by 2 log units to account for enhanced solubility in saturated NaCl solutions. See text for discussion.
[b]Value is obtained from extrapolation of regression analysis to pH 6.1.

TABLE H-2

REGRESSION STATISTICS ON ACTINIDE SOLUBILITY

1) Regression Analysis - Thorium

Dependent variable: Th(IV) Oxide   Independent variable: pH

Parameter	Estimate	Standard Error	T Value	Prob. Level
Intercept	-6.5484	0.847345	-7.72814	0.00000
Slope	-0.0889158	0.103675	-0.857641	0.40077

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	Prob. Level
Model	0.2137470	1	0.2137470	0.735547	0.40077
Residual	6.102514	21	0.290596		
Total (Corr.)	6.316261	22			

Correlation Coefficient = -0.183959   R-squared = 3.38 percent
Stnd. Error of Est. = 0.539069

2) Regression Analysis - Uranium

Dependent variable: UO3+2H2O   Independent variable: pH

Parameter	Estimate	Standard Error	T Value	Prob. Level
Intercept	-1.21305	0.337455	-3.59472	0.00207
Slope	-0.382024	0.0515513	-7.41055	0.00000

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	Prob. Level
Model	14.674492	1	14.674492	54.91627	0.00000
Residual	4.809883	18	0.267216		
Total (Corr.)	19.484375	19			

Correlation Coefficient = -0.867837   R-squared = 75.31 percent
Stnd. Error of Est. = 0.516929

M

TABLE H-2 (CONTINUED)

REGRESSION STATISTICS ON ACTINIDE SOLUBILITY

3) Regression Analysis - Neptunium

Dependent variable: NpO2OH — Independent variable: pH

Parameter	Estimate	Standard Error	T Value	Prob. Level
Intercept	4.30194	0.476779	9.02292	0.00000
Slope	-0.964968	0.052955	-18.2224	0.00000

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	Prob. Level
Model	41.85969	1	41.85969	332.0563	0.00000
Residual	3.655798	29	0.126062		
Total (Corr.)	45.515484	30			

Correlation Coefficient = -0.959 — R-squared = 91.97 percent

Stnd. Error of Est. = 0.355052

4) Regression Analysis - Plutonium

Dependent variable: Pu(OH)3 — Independent variable: pH

Parameter	Estimate	Standard Error	T Value	Prob. Level
Intercept	7.07026	0.928106	7.61794	0.00000
Slope	-1.69515	0.121298	-13.9751	0.00000

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	Prob. Level
Model	83.86683	1	83.86683	195.3031	0.00000
Residual	15.029656	35	0.429419		
Total (Corr.)	98.896481	36			

Correlation Coefficient = -0.920883 — R-squared = 84.80 percent

Stnd. Error of Est. = 0.6553

TABLE H-2 (CONTINUED)

REGRESSION STATISTICS ON ACTINIDE SOLUBILITY

5) Regression Analysis - Americium

Dependent variable: AmO2OH — Independent variable: pH

Parameter	Estimate	Standard Error	T Value	Prob. Level
Intercept	4.64694	0.284859	16.3131	0.00000
Slope	-0.983439	0.0269227	-36.5283	0.00000

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	Prob. Level
Model	53.1451	1	53.1451	1334.316	0.00000
Residual	0.716930	18	0.039829		
Total (Corr.)	53.862000	19			

Correlation Coefficient = -0.993322 — R-squared = 98.67 percent

Stnd. Error of Est. = 0.199573

THIS PAGE INTENTIONALLY LEFT BLANK

M