548788

Sensitivity of the CRA-2009 Performance Assessment Calculation Releases to Parameters

Author:	Thomas Kirchner (6711)	Thomas B Kirchner	5/2/08
	Print	Signature	Date
Technical Review:	Sean Dunagan (6711)	Sean Dunagan	5/5/08
	Print	Signature	Date
Management Review:	Moo Lee (6711)	MLee	5/5/08
	Print	Signature	Date
QA Review:	Mario Chavez (6710)	Mario Chavez	5/5/08
	Print	Signature	Date

WIPP: 1.2.5: PA: QA-L: 547488

Table of Contents

Introduction

AP-137 Revision 1, "Analysis Plan for the Performance Assessment for the 2009 Compliance Recertification Application " (Clayton 2008a) , specified that a sensitivity analysis be conducted on the results of the CRA-2009 Performance Assessment (CRA-2009 PA). This report documents the results of that analysis. The code STEPWISE version 2.21 was used to determine the relative importance of the sampled parameters in the CRA-2009 PA. STEPWISE receives sampled input parameter values and calculated release data that correspond to those input data. STEPWISE relates the sampled input parameter values to the calculated release data by performing a multiple regression analysis and reporting the results in tables. Additional analyses were performed on selected subsets of the data using Microsoft® Excel. The spreadsheets are provided on the attached are also stored in the CMS library LIBCRA09 in the compressed file AP137SimpleRegressionSpeadsheets.zip (Attachment 1).

WIPP PA employs stepwise linear multiple regression to evaluate the relative importance of the various sampled parameters on the estimates of potential releases. In the forward stepwise approach used by STEPWISE, a sequence of regression models is constructed, starting with the input parameter that has the strongest simple correlation with the output variable. Partial correlations between the output and the remaining variables are then computed. The partial correlations remove the linear effects of variables already included in the model. The variable having the largest significant partial correlation coefficient is added next, and the partial correlations for the remaining input variables are recomputed. Significance is determined using an F-test, and the significance level for adding an input variable to the model is $1-\alpha_{in}$, where α_{in} is the significance level for a Type I error that is set by the analyst. The F-test compares the variability contributed by the variable to the variability not accounted for by the regression, i.e. the variability of the residuals. By default STEPWISE sets $\alpha_{in} = 0.05$, so that one is 95% confident that there is a partial correlation between the input and output variables. This process is repeated until there are no variables remaining having significant correlations with the output variable. Variable excluded from the regression model contribute no significant information in relation to the unexplained variability and hence the results are judged to be relatively insensitive to those parameters.

Input variables that are added to the regression model are not necessarily retained. For an input variable to be retained, its regression coefficient, i.e. the linear contribution of an input to the prediction of the output variable, must be statistically distinguishable from zero. A t-test is used to determine whether a regression coefficient is significantly different than zero. The t-test evaluates the null hypothesis that the regression coefficient is zero. The hypothesis is not rejected when random effects can give rise to the observed regression coefficient with probability α_{out}. The random effects are caused by the stochastic variability contributed by the input variables not in the regression model. In other words, the hypothesis is rejected, and the variable is included in the model when the $1-\alpha_{out}$ confidence interval of the regression coefficient does not encompass zero. By default the STEPWISE α_{out} -value for allowing a variable to enter the regression model is 0.05. Thus, in the default case, one is 95% confident that the input variables make a linear contribution to the response of the output variable. The user may specify different α-values in the input control file. However, the value allowing a variable to enter the model, α_{in}, must be less than or equal to the value by which a variable is allowed to leave the model, α_{out}, to avoid looping. In the following analyses, α_{in} was 0.05, and α_{out} was 0.05.

Setting α_{out} to α_{in} maximizes the number of variables in the model (as requested by EPA in C-23-18, U.S. Environmental Protection Agency 2004) but can increase the number of spurious correlations (Kirchner 2004a, Kirchner 2004b). A spurious correlation implies a linear relationship exists between two variables but in reality no such relationship exists. The predicted error sum of squares (PRESS) was computed to detect over-fitting of the regression model to the data. Over-fitting can occur when the regression methodology causes the fit to favor specific points rather than the general shape of the data curve. In such a case the minimum value of PRESS may occur earlier than the last step in the regression analysis. No such condition was observed in any of the rank correlation analyses reported herein.

The STEPWISE procedure constructs a multivariate linear regression model. One of the assumptions of this statistical model is that the dependent (output) variable shows a linear response to the dependent (input) variables. In cases where the response is non-linear but monotonic, replacing the values of the data with their ranks tends to linearize the response curves and standardizes the variability in the outputs and parameters by mapping the data into identical ranges. The rank of a value is an integer representing its position in the sorted list of the values. Ranking also tends to de-emphasize the impact of "outliers," which are points having considerably larger or smaller values than the remainder of the sample population. Although the use of ranks precludes using the model to predict values of an output variable given an input variable, the results are usually well suited for ranking the importance of the contributions of the input variables to the response of the output variable. The STEPWISE procedure has the functionality to perform ranked regressions. For the cases described below, the ranked regressions showed stronger correlations than the regressions based on the unranked data. This result suggests that there are non-linear relationships between the dependent and independent variables, but it does not eliminate the possibility that there are also non-monotonic relationships.

Ranked regression was used to evaluate the sensitivity of the output variables to the sampled parameters. Scatter plots of the dependent versus independent ranked variables resulting from the analysis were examined to determine if there were any obvious non-monotonic relationships. Obvious non-monotonic relationships were not found although there are cases involving inputs that are categorized as discrete variables (e.g. GLOBAL:OXSTAT, which is sampled as a uniform distribution but is then mapped to one of two discrete values) and cases where there are large proportions of the vectors showing no release (e.g. releases from the Culebra). Application of linear regression to such cases is somewhat problematic in terms of the assumptions of normally-distributed residuals and homogeneous variance among the residuals. However, in terms of ranking the relative importance of the parameters these issues are probably not significant.

Most of the regression models produced by STEPWISE do not include all of the variables, even after ranking the data. This simply indicates that the uncertainties in many of the parameters have statistically insignificant effects on the output variable. Statistical insignificance can arise because the output variable has a low functional response to the input variable, because the magnitude of uncertainty in the input variable is small relative to the other inputs, or from a combination of both conditions. This is not to say that these non-significant variables have no influence on the releases. Their exclusion from the tables reflects the inability of this statistical technique to rank their importance with an acceptable degree of confidence. For example, if the response of the output variable to an input variable was non-monotonic then the regression analysis might fail to properly identify that variable's importance. This possibility is unlikely for total releases and cuttings and cavings releases because the R^2 value indicates that more than 88% of the variability in the output variables has been accounted for by the listed input variables.

Several of the parameters that appear in the model often contribute very little to the R^2 value and, therefore, explain very little of the variability in the output variable. Parameters that have minor contributions can appear by chance, simply due to random correlations. Many of the parameters that account for only a few percent to the variability in an output from one replicate may show different rankings, or can even be absent, in another replicate. Thus, it is difficult to assess the importance of the parameters that improve the regression model very little and, in reality, they may have no importance at all. Therefore, only the parameters that appear to have significant impacts on the regression model will be explained in detail.

Analyses were performed using Excel to supplement the STEPWISE results for spallings and Culebra releases because the majority of those releases were zero. These analyses used the data from CRA-2009 PA Replicate 1 to illustrate the relationship between the non-zero releases and various parameters. Simple linear regression analyses were conducted on the ranks of the parameters and non-zero releases. Significance of the regressions is denoted by the probability that the correlation is due to random error, as in $p = 0.09$.

This report documents the results of the CRA-2009 PA sensitivity analysis and shows, for comparison, the results obtained for the CRA-2004 PABC analysis (Kirchner 2005). The details of run control for the PRECCDFGF and CCDFGF results presented herein are documented in Long (2008). The files associated with the CRA-2009 PA STEPWISE analysis can be found in the CMS libraries LIBCRA09_EVAL, LIBCRA09_STP, LIBCRA09_LHS, LIBSTP and LIBCRA09_CCGF (Table 1). These files include the run script (EVAL_STP_CRA09_ALL.COM), the CCGFDF output files CCGF_CRA09_R*r*.OUT, the input files containing the mean values of the model output variables (CRA09_STEPWISE_MEANS_R*r*.TRN), the STEPWISE control files (STP_CRA09_RAW_ALL_R*r*.INP and STP_CRA09_RANK_ALL_R*r*.INP), the input file containing the sampled parameters (CRA09_STEPWISE_LHS_R*r*.TRN) and the results of the regression analysis (STP_CRA09_RANK_R*r*.TXT and STP_CRA09_RAW_R*r*.TXT), where the *r* in the file names represents replicate number 1, 2 or 3.

The CRA09_STEPWISE_LHS_R*r*.TRN files are created from the LHS2_CRA09_R*r*_CON.TRN files using the Microsoft® Access database CCDFGF_Analysis.mdb which links to the Access database CCDFGF_Data.mdb. The files of mean values of releases, CRA09_STEPWISE_MEANS_R*r*.TRN, were also generated from the CCGF_CRA09_R*r*.OUT files using the Access database CCDFGF_Analysis.mdb. The CCDFGF output files store the mean values and those mean values are extracted from those files along with the "binned" data. The "bins" are a series of equally spaced intervals on a logarithmic scale that encompass the data. Each datum is assigned to a bin and the frequency of values within the bin tabulated. The databases are provided on the attached CD in zipped format and the AP137_ANALYSISDATABASE.ZIP file is also stored in the CMS library LIBCRA09_STP (Attachment 1).

EVAL_STP_CRA109_ALL.COM runs STEPWISE.EXE to perform the regression analysis for all three replicates on both the ranked and raw (un-ranked) data.

Table 1. Files used in the CRA-2009 PA sensitivity analysis.

File	Library
CCDFGF_Analysis.mdb[1]	LIBCRA09_STP
CCDFGF_Data.mdb[1]	LIBCRA09_STP
CCGF_CRA09_R1.OUT	LIBCRA09_CCGF
CCGF_CRA09_R2.OUT	LIBCRA09_CCGF
CCGF_CRA09_R3.OUT	LIBCRA09_CCGF
LHS1_CRA09_R1.TRN	LIBCRA09_LHS
LHS1_CRA09_R2.TRN	LIBCRA09_LHS
LHS1_CRA09_R3.TRN	LIBCRA09_LHS
LHS2_CRA09_R1_CON.TRN	LIBCRA09_LHS
LHS2_CRA09_R2_CON.TRN	LIBCRA09_LHS
LHS2_CRA09_R3_CON.TRN	LIBCRA09_LHS
CRA09_LHS_PARAMNAMES.INP	LIBCRA09_STP
EVAL_STP_CRA09_ALL.COM	LIBCRA09_STP
CRA09_STEPWISE_LHS_R1.TRN	LIBCRA09_STP
CRA09_STEPWISE_LHS_R2.TRN	LIBCRA09_STP
CRA09_STEPWISE_LHS_R3.TRN	LIBCRA09_STP
CRA09_STEPWISE_MEANS_R1.TRN	LIBCRA09_STP
CRA09_STEPWISE_MEANS_R2.TRN	LIBCRA09_STP
CRA09_STEPWISE_MEANS_R3.TRN	LIBCRA09_STP
STP_CRA09_RANK_ALL_R1.INP	LIBCRA09_STP
STP_CRA09_RANK_ALL_R2.INP	LIBCRA09_STP
STP_CRA09_RANK_ALL_R3.INP	LIBCRA09_STP

File	Library
STP_CRA09_RANK_R1.TXT	LIBCRA09_STP
STP_CRA09_RANK_R1.SP	LIBCRA09_STP
STP_CRA09_RANK_R2.TXT	LIBCRA09_STP
STP_CRA09_RANK_R2.SP	LIBCRA09_STP
STP_CRA09_RANK_R3.TXT	LIBCRA09_STP
STP_CRA09_RANK_R3.SP	LIBCRA09_STP
STP_CRA09_RAW_ALL_R1.INP	LIBCRA09_STP
STP_CRA09_RAW_ALL_R2.INP	LIBCRA09_STP
STP_CRA09_RAW_ALL_R3.INP	LIBCRA09_STP
STP_CRA09_RAW_R1.TXT	LIBCRA09_STP
STP_CRA09_RAW_R1.SP	LIBCRA09_STP
STP_CRA09_RAW_R2.TXT	LIBCRA09_STP
STP_CRA09_RAW_R2.SP	LIBCRA09_STP
STP_CRA09_RAW_R3.TXT	LIBCRA09_STP
STP_CRA09_RAW_R3.SP	LIBCRA09_STP
STP_CRA09_SCATTER_RANK_R1.PS	LIBCRA09_STP
STP_CRA09_SCATTER_RANK_R2.PS	LIBCRA09_STP
STP_CRA09_SCATTER_RANK_R3.PS	LIBCRA09_STP
TP_CRA09_SCATTER_RAW_R1.PS	LIBCRA09_STP
TP_CRA09_SCATTER_RAW_R2.PS	LIBCRA09_STP
TP_CRA09_SCATTER_RAW_R3.PS	LIBCRA09_STP
STEPWISE_PA96_2.EXE	LIBSTP

[1] **Compressed and stored in AP137_ANALYSISDATABASE.ZIP**

The input files for STEPWISE use short names rather than material:property designations. These short names are required because of a limitation in the length of variable names in STEPWISE. Table 2 associates these names with the material and property names.

Table 2. Material and property values associated with the variable names used in the CRA-2009 PA sensitivity analysis. References in this table refer to DOE-2004.

Variable Name	Material Name	Property Name	Description
CMKDAM3	AM+3	MKD_AM	Matrix distribution coefficient (m3/kg) for Am in +3 oxidation state. Defined in Equation (231).
BHPERM	BH_SAND	PRMX_LOG	Logarithm of intrinsic permeability (m2) of the silty sand-filled borehole (Table PA-5). Used in regions Upper Borehole and Lower Borehole in Figure PA-8.
DOMEGA	BOREHOLE	DOMEGA	Drill string angular velocity (rad/s).

			Defined in Equation (112b).
WTAUFAIL	BOREHOLE	TAUFAIL	Shear strength of waste (Pa). Defined in Equation (111).
BPCOMP	CASTILER	COMP_RCK	Bulk compressibility (Pa–1) of Castile brine reservoir. Defined fB in Equation (29) for region CASTILER of Figure PA-8
BPINTPRS	CASTILER	PRESSURE	Initial brine pore pressure in the Castile brine reservoir. Defined in Equation (50) for region CASTILER in Figure PA-8.
BPPRM	CASTILER	PRMX_LOG	Logarithm of intrinsic permeability (m2) *of the Castile brine reservoir*. Used *in* region CASTILER in Figure PA-8.
WFBETCEL	CELLULS	FBETA	Scale factor used in definition of stoichiometric coefficient for microbial gas generation (dimensionless). Defined in Equation (77) for areas Waste Panel, South RoR, and North RoR, in Figure PA-8.
CONBCEXP	CONC_PCS	PORE_DIS	Brooks-Corey pore distribution parameter (dimensionless) for panel closure concrete (Section PA-4.2.8.1). Defined in Equation (32) for region CONC_PCS of Figure PA-8 for use with Brooks-Corey model; defined in Equation (36) for use with van Genuchten-Parker model in region CONC_PCS.
CONPRM	CONC_PCS	PRMX_LOG	Logarithm of intrinsic permeability (m2) for the concrete portion of the panel closure. (Section PA-4.2.8.1). Used in region CONC_PCS in Figure PA-8.
CONBRSAT	CONC_PCS	SAT_RBRN	Residual brine saturation (dimensionless) in panel closure concrete (Section PA-4.2.8.1). Sbr defined in Equation (35) for use in region CONC_PCS in Figure PA-8.
CONGSSAT	CONC_PCS	SAT_RGAS	Residual gas saturation (dimensionless) in panel closure concrete (Section PA-4.2.8.1). Sgr defined in Equation (35) for area CONC_PCS in Figure PA-8.
PLGPRM	CONC_PLG	PRMX_LOG	Logarithm of intrinsic permeability (m2) of the concrete borehole plugs (Table PA-5). Used in region Borehole Plugs in Figure PA-8.
CFRACPOR	CULEBRA	APOROS	Culebra fracture (i.e., advective) porosity (dimensionless). Defined *in Equation* (223).
CMTRXPOR	CULEBRA	DPOROS	Culebra matrix (i.e., diffusive) porosity (dimensionless). Defined in Equation (230).
CFRACSP	CULEBRA	HMBLKLT	Culebra fracture spacing (m). Equal to half the distance between fractures (i.e., the Culebra half matrix block length). Defined in Equation (236) and Figure PA-26.

CTRANSFM	CULEBRA	MINP_FAC	Multiplier (dimensionless) applied to transmissivity of the Culebra within the land withdrawal boundary after mining of potash reserves. MF defined in Equation (216) (see section PA-4.8.2).
DRZPRM	DRZ_1	PRMX_LOG	Logarithm of intrinsic permeability (m2) of the DRZ. Used in regions Upper DRZ and Lower DRZ in Figure PA-8.
DRZ2PRM	DRZ_2	PRMX_LOG	Logarithm of intrinsic permeability in X-dimension of disturbed rock zone after completion of healing.
DRZPCPRM	DRZ_PCS	PRMX_LOG	Logarithm of intrinsic permeability (m2) of the DRZ immediately above the panel closure concrete (Section PA-4.2.8.3). Used in region DRZ_PCS in Figure PA-8.
CCLIMSF	GLOBAL	CLIMTIDX	Climate scale factor (dimensionless) for Culebra flow field. SFC defined in Equation (221).
WOXSTAT	GLOBAL	OXSTAT	Indicator variable for elemental oxidation states (dimensionless). WOXSTAT ≤0.5 indicates use of CMKDPU3, CMKDU4, WSOLPU3C, WSOLPUS, , WSOLU4C, and WSOLU4S. WOXSTAT >0.5 implies use of CMKDPU4, CMKDU6, WSOLPU4C, WSOLPU4S, WSOLU6C, and WSOLU6S.
BPPROB	GLOBAL	PBRINE	Probability that a drilling intrusion penetrates pressurized brine in the Castile Formation. See Section PA-3.5 for definition.
CTRAN	GLOBAL	TRANSIDX	Indicator variable for selecting transmissivity field. *See Section* PA-4.8.2.
WPHUMOX3	PHUMOX3	PHUMCIM	Ratio (dimensionless) of concentration of actinides attached to humic colloids to dissolved concentration of actinides for oxidation state +III in Castile brine. Defined SFHum(Castile, +3, Am) and SFHum(Castile, +3, Pu) for Equation (90).
CMKDPU3	PU+3	MKD_PU	Matrix distribution coefficient (m3/kg) for Pu in +3 oxidation state. Defined in Equation (231).
CMKDPU4	PU+4	MKD_PU	Matrix distribution coefficient (m3/kg) for Pu in +4 oxidation state. Defined in Equation (231).
HALCROCK	S_HALITE	COMP_RCK	Bulk compressibility of halite (Pa–1). Defined fB in Equation (31) for region Salado of Figure PA-8.
HALPOR	S_HALITE	POROSITY	Halite porosity (dimensionless). Defined in Equation (25g) for region Salado in Figure PA-8.
SALPRES	S_HALITE	PRESSURE	Initial brine pore pressure (Pa) in the

			Salado halite, applied at an elevation consistent with the intersection of MB 139. Defined in Equation (49) for region Salado. See Figure PA-8.
HALPRM	S_HALITE	PRMX_LOG	Logarithm of intrinsic halite permeability (m2). Used in region Salado in Figure PA-8.
ANHBCEXP	S_MB139	PORE_DIS	Brooks-Corey pore distribution parameter for anhydrite (dimensionless). Defined in Equation (32) for regions MB 138, Anhydrite AB, and MB 139 of Figure PA-8 for use with Brooks-Corey model; defined in Equations (36) for use with van Genuchten-Parker model in the same regions.
ANHPRM	S_MB139	PRMX_LOG	Logarithm of intrinsic anhydrite permeability (m2). Used in regions MB 138, Anhydrite AB, and MB 139 in Figure PA-8.
ANHBCVGP	S_MB139	RELP_MOD	Indicator for relative permeability model (dimensionless) for regions MB 138, Anhydrite AB and MB 139 in Figure PA-8. See Table PA-3.
ANRBRSAT	S_MB139	SAT_RBRN	Residual brine saturation in anhydrite (dimensionless). Sbr defined in Equation (35) for regions MB 138, Anhydrite AB, and MB 139 in Figure PA-8.
SHLPRM2	SHFTL_T1	PRMX_LOG	Logarithm of intrinsic permeability (m2) of lower shaft seal materials for the first 200 years after closure. Used in region Lower Shaft in Figure PA-8.
SHLPRM3	SHFTL_T2	PRMX_LOG	Logarithm of intrinsic permeability (m2) of lower shaft seal materials from 200 years to 10,000 years after closure. Used in region Lower Shaft in Figure PA-8.
SHUPRM	SHFTU	PRMX_LOG	Logarithm of intrinsic permeability (m2) of upper shaft seal materials. Used in region Upper Shaft in Figure PA-8.
SHURBRN	SHFTU	SAT_RBRN	Residual brine saturation in upper shaft seal materials (dimensionless). Sbr defined in Equation (35) for region Upper Shaft in Figure PA-8.
SHURGAS	SHFTU	SAT_RGAS	Residual gas saturation in upper shaft seal materials (dimensionless). Sgr defined in Equation (34) for region Upper Shaft in Figure PA-8.
WSOLVAR3	SOLMOD3	SOLVAR	Solubility multiplier for +III oxidation states
WSOLVAR4	SOLMOD4	SOLVAR	Solubility multiplier for +IV oxidation states
SPLPTDIA	SPALLMOD	PARTDIAM	Particle diameter of disaggregated waste.
REPIPERM	SPALLMOD	REPIPERM	Waste permeability of gas local to intrusion borehole.

SPLRPOR	SPALLMOD	REPIPOR	Waste porosity at time of drilling intrusion
TENSLSTR	SPALLMOD	TENSLSTR	Tensile strength of waste.
WGRCOR	STEEL	CORRMCO2	Rate of anoxic steel corrosion (m/s) under brine inundated conditions and with no CO2 present. Defined Rci in Equation (59) for areas Waste Panel, South RoR, and North RoR in Figure PA-8.
CMKDTH4	TH+4	MKD_TH	Matrix distribution coefficient (m3/kg) for Th in +4 oxidation state. Defined in Equation (231).
CMKDU4	U+4	MKD_U	Matrix distribution coefficient (m3/kg) for U in +4 oxidation state. Defined in Equation (231).
CMKDU6	U+6	MKD_U	Matrix distribution coefficient (m3/kg) for U in +6 oxidation state. Defined in Equation (231).
WBIOGENF	WAS_AREA	BIOGENFC	Probability of obtaining sampled microbial gas generation rates.
WGRMICH	WAS_AREA	GRATMICH	Rate of CPR biodegradation (mol C6H10O5 / kg C6H10O5 / s) under anaerobic, humid conditions. Defined Rmh in Equation (61) for areas Waste Panel, South RoR, and North RoR, in Figure PA-8.
WGRMICI	WAS_AREA	GRATMICI	Rate of CPR biodegradation (mol C6H10O5 / kg C6H10O5 / s) under anaerobic, brine-inundated conditions. Defined Rmi in Equation (61) for areas Waste Panel, South RoR, and North RoR, in Figure PA-8.
WMICDFLG	WAS_AREA	PROBDEG	Index for model of microbial degradation of CPR materials (dimensionless). Used in areas Waste Panel, South RoR, and North RoR in Figure PA-8.
WRBRNSAT	WAS_AREA	SAT_RBRN	Residual brine saturation in waste (dimensionless). Defined Sbr in Equation (34) for areas Waste Panel, South RoR, and North RoR, in Figure PA-8; also used in waste material in Figure PA-20 for calculation of DBR; see Section PA-4.7.
WRGSSAT	WAS_AREA	SAT_RGAS	Residual gas saturation in waste (dimensionless). Sgr defined in Equation (35) for areas Waste Panel, South RoR, and North RoR in Figure PA-8; also used in waste material in Figure PA-20 for calculation of DBR; see Section PA-4.7.
WASTWICK	WAS_AREA	SAT_WICK	Increase in brine saturation of waste due to capillary forces (dimensionless). Defined Swick in Equation (78) for areas Waste Panel, South RoR, and North RoR, in Figure PA-8.

In addition, three variables are created in STEPWISE through transformation of the variable GLOBAL:OXSTAT, the indicator variable for oxidation states of uranium and plutonium. GLOBAL:OXSTAT is sampled as a [0,1) uniform distribution but is treated in the code as a Bernoulli distribution (a distribution having only two discrete states). The computed variable OXSTAT is assigned 0 if GLOBAL:OXSTAT is less than 0.5 and is assigned 1 otherwise. The other two computed variables represent the K_ds for the +VI and +IV oxidation states of uranium and plutonium, respectively. A K_d value represents the matrix:water partitioning coefficient. If GLOBAL:OXSTAT is 0 then CMKDU is assigned U+6:MKD_U and CMKDPU is assigned PU+4:MKD_PU. If GLOBAL:OXSTAT is 1 then CMKDU is assigned U+4:MKD_U and CMKDPU is assigned PU+3:MKD_PU, i.e. the K_ds for the +IV and +III oxidation states of uranium and plutonium, respectively. In the discussion below these variables are referenced as Composite:MKD_U and Composite:MKD_PU in order to denote their status as composites of pairs of sampled parameters.

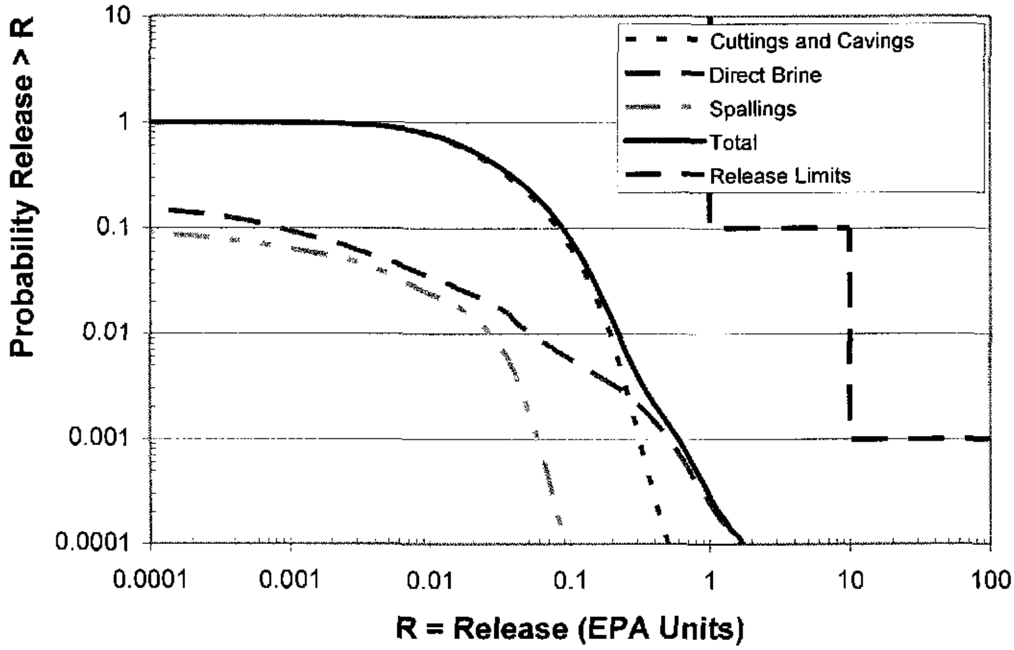
Changes from the CRA-2004 PABC

The sensitivity analysis performed using stepwise regression cannot be used to explain sensitivity of the results to the changes implemented in the models, parameters and inventory following the CRA-2004 PABC. The sensitivity analysis can only attempt to resolve the question of which *sampled* parameters contribute the most to the variability (uncertainty) observed in the mean releases for each vector.

The changes implemented in the CRA-2009 PA analysis from the CRA-2004 PABC analysis are documented in Clayton (2008a). Overall, these changes had little impact on total releases but did result in an upward shift of the CCDF for direct brine releases (DBR) (Figure 1). The changes included using a new version of BRAGFLO (version 6.0) for the CRA-2009 PA. Although this version contains many new capabilities, most of these were not used. Changes were made to the capillary pressure and relative permeability calculations, to implement a cut-off saturation to suspend chemical reactions, to control the initial conditions in the repository, and to modify the concentrations or materials as the result of intrusions. These changes are unlikely to contribute much to the change in DBR. A small reduction (0.3%) was made to the fraction of the repository volume occupied by waste which could reduce DBR releases slightly. There were errors in the input files to NUTS and ALGEBRACDB, including the assignment of the wrong intrusion time to a DBR release calculation. These errors were shown to have no significant impact on releases (Clayton 2007, Ismail 2007). There was also an increase in the inventory of CPR to correct for the omission of emplacement materials in the CRA-2004 PABC and as a result the higher CPR could slightly increase gas pressures in the repository.

Fifteen parameters were changed from their CRA-2004 PABC values and 90 new parameters were defined (Clayton 2008a), although most of the new parameters were associated with the new BRAGFLO capabilities and thus were not used. Six new parameters were also added to represent the density of cellulose, plastics and rubber (CPR) in the emplacement materials (Clayton 2008a). Two of the parameter changes undoubtedly have a direct impact on DBR. One change that should have led to lower releases was that the maximum duration of releases was reduced from 11 days to 4.5 days (Kirkes 2007). It is not expected that this would lead to a proportional change in DBRs because the typical pattern of release is to have an initial high flow rate followed by a lower flow rate, thus causing most of a release to occur early in the release event. Drilling rate was increased from 5.25E-3 to 5.85E-3 boreholes per km^2 per year and releases should increase proportional to this rate, or by about 11 %. Changes in these parameters would not be reflected in this sensitivity analysis because the parameters are not sampled by LHS; these changes would be reflected in a systematic change in the model results rather than a change in the uncertainty in the results.

A



B

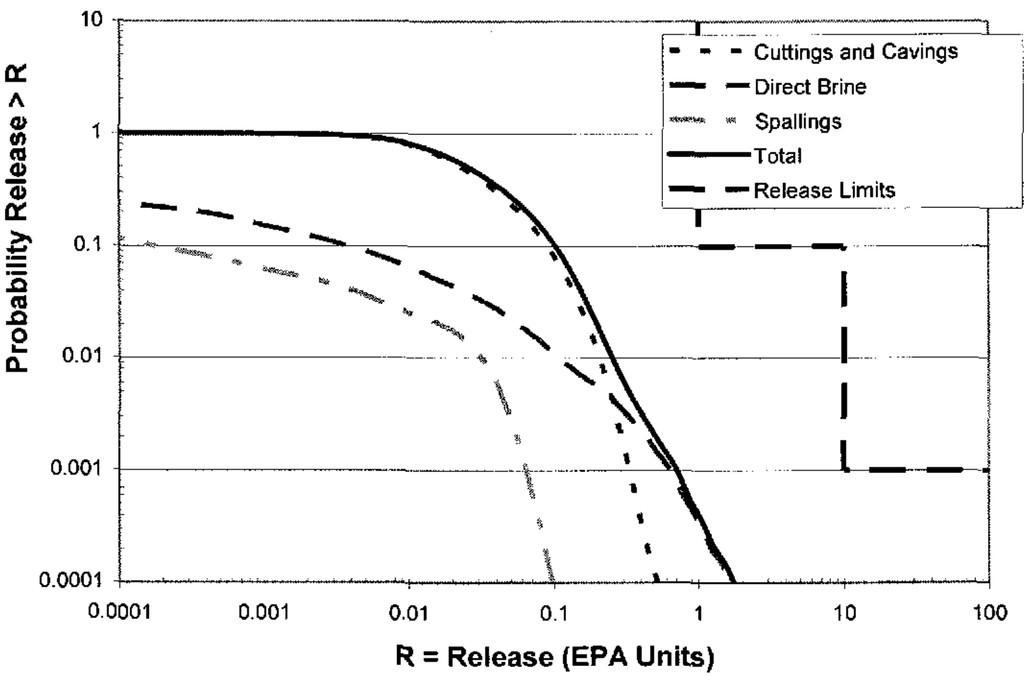


Figure 1. Mean CCDFs from the CRA-2004 PABC analysis (A) and CRA-2009 PA analysis (B).

It should be noted that only two sampled parameters were changed from the CRA-2004 PABC. The first sampled parameter, the maximum value of halite porosity (S_HALITE:POROSITY), was increased from 0.0300 to 0.0519. This change is thought to be the cause of an increase in DBR volumes as compared to the CRA-2004 PABC (Clayton 2008b). The second sampled parameter, the rate of CPR under anaerobic, humid conditions (WAS_AREA:GRATMICH) was previously made conditional to WAS_AREA:GRATMICI by setting WAS_AREA:GRATMICH to the value of WAS_AREA:GRATMICI if WAS_AREA:GRATMICH exceeded WAS_AREA:GRATMICI in any particular vector. For the CRA-2009 PA it was assumed that WAS_AREA:GRATMICH was uniformly distributed between 0 and the minimum of either 1.02717E-9 (the upper level of the uniform distribution specified in the parameter data base for the variable) and the value selected for WAS_AREA:GRATMICI. The CRA-2009 PA showed more non-zero DBR volumes than observed in the CRA-2004 PABC but also showed a decrease in the maximum DBR volume (Clayton 2008b).

Although these changes in the sampled parameters are expected to explain most of the difference in releases between the CRA-2004 PABC and the CRA-2009 PA, quantifying the relationships statistically is a challenge. Scatter plots of the change in DBR versus the change in the parameter values for the results paired by vector show no obvious patterns (Figs. 2 and 3). One possible reason for this is that these analyses are based on the mean releases for each vector, and the small releases that show the most change contribute little to the mean. Another possible explanation is that there are interactions between the parameters. Unfortunately, the sampling design used for evaluating uncertainty is not suited for quantifying interactions among the fifty-nine parameters; there would be 1711 first-order interactions alone to be considered but only 100 observations. Thus there would be more coefficients to fit in the full model than data points, leading to an indeterminate solution. Limiting the interactions to those with just the two sampled parameters that changed would still involve 118 first-order interactions.

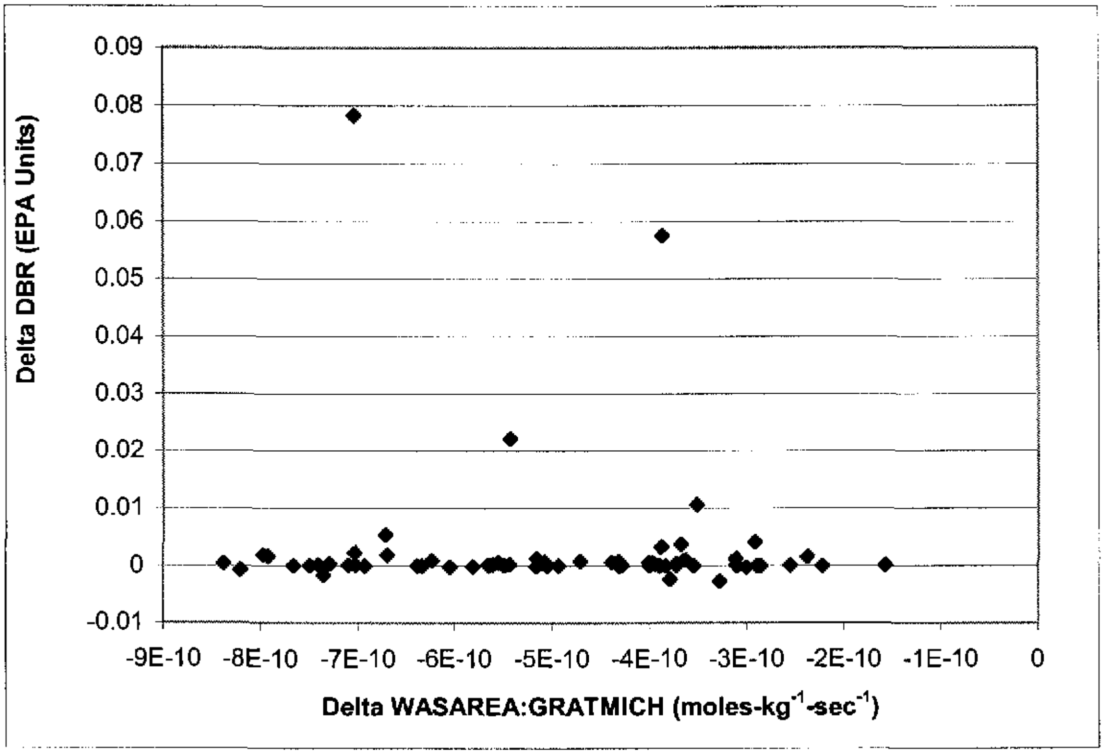


Figure 2. The scatter plot of the change in DBR versus the change in WASAREA:GRATMICH paired by vector.

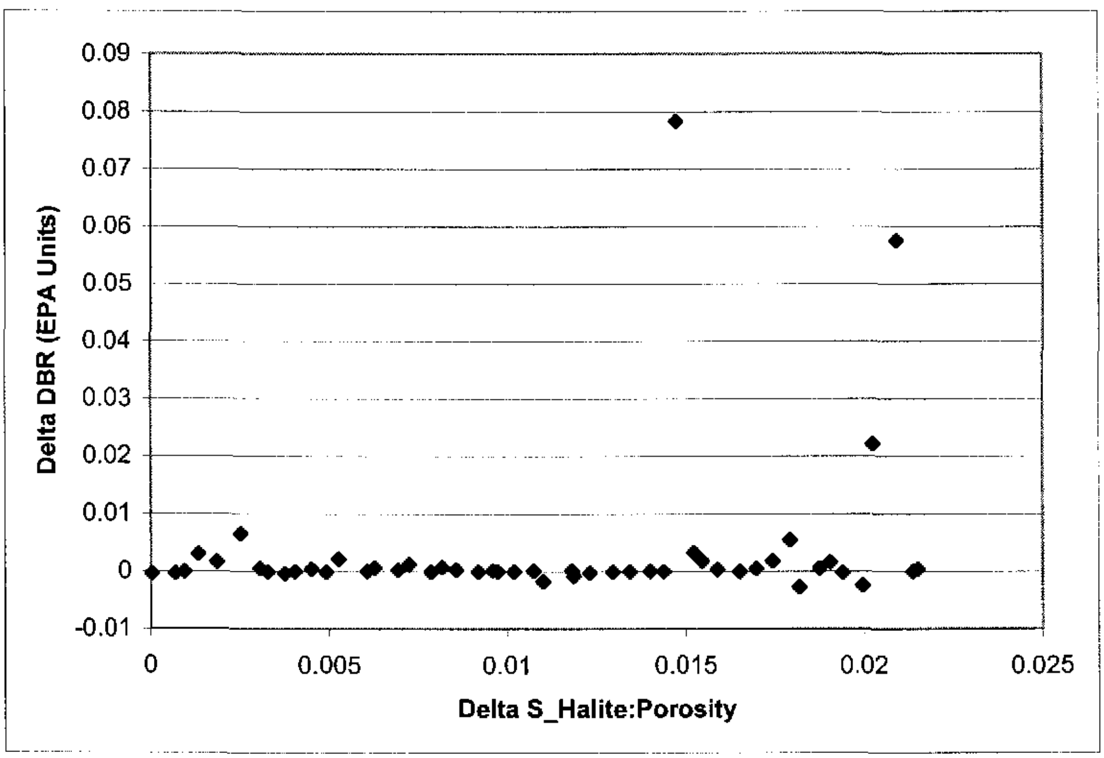


Figure 3. The scatter plot of the change in DBR versus the change in S_HALITE:POROSITY (a unitless parameter) paired by vector.

Total Releases

As in the CRA-2004 PABC (Kirchner 2005), cuttings and cavings, direct brine and spallings releases account for the majority of the total releases estimated in the CRA-2009 PA. In both analyses uncertainty in total normalized releases is largely due to uncertainty in waste shear strength (BOREHOLE:TAUFAIL). The volumes of cuttings and cavings are primarily controlled by shear strength, and the negative correlation found in the analysis is expected. SOLMOD3:SOLVAR remained the second most dominant parameter contributing to variability in Total releases in all replicates. SOLMOD3:SOLVAR is a "solubility multiplier" which represents uncertainty in solubilities for all actinides in the +III oxidation state (Xiong et al. 2005). Solubility of actinides impacts their concentration in direct brine releases (DBR). The variability in Total releases explained by BOREHOLE:TAUFAIL in the CRA-2009 PA dropped from previous levels. BOREHOLE:TAUFAIL only accounts for about 81% of the total variability in total releases in the CRA-2009 PA Replicate 1, whereas in the CRA-2004 PABC it accounted for 88% of the variability (Table 3). The R^2 value is called the coefficient of determination and represents the proportion of total variation explained by the fitted regression. The difference in the R^2 values between the CRA-2004 PABC and the CRA-2009 PA is small.

In replicate 1 of the CRA-2004 PABC the drill string angular velocity (BOREHOLE:DOMEGA), which is also used in computing cuttings and cavings, was ranked third in importance with the remaining parameters explain less than 1 % of the variability in the total releases. CELLULS:FBETA was ranked third in the CRA-2009 PA analysis with BOREHOLE:DOMEGA following it in fourth place. CELLULS:FBETA is a dimensionless scale factor used in definition of stoichiometric coefficient for microbial gas generation. It is assigned a uniform distribution from 0.0 to 1.0. In the CRA-2009 PA the remaining parameters contribute only about a percent or less each (Tables 3 and 4). CASTILER:PRESSURE is the initial brine pore pressure in the Castile. GLOBAL:PBRINE is the probability that a drilling intrusion penetrates pressurized brine in the Castile Formation. Therefore, CASTILER:PRESSURE and GLOBAL:PBRINE control the frequency with which Castile brine intrudes the repository due to a drilling event and the initial pressure of that brine thus influencing direct brine releases. Each of the variables ranked fifth or lower contributes only a few percent to the explanation of variance and their ordering should not be considered particularly informative.

Table 3. Stepwise ranked regression analysis for mean Total releases, replicate 1 of CRA-2009 PA and CRA-2004 PABC.

	Expected Normalized Release					
	CRA-2009 PA Replicate 1			CRA-2004 PABC Replicate 1		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	BOREHOLE:TAUFAIL	0.80703	-0.90051	BOREHOLE:TAUFAIL	0.88296	-0.94369
2	SOLMOD3:SOLVAR	0.83811	0.17323	SOLMOD3:SOLVAR	0.90538	0.14323
3	CELLULS:FBETA	0.85294	-0.11499	BOREHOLE:DOMEGA	0.91685	0.10497
4	BOREHOLE:DOMEGA	0.86422	0.10214	CELLULS:FBETA	0.92613	-0.09382
5	SHFTU:SAT_RGAS	0.87297	-0.08272	CASTILER:PRESSURE	0.93357	0.08050
6	CASTILER:PRESSURE	0.87981	0.07442	GLOBAL:PBRINE	0.93852	0.07062
7	GLOBAL:PBRINE	0.88536	0.08080	SHFTU:SAT_RGAS	0.94285	-0.06464
8	GLOBAL:TRANSIDX	0.89062	0.08172	SHFTL_T1:PRMX_LOG	0.94638	0.05951
9	S_MB139:RELP_MOD	0.89541	-0.07098			

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Table 4. Stepwise ranked regression analysis for mean Total releases, replicates 2 and 3 of CRA-2009 PA.

	Expected Normalized Release					
	CRA-2009 PA Replicate 2			CRA-2009 PA Replicate 3		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	BOREHOLE:TAUFAIL	0.78502	-0.88391	BOREHOLE:TAUFAIL	0.85669	-0.91852
2	SOLMOD3:SOLVAR	0.80538	0.14594	SOLMOD3:SOLVAR	0.87989	0.14666
3	CASTILER:PRESSURE	0.82772	0.14173	GLOBAL:PBRINE	0.89214	0.11080
4	CULEBRA:MINP_FAC	0.83829	-0.10057	CASTILER:PRESSURE	0.90322	0.10350
5	S_HALITE:POROSITY	0.84819	0.09468	BOREHOLE:DOMEGA	0.91319	0.10249
6	CONC_PCS:PRMX_LOG	0.85580	0.08758	S_HALITE:POROSITY	0.91743	0.06536
7	DRZ_1:PRMX_LOG	0.86291	0.08338			
8	CONC_PCS:PORE_DIS	0.86922	-0.08003			
9	SPALLMOD:REPIPERM	0.87485	-0.07397			
10	GLOBAL:PBRINE	0.88026	0.07374			

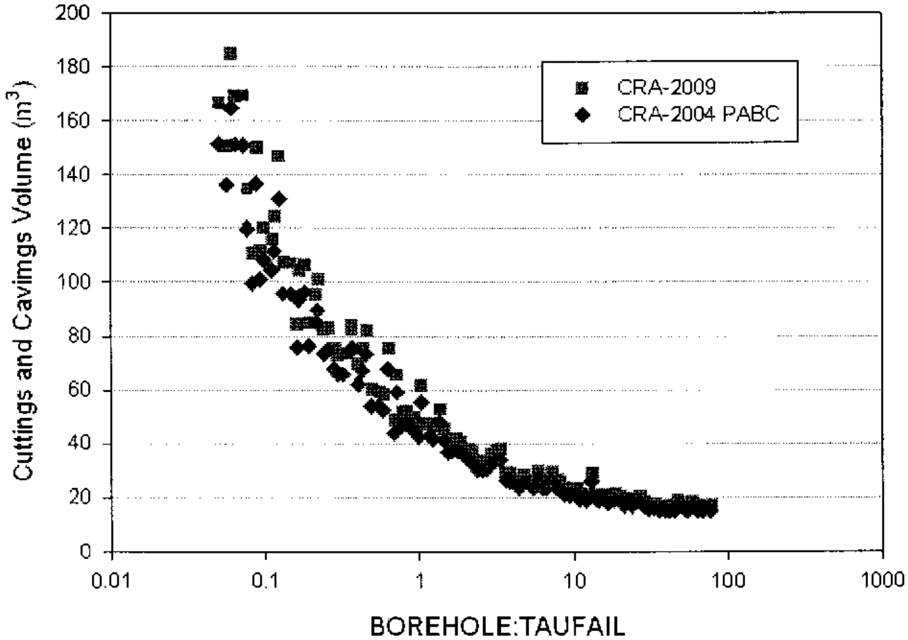
[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Cuttings and Cavings Releases

Tables 5 and 6 list the parameters that showed significant correlations to cuttings and cavings releases based on a stepwise regression using ranked data. The uncertainty in mean cuttings and cavings releases is primarily due to the uncertainty in the cuttings and cavings volume, as described in CRA-2004 Appendix PA (DOE, 2004, Figure PA-105). Waste shear strength (BOREHOLE:TAUFAIL) controls about 98 % of the variability in mean cuttings and cavings releases in replicate 1 of both the CRA-2004 PABC and the three CRA-2009 PA replicates. Cuttings and caving releases are primarily controlled by the volume of cuttings and cavings produced, which in turn is a highly non-linear function of BOREHOLE:TAUFAIL (Fig. 4). The drill string angular velocity

(BOREHOLE:DOMEGA) has a very minor contribution as well, and is discussed in Dunagan (2004) and Ismail (2008). The remaining parameters in Tables 5 and 6 explain less than about 0.2% of the variability in cuttings and cavings. Correlations that explain only a few percent of the variation can occur due to random sampling and may well be spurious.

A.



B.

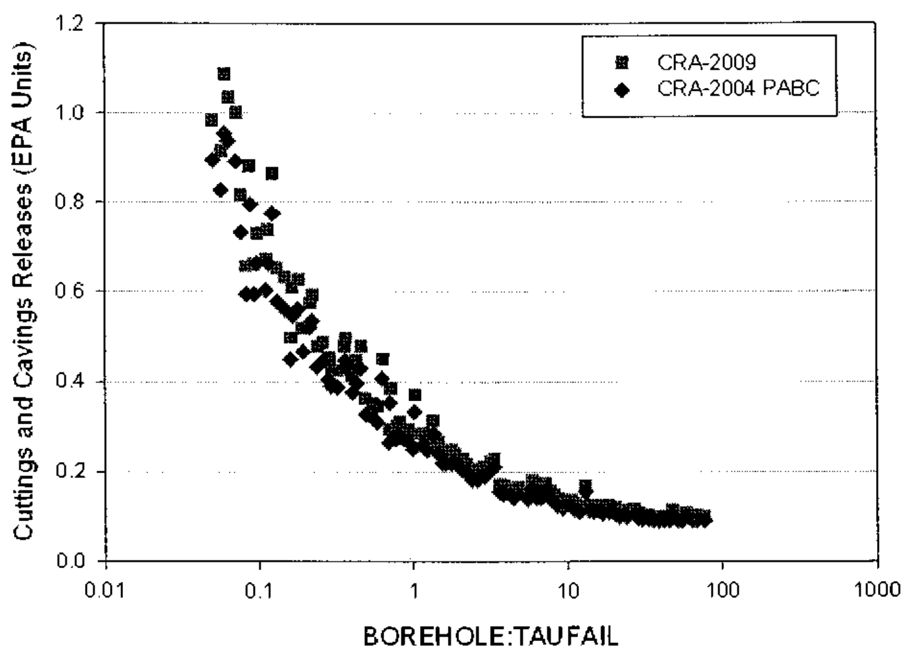


Figure 4. The volume of cuttings and cavings (A) and releases from cuttings and cavings (B) show significant nonlinear responses to WTAUFAIL. Note that the response is a curve even though BOREHOLE:TAUFAIL is plotted using a logarithmic axis.

Table 5. Stepwise ranked regression analysis for mean Cuttings and Cavings releases.

	Expected Normalized Release					
	CRA-2009 PA			CRA-2004 PABC		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	BOREHOLE:TAUFAIL	0.98113	-0.98784	BOREHOLE:TAUFAIL	0.98182	-0.98930
2	BOREHOLE:DOMEGA	0.99363	0.11414	BOREHOLE:DOMEGA	0.99314	0.10743
3	SHFTL_T1:PRMX_LOG	0.99401	0.01952	(Composite):OXSTAT	0.99363	-0.01943
4	DRZ_PCS:PRMX_LOG	0.99433	0.01773	SHFTL_T1:PRMX_LOG	0.99402	0.01934
5	CASTILER:PRESSURE	0.99458	-0.01571	CULEBRA:HMBLKLT	0.99434	0.01724
6	GLOBAL:CLIMTIDX	0.99482	-0.01539	DRZ_PCS:PRMX_LOG	0.99458	0.01550
7	CULEBRA:HMBLKLT	0.99505	0.01522			

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Table 6. Stepwise ranked regression analysis for mean Cuttings and Cavings releases, replicates 2 and 3 of CRA-2009 PA.

	Expected Normalized Release					
	CRA-2009 PA Replicate 2			CRA-2009 PA Replicate 3		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	BOREHOLE:TAUFAIL	0.98207	-0.99260	BOREHOLE:TAUFAIL	0.97896	-0.98438
2	BOREHOLE:DOMEGA	0.99280	0.10370	BOREHOLE:DOMEGA	0.99292	0.11829
3	SHFTL_T1:PRMX_LOG	0.99340	-0.02452	SOLMOD3:SOLVAR	0.99339	-0.02119
4				CULEBRA:APOROS	0.99375	0.01922

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Spallings Releases

Compared to the CRA-2004 PABC the CRA-2009 PA had higher probabilities of releases of activity and spalling volumes across much of the range of releases (Figs. 5 and 6). Thirty-four percent of the replicate 1 vectors, 41 % of the replicate 2 vectors and 36 % of the replicate 3 vectors in the CRA-2009 PA showed spallings releases. In comparison, the CRA-2004 PABC had 34 %, 37 %, and 31 % of the vectors in the three replicates showing no releases, so there was a small overall increase in the number of spalling events. These differences probably reflect somewhat higher pressures or larger volumes of gas flow that resulted from the increase in CPR and the increase in maximum halite porosity in the CRA-2009 PA analysis (Ismail 2008). Nevertheless, the number of zero-releases was high enough to reduce the effectiveness of the regression analysis.

Table 6 lists the parameters that showed correlation to mean spallings releases after a stepwise ranked regression using data from replicate 1. The first five parameters from the analysis of the first in both the CRA-2004 PABC and CRA-2009 PA analyses are the same although the rank order is different. However, the third-ranked parameter, CULEBRA:DPOROS, is found only in the analysis of the first replicate. Thus little confidence should be placed on inferences drawn about the impact of parameters ranked third or greater.

The dominant parameter with regard to controlling spallings releases in the CRA-2009 PA assessment is S_HALITE:POROSITY, the effective porosity in intact halite. Its higher ranking in the CRA-2009 PA analysis compared to the CRA-2004 PABC analysis may be due to the increase in the maximum value of its distribution. The analysis of the non-zero releases shows much similarity between the scatter plots of the ranks of spallings releases versus S_HALITE:POROSITY (Figure 7; from spreadsheet SpallingsVsHalitePorosityPABC_CRA09.xls). Nevertheless, the regression on the CRA-2009 PA data is significant (p=0.018) whereas the regression on the CRA-2004 PABC data is not (p=0.23). The positive correlation is likely to be due to having greater gas pressures under higher porosities due to greater brine flow into the repository. SPALLMOD:PARTDIAM is the particle diameter for disaggregated waste. A negative correlation is observed in regressions whether releases of zero are included (Tables 6 and 7) or omitted (p = 0.104, Fig. 8; from spreadsheet SpallVsSpallmod_partDiam.xls). The negative correlation with SPALLMOD:PARTDIAM is probably due to the tendency to have greater fluidization at smaller particle diameters.

There is little consistency among the replicates in the lists of the variables of step 3 or higher. These variables contribute only a few percent to the total variability and the ranking of their importance to determining spallings releases is thus questionable. For example, CULEBRA:DPOROS is the diffusive porosity of the Culebra dolomite. Taking only the ranks of the non-zero releases and the corresponding ranks of the spallings releases shows no correlation (p=0.41, Fig. 9; spreadsheet CRA2009_AP137_Cul_DPOROSvsSpall_REL.xls). CASTILER:PRESSURE is the far-field pore pressure in the Castile brine reserve. S_MB139:PRMX_LOG is the intrinsic permeability of the Salado marker bed. DRZ_PCS:PRMX_LOG is the logarithm of intrinsic permeability (m^2) of the DRZ immediately above the panel closure concrete. SHFTL_T2:PRMX_LOG is the logarithm of intrinsic permeability (m^2) of lower shaft seal materials from 200 years to 10,000 years after closure. WAS_AREA:PROBDEG is the index for microbial degradation of CPR materials (dimensionless). While these parameters can perhaps influence pressures in the repository, and hence spallings releases, the lack of consistency across the replicates should be interpreted as evidence against their importance in controlling spallings.

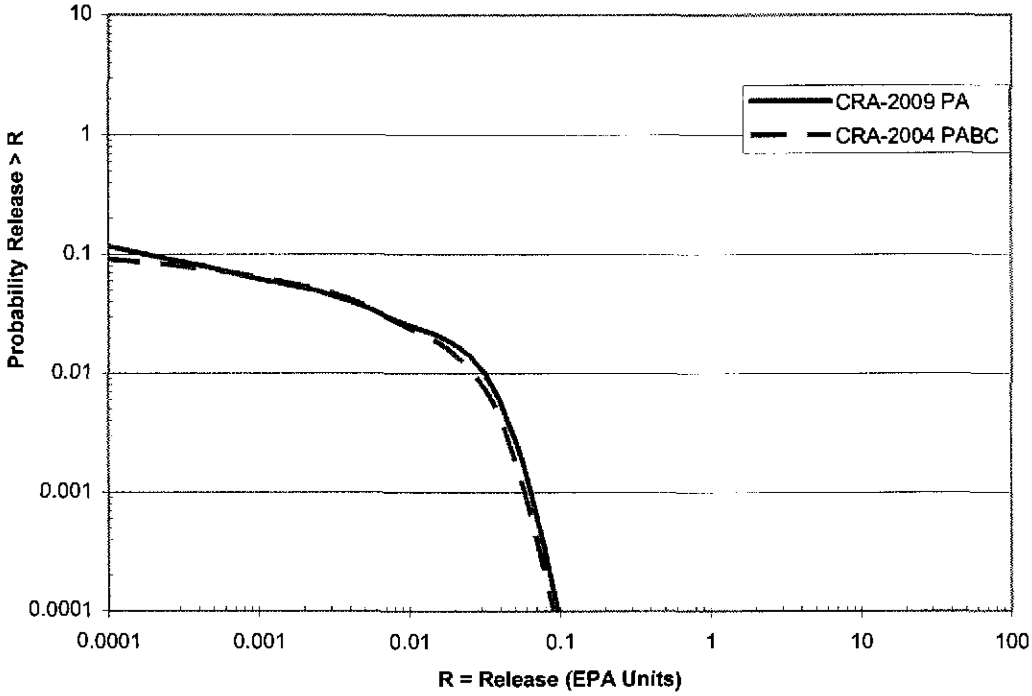


Figure 5. Spalling releases from the CRA-2004 PABC and CRA 2009 analyses.

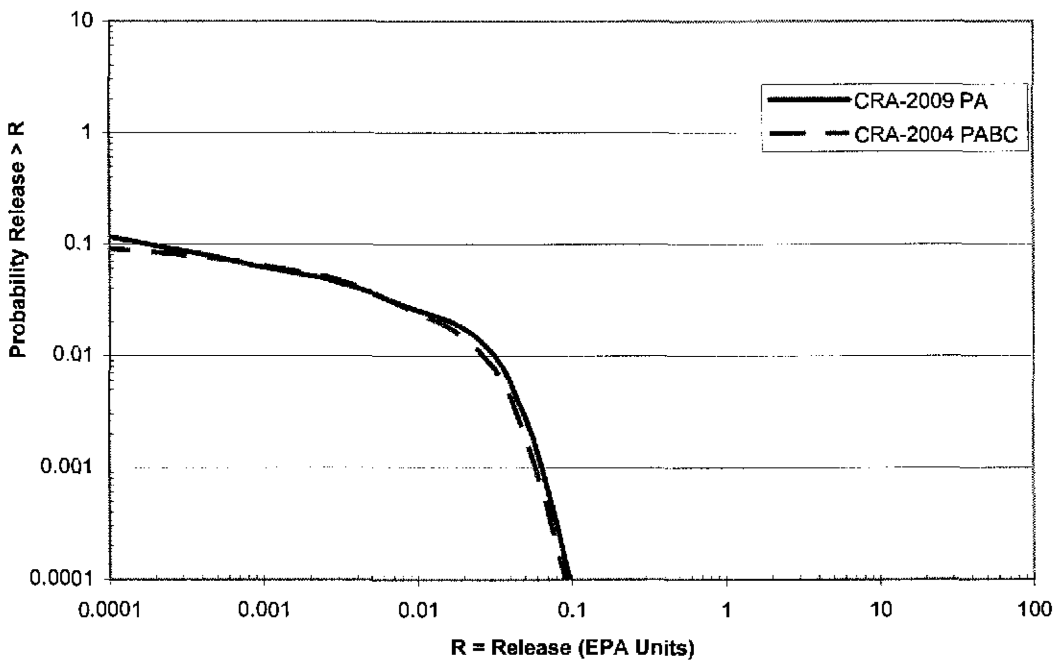


Figure 6. Spalling volume releases from the CRA-2004 PABC and CRA 2009 analyses.

Table 6. Stepwise ranked regression analysis for mean Spallings releases, replicate 1.

	Expected Normalized Release					
	CRA-2009 PA			CRA-2004 PABC		
Step[a]	Variable	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	S_HALITE:POROSITY	0.12435	0.31752	SPALLMOD:PARTDIAM	0.11478	-0.31336
2	SPALLMOD:PARTDIAM	0.21539	-0.29383	S_HALITE:POROSITY	0.16750	0.22391
3	CULEBRA:DPOROS	0.26409	-0.23128	CULEBRA:DPOROS	0.21204	-0.22767
4	CASTILER:PRESSURE	0.29852	0.18369	S_MB139:PRMX_LOG	0.25605	0.20906
5	S_MB139:PRMX_LOG	0.33061	0.17654	CASTILER:PRESSURE	0.28990	0.19710
6	DRZ_PCS:PRMX_LOG	0.36146	-0.16959	WAS_AREA:BIOGENFC	0.32333	0.17681
7	SHFTL_T2:PRMX_LOG	0.38844	-0.17081	PU+3:MKD_PU	0.35176	0.17213
8	WAS_AREA:PROBDEG	0.41571	0.16758	BH_SAND:PRMX_LOG	0.38021	-0.16926

a Steps in stepwise regression analysis
b Variables listed in order of selection
c Cumulative R^2 value with entry of each variable into regression model
d Standardized Rank Regression Coefficient

Table 7. Stepwise ranked regression analysis for mean Spallings releases, replicates 2 and 3 of CRA-2009 PA.

	Expected Normalized Release					
	CRA-2009 PA Replicate 2			CRA-2009 PA Replicate 3		
Step[a]	Variable	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	S_HALITE:POROSITY	0.15897	0.41243	S_HALITE:POROSITY	0.21810	0.45662
2	SPALLMOD:PARTDIAM	0.26504	-0.33108	SPALLMOD:PARTDIAM	0.28191	-0.22810
3	CASTILER:PRESSURE	0.30913	0.20513	PHUMOX3:PHUMCIM	0.31567	-0.17847
4	BH_SAND:PRMX_LOG	0.34709	-0.19489	WAS_AREA:GRATMICI	0.35023	0.20958
5				WAS_AREA:SAT_RGAS	0.38244	-0.15805
6				CASTILER:PRESSURE	0.40971	0.19749
7				GLOBAL:CLIMTIDX	0.43803	0.17868
8				CONC_PCS:PRMX_LOG	0.46365	-0.15955
9				WAS_AREA:PROBDEG	0.48865	0.16418

a Steps in stepwise regression analysis
b Variables listed in order of selection
c Cumulative R^2 value with entry of each variable into regression model
d Standardized Rank Regression Coefficient

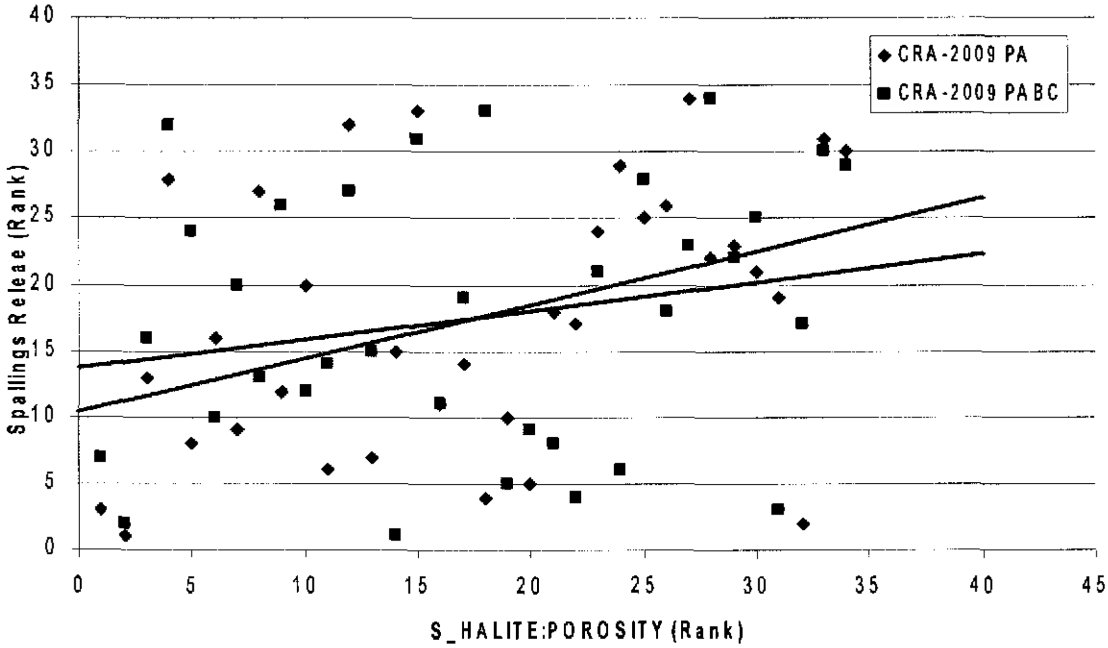


Figure 7. Correlation between non-zero spallings releases and S_HALITE:POROSITY from the CRA-2009 PA and CRA-2004 PABC.

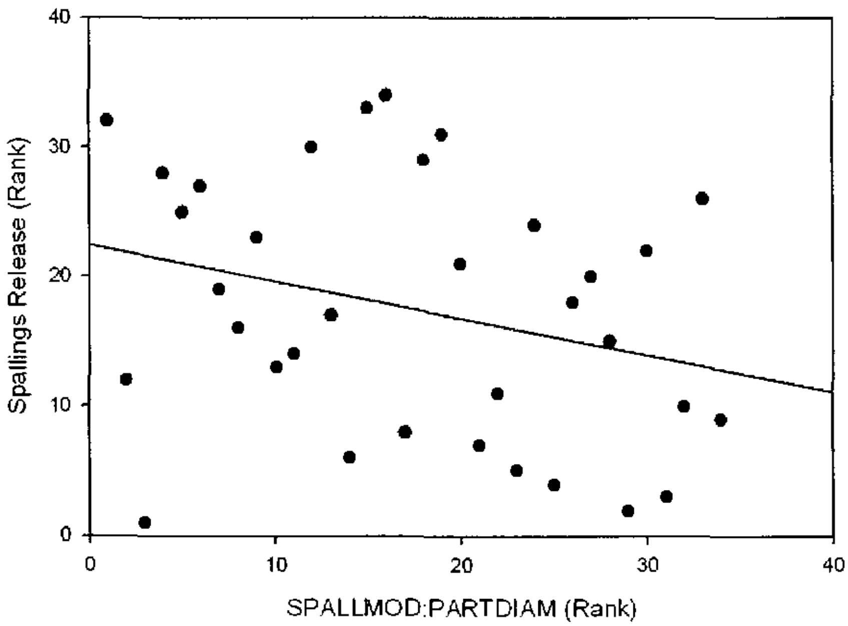


Figure 8. Non-zero spallings release shows a negative correlation (p = 0.104) with SPALLMOD:PARTDIAM.

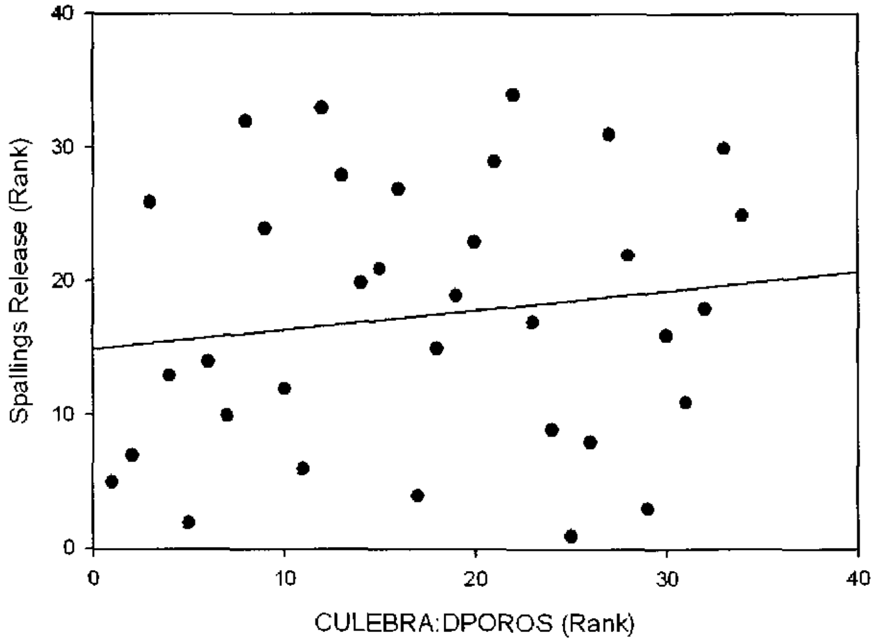


Figure 9. Ranks of the non-zero spallings releases for replicate 1 show no correlation (p = 0.41) with the ranks of CULEBRA:DPOROS.

Direct Brine Release

Direct brine releases (DBR) are releases of contaminated brine originating in the repository and flowing up an intrusion borehole during the period of drilling and before the hole is plugged. In order for DBR to occur volume-averaged pressure near the borehole must exceed 8 MPa and brine saturation in the repository must exceed the residual saturation of the waste material (Clayton 2008b). The CRA-2009 PA analysis shows that four variables (SOLMOD3:SOLVAR, CASTILER:PRESSURE, STEEL:CORRMCO2 and GLOBAL:PBRINE) account for more than 50 % of the uncertainty in DBR (Tables 8 and 9). These variables are also important in the CRA-2004 PABC analysis although the third- and fourth-ranked variables are in reverse order relative to the CRA-2009 PA replicates. SOLMOD3:SOLVAR is a "solubility multiplier" that represents uncertainty in solubilities for all actinides in the +III oxidation state (Xiong et al. 2005). STEEL:CORRMCO2 is the inundated corrosion rate for steel in the absence of CO_2. The corrosion of iron is expected to produce hydrogen but at the same time it consumes water. When the repository is flooded with brine from the intrusion of a brine pocket it is likely that the influence of STEEL:CORRMCO2 on DBR would be positive since the production of hydrogen would outweigh the minimal impact of the consumption of water. However, a negative correlation is observed between the ranked variables (Fig. 10 and Table 8), suggesting that the corrosion of steel is having its strongest influence when the repository is not saturated and DBR releases are expected to be small. CASTILER:PRESSURE is the initial brine pore pressure in the Castile. CASTILER:PRESSURE and GLOBAL:PBRINE control the frequency with which Castile brine intrudes the repository due to a drilling event and the initial pressure of that brine, thus their positive correlation with DBR is obvious.

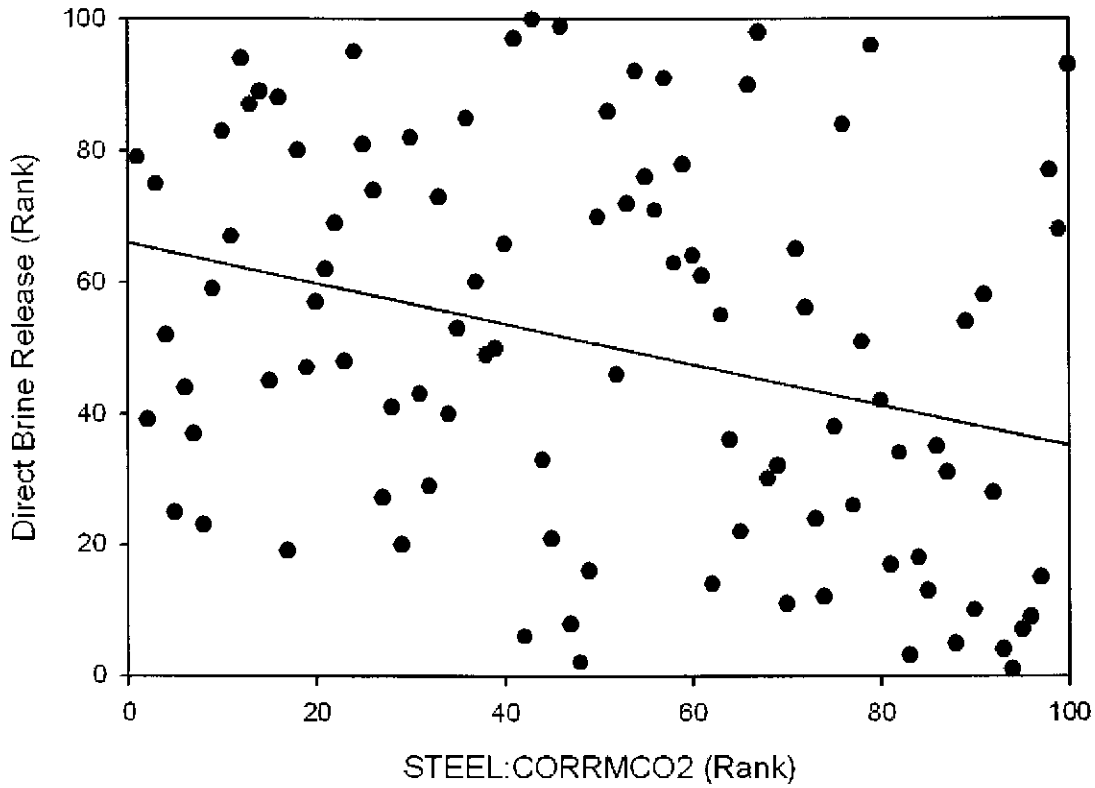


Figure 10. A significant (p = 0.0018) negative correlation is observed between the ranks of Direct Brine Releases and the corrosion rate of steel (STEEL:CORRMCO2).

There is no consistency in the names and orderings of the remaining parameters contributing to the uncertainty in DBR. These parameters each contribute about 1-4 % of the variability in DBR. Once the contribution of a parameter to the uncertainty drops to a few percent the correlation could arise from sampling rather than a mechanistic relationship. That is not to say that all of these correlations are spurious. For example, WAS_AREA:SAT_WICK is the parameter that represents the increase in brine saturation due to capillary forces. The negative correlation is the result of an increase in iron consumption and thereby a reduction of brine as WAS_AREA:SAT_WICK increases. However, although WAS_AREA:SAT_WICK was ranked fifth or sixth in importance in the CRA-2004 PABC analysis and in two of the three replicates of this analysis, it was not identified as a contributor to uncertainty in DBR in the third replicate. Thus, although it is logical for WAS_AREA:SAT_WICK to exert some control on DBR and it can be shown to contribute to the uncertainty in DBR in some replicate samples, its influence is small and can be undetectable in other samples.

Table 8. Stepwise ranked regression analysis for mean Direct Brine releases.

	Expected Normalized Release					
	CRA-2009 PA			CRA-2004 PABC		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	SOLMOD3:SOLVAR	0.22767	0.45502	SOLMOD3:SOLVAR	0.23843	0.46604
2	CASTILER:PRESSURE	0.36876	0.38668	CASTILER:PRESSURE	0.39692	0.40193
3	STEEL:CORRMCO2	0.48679	-0.32646	GLOBAL:PBRINE	0.50630	0.32009
4	GLOBAL:PBRINE	0.57557	0.29956	STEEL:CORRMCO2	0.60015	-0.29065
5	S_MB139:RELP_MOD	0.60706	-0.18146	BH_SAND:PRMX_LOG	0.63488	-0.18391
6	WAS_AREA:SAT_WICK	0.63227	-0.15672	WAS_AREA:SAT_WICK	0.66502	-0.16704
7	S_MB139:PRMX_LOG	0.65414	0.13950	DRZ_PCS:PRMX_LOG	0.68997	0.15075
8	WAS_AREA:BIOGENFC	0.67239	0.13434	S_MB139:RELP_MOD	0.71125	-0.16775
9	CONC_PCS:SAT_RGAS	0.68727	-0.12225	S_MB139:PRMX_LOG	0.72506	0.11787
10				S_HALITE:COMP_RCK	0.73878	-0.11038
11				CONC_PCS:SAT_RGAS	0.75012	-0.10979
12				PHUMOX3:PHUMCIM	0.76174	0.10862

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Table 9. Stepwise ranked regression analysis for mean Direct Brine releases, replicates 2 and 3 of CRA-2009 PA.

	Expected Normalized Release					
	CRA-2009 PA Replicate 2			CRA-2009 PA Replicate 3		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	SOLMOD3:SOLVAR	0.23180	0.50565	SOLMOD3:SOLVAR	0.25647	0.48751
2	CASTILER:PRESSURE	0.39685	0.34590	CASTILER:PRESSURE	0.46183	0.43240
3	STEEL:CORRMCO2	0.51867	-0.37838	STEEL:CORRMCO2	0.54393	-0.27362
4	GLOBAL:PBRINE	0.59555	0.27108	GLOBAL:PBRINE	0.61169	0.25813
5	WAS_AREA:SAT_WICK	0.63686	-0.18369	DRZ_PCS:PRMX_LOG	0.64418	0.17376
6	WAS_AREA:PROBDEG	0.66803	-0.20720	CULEBRA:MINP_FAC	0.66908	0.16748
7	S_MB139:PRMX_LOG	0.70087	0.19048	WAS_AREA:SAT_RBRN	0.69219	-0.16372
8	DRZ_PCS:PRMX_LOG	0.72778	0.16471	CASTILER:COMP_RCK	0.71418	-0.15355
9	BH_SAND:PRMX_LOG	0.74580	-0.13565	SPALLMOD:REPIPOR	0.73368	0.14206
10	WAS_AREA:SAT_RBRN	0.76417	-0.13700	WAS_AREA:PROBDEG	0.75280	-0.14492
11	S_HALITE:POROSITY	0.78192	0.14266	SHFTL_T2:PRMX_LOG	0.76493	-0.11074
12	SPALLMOD:REPIPOR	0.79644	0.11723			
13	WAS_AREA:GRATMICI	0.80877	-0.11598			
14	S_MB139:SAT_RBRN	0.81853	0.09399			
15	TH+4:MKD_TH	0.82705	0.09373			

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Although the sensitivity analyses for the DBR for the CRA-2004 PABC and the CRA-2009 PA were very similar, there were some significant differences in the magnitudes of the DBR releases. These differences were most prominent for small releases (Fig. 11).

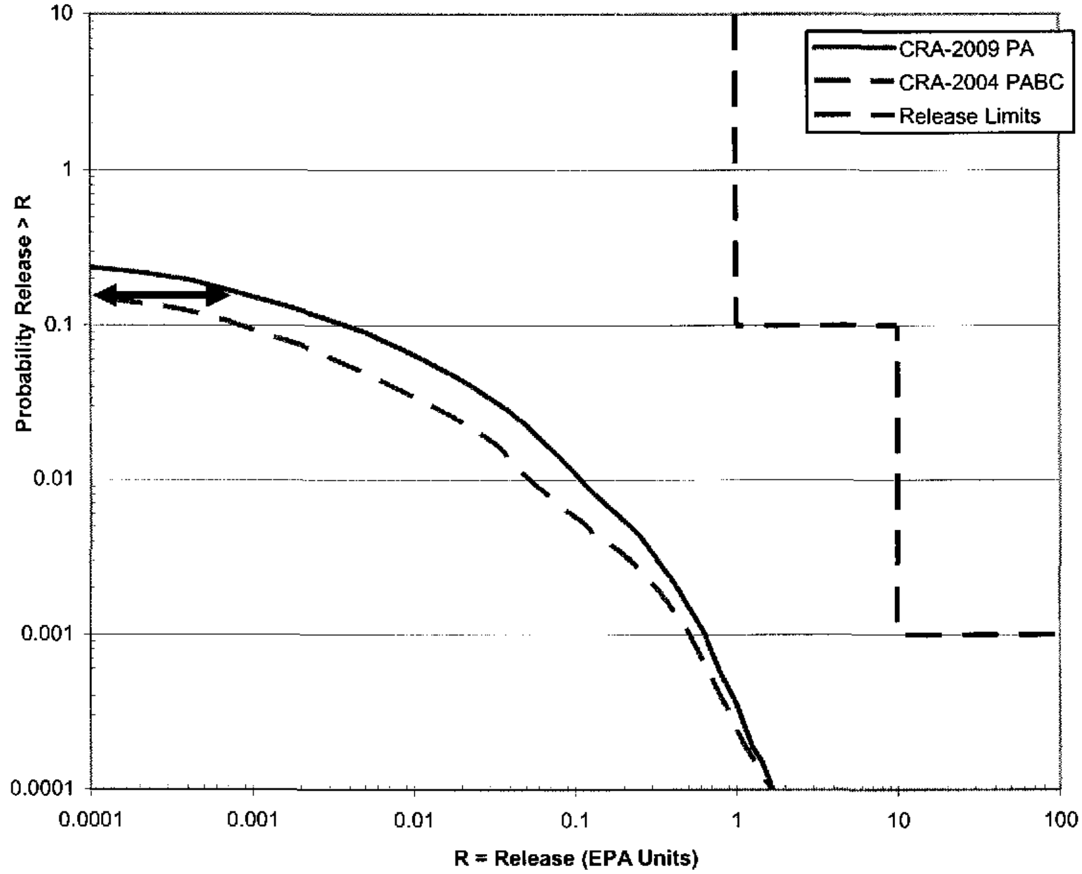


Figure 11. DBR in the CRA-2009 PA analysis exceeded DBR from the CRA-2004 PABC analysis by about a factor of 10 at a probability of about 0.15.

Culebra Releases

The analysis of the sensitivity of Culebra releases to the input parameters using linear regression is problematic. In the CRA-2004 PABC and the CRA-2009 PA 84 % of the vectors for replicate 1, and 83 % across all three replicates, had Culebra releases of zero. The releases of zero are found across the entire range of every parameter. This is undoubtedly due, for the most part, to transport rates frequently being too small to enable contaminants to reach the boundary within the simulation period, 10,000 years. Thus the release data are strongly censored. The times of the intrusions giving rise to flows to the Culebra are also likely to influence whether or not such releases occur. These times are not represented in the "sampled" input parameters and thus cannot be associated with the releases. In addition, the preponderance of zero values tends to negate the assumption of linear regression that errors (residuals) are normally distributed. In many cases it appears that it is the distribution of zeros along the independent axis that determines whether a positive or negative correlation is observed (e.g. Figure 12). Because of these issues, the linear ranked regression analysis is unlikely to yield a definitive identification of the sensitivity of Culebra releases to the sampled parameters. Most of the variability in Culebra releases remains unexplained by the regression model (Tables 10 and 11).

Table 10. Stepwise ranked regression analysis for mean Culebra releases.

	Expected Normalized Release					
	CRA-2009 PA			CRA-2004 PABC		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	(Composite):MKD_U	0.11631	-0.54630	(Composite):MKD_U	0.11743	-0.67499
2	CULEBRA:APOROS	0.16571	-0.22797	CULEBRA:APOROS	0.16497	-0.24101
3	GLOBAL:CLIMTIDX	0.20527	0.19676	GLOBAL:CLIMTIDX	0.20536	0.19271
4	TH+4:MKD_TH	0.24520	-0.23734	TH+4:MKD_TH	0.24594	-0.25632
5	(Composite):MKD_PU	0.28256	-0.26966	(Composite):MKD_PU	0.28274	-0.36378
6	STEEL:CORRMCO2	0.31946	-0.18489	U+4:MKD_U	0.32524	0.22376
7	SHFTU:SAT_RGAS	0.35152	-0.17254	STEEL:CORRMCO2	0.36337	-0.18764
8	BOREHOLE:TAUFAIL	0.38173	0.18433	SHFTU:SAT_RGAS	0.39699	-0.17682
9	BH_SAND:PRMX_LOG	0.40972	0.18129	BOREHOLE:TAUFAIL	0.42776	0.18535
10	WAS_AREA:GRATMICI	0.43864	0.17140	BH_SAND:PRMX_LOG	0.45208	0.16974
11				WAS_AREA:GRATMICI	0.47953	0.16705

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Table 11. Stepwise ranked regression analysis for mean Culebra releases, replicates 2 and 3 of CRA-2009 PA.

	Expected Normalized Release					
	CRA-2009 PA Replicate 2			CRA-2009 PA Replicate 3		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable	R^2	SRRC
1	(Composite):MKD_U	0.15714	-0.35236	(Composite):MKD_U	0.10435	-0.77531
2	CULEBRA:MINP_FAC	0.22823	-0.24288	(Composite):OXSTAT	0.16550	-0.51265
3	CASTILER:PRMX_LOG	0.28678	-0.24058	SOLMOD4:SOLVAR	0.20723	0.20074
4	CULEBRA:DPOROS	0.33109	-0.21516	CULEBRA:MINP_FAC	0.24674	-0.20027
5	WAS_AREA:BIOGENFC	0.36721	-0.19029	DRZ_PCS:PRMX_LOG	0.28497	-0.19055
6				CONC_PLG:PRMX_LOG	0.31582	0.17624

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

In a stepwise regression analysis restricted to the 16 vectors of the CRA-2004 PABC having non-zero releases there were no significant correlations found with any of the parameters (Kirchner 2005). The releases of the CRA-2009 PA analysis are nearly identical to the CRA-2004 PABC releases (Fig. 13; spreadsheet CulRel2009AndPABCVsCul_aporos.xls) and would undoubtedly show the same lack of correlation. Only one vector, 57, showed a significant change from the CRA-2004 PABC and that appears to be due to a change from having one release event to two release events out of 10,000 trials. Restricting the analysis to the non-zero results is not entirely a fair treatment of the data because releases of zero are due to the associated sampled values for parameters that affect the transport rates, such as K_ds. The following analysis examines the sensitivities both including and excluding the vectors having releases of zeros. Single-parameter rank regression analyses of the non-zero releases were carried out using Excel in order to evaluate the probability of significance of those relationships.

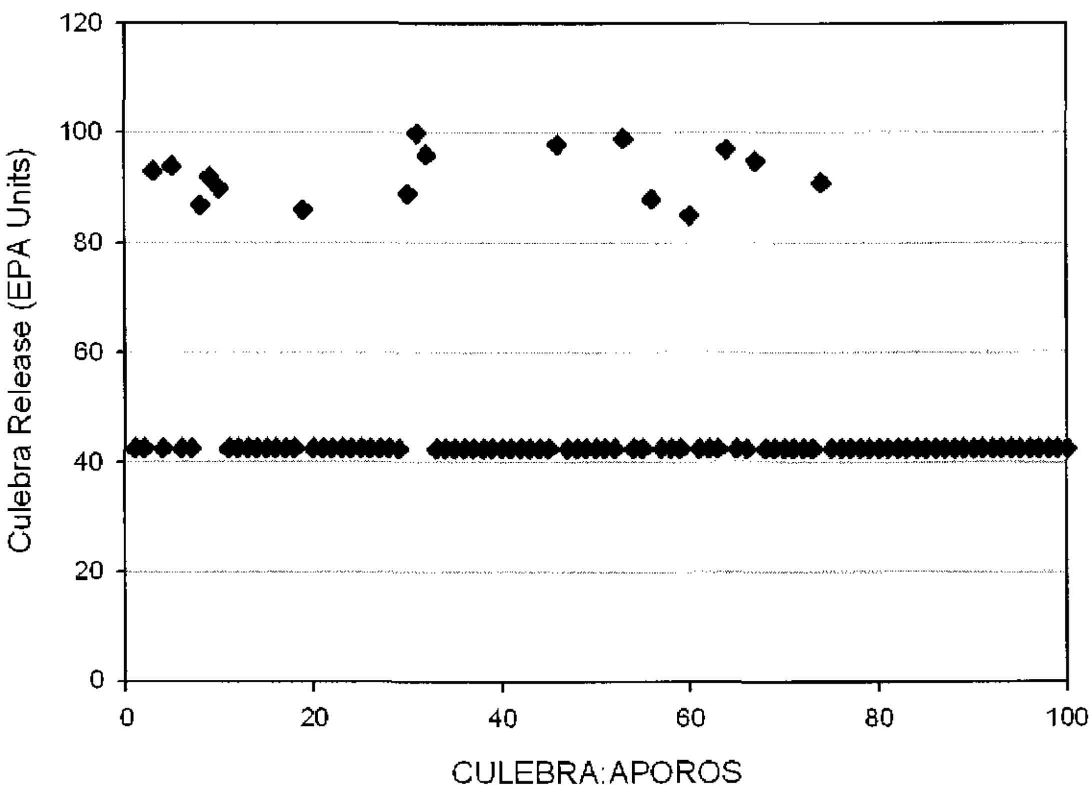


Figure 12. The preponderance and distribution of zero releases (average rank = 42) can control the regression, as in the case of the ranks of CULEBRA:APOROS (a unitless parameter) and Culebra Release.

The results from the CRA-2004 PABC and the CRA-2009 PA are nearly identical though the topmost five parameters (Table 9). The Culebra releases appear to be sensitive to the matrix/water partition coefficients (K_ds) for uranium, thorium and plutonium, to the Culebra advective porosity (the fracture volume per unit volume of porous media, CULEBRA:APOROS), and to GLOBAL:CLIMTIDX, the climate scale factor for the

Culebra flow field. The climate scale factor accounts for uncertainty in the climate that could result in increased precipitation. These correlations make sense logically because the parameters help control the rate of movement of the metals through the groundwater system.

The dominant parameter in replicate 1 of the CRA-2004 PABC and all three replicates of the CRA-2009 PA is Composite:MKD_U, the matrix partition coefficient, K_d, for uranium. Composite:MKD_U is assigned U+6:MKD_U or U+4:MKD_U depending on the oxidation state for uranium. These are the K_ds for the +VI and +IV oxidation states of uranium, respectively. The negative correlation between Culebra releases and Composite:MKD_U was expected because high values of K_d suppress the transport of the radionuclide. However, this correlation may nevertheless be an artifact related to the low frequency of non-zero results. Neither of the ranked correlations between only the non-zero Culebra releases and the parameters U+4:MKD_U ($p = 0.29$; spreadsheet CRA2009_AP137_CMKDU4VsCulRel.xls) and U+6:MKD_U ($p = 0.31$; spreadsheet CRA2009_AP137_CMKDU6VsCulRel.xls) were significant. The K_ds for thorium +IV and plutonium (also a composite variable) are ranked fourth and fifth, respectively, but their correlations are probably spurious as well.

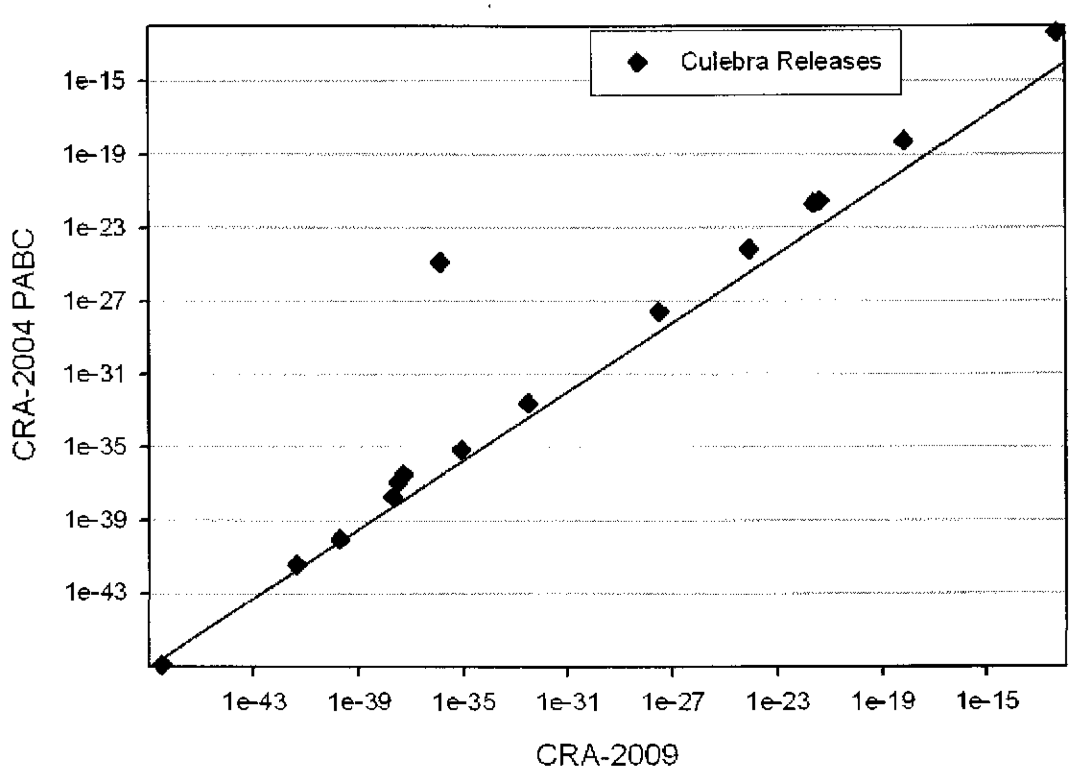


Figure 13. Culebra releases (EPA Units) from the CRA-2004 PABC and the CRA-2009 PA analyses are nearly identical. The line marks the location of equality.

CULEBRA:APOROS is the second most important parameter in both the CRA-2004 PABC and the CRA-2009 PA. Positive correlations are expected for this variable, thus the negative correlation between Culebra releases and CULEBRA:APOROS is counterintuitive. The negative correlation results from the preponderance of zero values in the results as illustrated in Fig. 12. A rank regression on the non-zero values only shows a non-significant ($p = 0.57$; spreadsheet CulRel2009VsCul_aporos.xls) but positive correlation between Culebra releases and CULEBRA:APOROS.

Because most release events do not reach the boundary of the WIPP site within 10,000 years, it can be helpful to look at the releases to the Culebra as well. The dominant parameter in the CRA-2009 PA analysis of releases to the Culebra is BH_SAND:PRMX_LOG, the logarithm of intrinsic permeability in the X-direction for a sand-filled borehole (Tables 12 and 13). Conceptually, the flow of brine up the borehole (and thus to the Culebra) should be positively influenced by increasing values for BH_SAND:PRMX_LOG (Stein and Zelinski 2003). However, only replicate 1 showed a significant rank correlation between BH_SAND:PRMX_LOG and the releases from the Culebra (Tables 10 and 11). BH_SAND:PRMX_LOG was ranked ninth and contributed only about 3% to the uncertainty in releases from the Culebra (Table 10). There is also no correlation between BH_SAND:PRMX_LOG and the releases to the Culebra ($p = 0.90$; spreadsheet CRA2009_AP137_CulRelVsBHPERM.xls) when only the non-zero release data for those replicates are analyzed.

In summary, in considering the processes that transport radionuclides to the Culebra from the repository and those that transport the radionuclides within the Culebra to the boundary of the WIPP site, BH_SAND:PRMX_LOG and the K_ds for the radionuclides are expected to be important factors controlling Culebra releases. However, a large majority of the vectors result in no releases being projected, and those that result in releases do not show the expected correlations. When the vectors resulting in releases of zero are included in the analyses, some significant correlations with parameters are found, including the K_ds for uranium but including BH_SAND:PRMX_LOG in only one replicate. However, a regression model is ill suited for evaluating sensitivities in such dichotomously distributed data and it is likely that many if not all of these correlations are spurious.

Table 12. Stepwise ranked regression analysis for mean releases to Culebra, replicate 1.

	Expected Normalized Release		
	CRA-2009 PA Replicate 1		
Step[a]	**Variable[b]**	**R^{2c}**	**SRRC[d]**
1	BH_SAND:PRMX_LOG	0.81327	0.90850
2	STEEL:CORRMCO2	0.83387	-0.14945
3	DRZ_PCS:PRMX_LOG	0.84807	0.12237
4	CASTILER:PRESSURE	0.86222	0.11919

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Table 13. Stepwise ranked regression analysis for mean releases to Culebra, replicates 2 and 3.

	Expected Normalized Release					
	CRA-2009 PA Replicate 2			CRA-2009 PA Replicate 3		
Step[a]	Variable[b]	R^{2c}	SRRC[d]	Variable[b]	R^{2c}	SRRC[d]
1	BH_SAND:PRMX_LOG	0.86341	0.93683	BH_SAND:PRMX_LOG	0.83603	0.90480
2	SOLMOD3:SOLVAR	0.88302	0.14128	STEEL:CORRMCO2	0.86167	-0.15276
3	CASTILER:PRESSURE	0.90053	0.11142	CASTILER:PRESSURE	0.88227	0.14211
4	STEEL:CORRMCO2	0.90870	-0.10167	CASTILER:COMP_RCK	0.89399	-0.10828
5	GLOBAL:PBRINE	0.91506	0.07282	WAS_AREA:SAT_WICK	0.89865	-0.06843
6	WAS_AREA:PROBDEG	0.91943	-0.07454			
7	S_HALITE:POROSITY	0.92345	-0.06088			
8	S_HALITE:PRMX_LOG	0.92718	-0.06076			
9	SPALLMOD:REPIPOR	0.93047	0.05762			

[a] Steps in stepwise regression analysis
[b] Variables listed in order of selection
[c] Cumulative R^2 value with entry of each variable into regression model
[d] Standardized Rank Regression Coefficient

Summary and Conclusions

Of the four pathways of release, direct brine releases showed the greatest difference between the CRA-2004 PABC and the CRA-2009 PA results. However, there were few differences in the sensitivity analysis for direct brine releases or Culebra releases. The changes in DBR are most likely due to the changes in the porosity of halite and the DRZ, although an increase in the drilling rate could account for as much as an 11 % increase in DBR releases.

Cuttings and caving releases are the primary contributor to total releases. BOREHOLE:TAUFAIL controls about 98 % of the variability in mean cuttings and cavings releases through its highly non-linear relationship with the volume of cuttings and cavings produced. The drill string angular velocity (BOREHOLE:DOMEGA) may also have a small effect on cuttings and cavings release volumes.

The analysis of spallings releases showed some differences as compared to the CRA-2004 PABC due to changes in BRAGFLO, primarily the increase in the porosity of halite and the DRZ and to the increase in CPR related to the emplacement materials. The CRA-2009 PA had more events that generated spallings because repository pressures tended to be higher and gas volumes larger. The larger gas pressures and volumes were most likely due to the change in halite porosity, which would lead to larger volumes of water in the repository, and perhaps to the increase in CPR.

The low frequency of vectors having non-zero releases from the Culebra or of spallings continues to make the regression analysis of those results problematic. Whereas the importance of the top one or two variables seems to agree with expectations based on the functioning of the models, it is evident that the assumptions required for linear regression are unlikely to be met and that the dichotomy between zero and non-zero releases is likely to be driving the correlations. Analyses involving only the non-zero data provided little support for justifying the ordering of the parameters with regard to their impact on releases.

References:

Clayton, D.J. 2007. Corrections to Input Files for DBR PABC Calculations. Sandia National Laboratories, Carlsbad, NM. ERMS 546311.

Clayton, D. 2008a. Analysis Plan for the Performance Assessment for the 2009 Compliance Recertification Application, Revision 1. Sandia National Laboratories. Carlsbad, NM. ERMS # 547905.

Clayton, D. 2008b. Analysis Package for Direct Brine Releases: Compliance Recertification Application – 2009. Sandia National Laboratories. Carlsbad, NM.

Dunagan, S. 2004. Analysis Package for Cuttings and Cavings: Compliance Recertification Application. Revision 1.0. Sandia National Laboratories. Carlsbad, NM. ERMS # 533541.

DOE. 2004. Title 40 CFR Part 191 Compliance Recertification Application for the Waste Isolation Pilot. DOE/WIPP 2004-3231, U.S. Department of Energy Waste Isolation Pilot Plant, Carlsbad Field Office, Carlsbad, NM.

Ismail, A. E. 2007. Errors in input files for Nuts for CRA-2004 PABC calculations. Sandia National Laboratories, Carlsbad, NM. ERMS 546200.

Ismail, A. E. 2008. Analysis Package for CUTTINGS_S: Compliance Recertification Application 2009, Revision 1. Sandia National Laboratories. Carlsbad, NM.

Kirchner, T. 2004a. "Stepwise regression analysis of the final release for each release mechanism, in response to C-23-18, Revision 2." Memorandum to David Kessel dated December 6, 2004. Sandia National Laboratories. Carlsbad, NM. ERMS #537992.

Kirchner, T. 2004b. CRA Reponse Activity, Tracking Number 09/02/04Q, DOE Response to Comment C-23-18, Revision 1, ERMS 538114

Kirchner, T. 2005. Sensitivity of performance assessment baseline calculation release to parameters. Sandia National Laboratories. Carlsbad, NM. ERMS #540767.

Kirkes, G.R. 2007. Evaluation of the Duration of Direct Brine Release in WIPP Performance Assessment. April 27, 2007. Sandia National Laboratories, Carlsbad, NM. ERMS 545988.

Long, J. J. 2008. Execution of Performance Assessment Codes for the 2009 Performance Assessment Baseline Calculation, Rev. 0. Sandia National Laboratories. Carlsbad, NM.

Stein, J. S., and W. Zelinski. 2003. Analysis Package for BRAGFLO: Compliance Recertification Application. Analysis Report. Sandia National Laboratories, Carlsbad, NM. ERMS # 530163,

U. S. Environmental Protection Agency (EPA). 2004. "EPA's Completeness Comments, 3rd Set." Letter from Elizabeth Cotsworth to Dr. R. Paul Detwiler. September 2, 2004. ERMS # 536771.

Xiong, Y., E. J. Nowak and L. H. Brush. 2005. Updated Uncertainty Analysis of Actinide Solubilities for the Response to EPA Comment C-23-16, Rev. 1. Sandia National Laboratories. Carlsbad, NM. ERMS# 539595.

Attachment 1. CD containing the CCDFGF_Analysis.mdb and CCDFGF_Data.mdb databases

The attached CD contains copies of the CCDFGF_Analysis.mdb and CCDFGF_Data.mdb databases within the WinZip file AP137_ANALYSISDATABASE.ZIP and the Excel spreadsheets in AP137SimpleRegressionSpeadsheets.zip. To utilize these databases or spreadsheets they must be extracted from the zip file and installed on a PC running Microsoft Windows XP or a compatible operating system. CCDFGF_Analysis.mdb contains the queries and code used in the analysis and contains links to some of the tables in CCDFGF_Data.mdb. CCDFGF_Data.mdb contains the data. These links currently assume that CCDFGF_Data.mdb is located in C:\Program Files\CCDFGF_Analysis. If CCDFGF_Data.mdb is placed in another location then the links in CCDFGF_Analysis will need to be updated. These links can be updated using the Access menu item Tools/Database Utilities/Linked Table Manager. The means are computed and the STEPWISE input file is generated by selecting the menu button in CCDFGF_Analysis labeled "Compute Vector Means for STEPWISE". This process also produces STEPWISE input files for the parameters. A copy of AP137_ANALYSISDATABASE.ZIP and AP137SimpleRegressionSpeadsheets.zip are also stored in the CMS library LIBCRA09_STP.