Department of Energy
Office of Civilian Radioactive Waste Management
Yucca Mountain Site Characterization Office
P.O. Box 364629
North Las Vegas, NV 89036-8629

QA: N/A

AUG 29 2002

OVERNIGHT MAIL

Janet R. Schlueter, Chief
High-Level Waste Branch
Division of Waste Management
Office of Nuclear Materials Safety
and Safeguards
U.S. Nuclear Regulatory Commission
Two White Flint North
Rockville, MD 20852

TRANSMITTAL OF REPORT ADDRESSING KEY TECHNICAL ISSUE (KTI) AGREEMENT ITEM TOTAL SYSTEM PERFORMANCE ASSESSMENT AND INTEGRATION (TSPAI) 3.37

This letter transmits a report entitled *Response to TSPAI 3.37, Adequacy of the BDCF Sampling Method and Correlation*, which satisfies the subject KTI agreement. Specifically, the KTI agreement states:

> TSPAI 3.37: "Provide a quantitative analysis that the sampling method including the correlations to NP[sic] used by the TSPA code to abstract the GENII-S process model code adequately represent the uncertainty and variability and correlations for the biosphere process model (DOSE 3.4.1) [sic].
> DOE will provide a quantitative analysis that the sampling method including the correlations between BDCFs utilized by the TSPA code to abstract the GENII-S process model data adequately represent the uncertainty and variability and correlations for the biosphere process model. This will be documented in Nominal Performance Biosphere Dose Conversion Factor Analysis (ANL-MGR-MD-000009), Disruptive Event Biosphere Dose Conversion Factor Analysis (ANL-MGR-MD-000003) or other document expected to be available to NRC in FY 2003. Results of these analyses will be documented in the TSPA for any potential license application expected to be available to NRC in FY 2003."

The enclosure to this letter is an assessment addressing the adequacy of the approach used in the Total System Performance Assessment (TSPA) model to abstract the biosphere dose conversion factor (BDCF) data and represent the uncertainty, variability, and correlation for the Biosphere model. As described in the assessment, the expected dose at a given time from a set of stochastic TSPA calculations is independent of the degree of correlation between the BDCFs. To ensure that the TSPA-Site Recommendation model did not systematically underestimate annual dose, and thus, adequately represented the uncertainty and variability for the biosphere process model, the conservative assumption that all BDCF data sets were correlated with the correlation coefficient of unity was used.

NMSS07

This letter contains no new regulatory commitments. Please direct any questions concerning this letter and its enclosure to Timothy C. Gunter at (702) 794-1343 or Mark C. Tynan at (702) 794-5457.

Joseph D. Ziegler
Acting Assistant Manager
Office of Licensing and Regulatory Compliance

OL&RC:TCG-1671

Enclosure:
Response to TSPAI 3.37, Adequacy of the BDCF Sampling Method and Correlation

cc w/encl:
J. W. Andersen, NRC, Rockville, MD
D. D. Chamberlain, NRC, Arlington, TX
R. M. Latta, NRC, Las Vegas, NV
H. J. Larson, ACNW, Rockville, MD
W. D. Barnard, NWTRB, Arlington, VA
W. C. Patrick, CNWRA, San Antonio, TX
Budhi Sagar, CNWRA, San Antonio, TX
Steve Kraft, NEI, Washington, DC
J. H. Kessler, EPRI, Palo Alto, CA
J. R. Egan, Egan & Associates, McLean, VA
R. R. Loux, State of Nevada, Carson City, NV
Marjorie Paslov-Thomas, State of Nevada, Carson City, NV
Irene Navis, Clark County, Las Vegas, NV
Alan Kalt, Churchill County, Fallon, NV
George McCorkell, Esmeralda County, Goldfield, NV
Leonard Fiorenzi, Eureka County, Eureka, NV
Michael King, Inyo County, Edmonds, WA
Andrew Remus, Inyo County, Independence, CA
Mickey Yarbro, Lander County, Battle Mountain, NV
Lola Stark, Lincoln County, Caliente, NV
Arlo Funk, Mineral County, Hawthorne, NV
L. W. Bradshaw, Nye County, Pahrump, NV
David Chavez, Nye County, Tonopah, NV
Josie Larson, White Pine County, Ely, NV
R. I. Holden, National Congress of American Indians, Washington, DC
Allen Ambler, Nevada Indian Environmental Coalition, Fallon, NV
CMS Coordinator, BSC, Las Vegas, NV
OL&RC Library
Records Processing Center = "19"

cc w/o encl:
C. W. Reamer, NRC, Rockville, MD
L. L. Campbell, NRC, Rockville, MD
N. K. Stablein, NRC, Rockville, MD
S. L. Wastler, NRC, Rockville, MD
Margaret Chu, DOE/HQ (RW-1), FORS
R. A. Milner, DOE/HQ (RW-2), FORS
A. B. Brownstein, DOE/HQ (RW-52), FORS
N. H. Slater-Thompson, DOE/HQ (RW-52), FORS
R. B. Murthy, DOE/OQA (RW-3), Las Vegas, NV
S. J. Cereghino, BSC, Las Vegas, NV
Donald Beckman, BSC, Las Vegas, NV
N. H. Williams, BSC, Las Vegas, NV
R. B. Bradbury, MTS, Las Vegas, NV
R. C. Murray, MTS, Las Vegas, NV
R. D. Rogers, MTS, Las Vegas, NV
Richard Goffi, BAH, Washington, DC
J. R. Dyer, DOE/YMSCO, Las Vegas, NV
D. G. Horton, DOE/YMSCO, Las Vegas, NV
W. J. Boyle, DOE/YMSCO, Las Vegas, NV
T. C. Gunter, DOE/YMSCO, Las Vegas, NV
C. L. Hanlon, DOE/YMSCO, Las Vegas, NV
G. W. Hellstrom, DOE/YMSCO, Las Vegas, NV
S. P. Mellington, DOE/YMSCO, Las Vegas, NV
C. M. Newbury, DOE/YMSCO, Las Vegas, NV
G. L. Smith, DOE/YMSCO, Las Vegas, NV
R. E. Spence, DOE/YMSCO, Las Vegas, NV
M. C. Tynan, DOE/YMSCO, Las Vegas, NV
J. D. Ziegler, DOE/YMSCO, Las Vegas, NV
C. A. Kouts, DOE/YMSCO (RW-2), FORS
R. N. Wells, DOE/YMSCO (RW-60), Las Vegas, NV

RESPONSE TO TSPAI 3.37, ADEQUACY OF THE BDCF SAMPLING METHOD AND CORRELATION

August 2002

Prepared by:

J.K. Prince, License Application Project — Date

A.J. Smith, Biosphere Department, Natural Systems — 28 Aug 2002 Date

Technical Review by:

M.A. Wasiolek, Biosphere Department, Natural Systems — 8/28/02 Date

Approval:

E. Hardin, Natural Systems — 8/28/02 Date

ENCLOSURE

ACRONYMS AND ABBREVIATIONS

BDCF	biosphere dose conversion factor
DOE	U.S. Department of Energy
NRC	U.S. Nuclear Regulatory Commission
TSPA	Total System Performance Assessment
SR	Site Recommendation

Adequacy of the BDCF Sampling Method and Correlation

This report describes the basis for resolving Total System Performance Assessment and Integration agreement item 3.37 associated with how the approach used in the total system performance assessment (TSPA) model to abstract the biosphere dose conversion factor (BDCF) data adequately represents the uncertainty, variability, and correlation for the biosphere model. The agreement specifies that the U.S. Department of Energy (DOE) provide clarification concerning how the BDCF abstraction was implemented in the TSPA model.

1. BACKGROUND

The DOE agreed to provide a quantitative analysis showing that the BDCF sampling method as used in TSPA for the Site Recommendation (SR) adequately represents the uncertainty, variability, and correlation for the biosphere process model. This approach is currently being considered for potential use in TSPA for License Application (LA). The biosphere modeling approach used in TSPA-LA will incorporate modifications to reflect use of reasonably maximally exposed individual and other regulatory requirements of 10 CFR 63. If this approach is to be used in support of LA, this analysis will be documented in a revision of *Nominal Performance Biosphere Dose Conversion Factor Analysis* ANL-MGR-MD-000009; *Disruptive Event Biosphere Dose Conversion Factor Analysis*, ANL-MGR-MD-000003; or other documentation. Should another more realistic, but less conservative method be selected for use in TSPA-LA, the NRC will be informed. This analysis documents the basis for the TSPA-SR correlation assumption.

1.1 NRC INITIAL COMMENTS

The U.S. Nuclear Regulatory Commission (NRC) noted that the approach used to propagate uncertainty in BDCFs for the biosphere abstraction in the TSPA-SR model introduces unnatural correlation (e.g., samples from radionuclide-specific BDCF distributions are correlated to the neptunium-237 BDCF distribution, and no justification for this approach was provided). The NRC further indicated that the biosphere factors that influence the magnitude of BDCFs vary by radionuclide, and the justification for the selected approach is not self-evident. In addition, the NRC noted that failure to maintain vectors from initial GENII-S BDCF modeling leads to inconsistencies in sampled biosphere and receptor parameters across radionuclides when resampling in TSPA-SR (BSC 2001).

1.2 DOE INITIAL COMMENTS

The DOE noted that the effect of the correlation between BDCFs on dose is only an issue for those times where multiple radionuclides are predicted to contribute appreciably to dose. In the case where two or more radionuclides contribute to dose, the assumption used in TSPA-SR is that the correlation between the BDCFs for each pair of radionuclides is unity. Neptunium was selected as the starting point for the correlation (i.e., the correlation between the BDCFs for radionuclide #1 and neptunium-237 is unity, and the correlation between the BDCFs for neptunium-237 and radionuclide #2 is also unity). This assumption could have an effect on the TSPA dose predictions. Statistical theory (Section 3.2) indicates that irrespective of the degree of correlation, the expected value of the average of a set of TSPA realizations is constant. The

DOE also indicated that the GENII-S code does not provide the capability to comprehensively consider correlation between parameters for multiple radionuclides (BSC 2001).

1.3 DEFINITION OF TECHNICAL TERMS

Biosphere Dose Conversion Factor (BDCF)—A radionuclide-specific parameter that accounts for all pathways contributing to radiation exposure for a defined receptor and environment. A BDCF is a scalar value that converts activity concentration (in a predefined unit, such as becquerel per cubic meter) into annual committed dose.

Correlation Coefficient—A dimensionless parameter that, with certain reservations, is used as a measure of the linear relationship between two variables.

Dose—The regulatory specified measure of exposure of the defined receptor to radionuclides introduced by a given release scenario into the accessible environment.

Expected Value—The expected value of a sample for a distribution is the average value (μ) of that distribution. If the probability distribution for x is $f(x)$, then:

$$\mu = \int x f(x) dx$$

Mean Value—The average value ($\bar{x}$) of a sample from a distribution. For a given distribution, $\bar{x}$ is an estimator of μ. If n samples are taken from the distribution, with x_i being the i^{th} sample, the mean value is given by:

$$\bar{x} = \frac{\sum_{1}^{n} x_i}{n}$$

Sample Variance— (s^2) The expected value of the square of the deviation from the mean value:

$$s^2 = \frac{1}{(n-1)} \sum_{1}^{n} (x_i - \bar{x})^2$$

Seed—An initial input parameter to the pseudo-random number generator used in computers to enable the user to define a sequence of random numbers. Using a fixed seed allows the user to generate an identical sequence of random numbers. This facility is used to ensure that the calculations for each radionuclide use the same set of random numbers.

Vector—A one-dimensional array that, in the context used herein, contains the ordered sequence of the stochastic BDCFs as generated by GENII-S. For each radionuclide, the i^{th} vector component (BDCF value) was generated by the same set of stochastic samples from the input parameters (i.e., each radionuclide run was initiated using the same seed).

2. APPLICABLE NUCLEAR SAFETY STANDARDS, REQUIREMENTS, AND GUIDANCE

2.1 APPLICABLE REQUIREMENTS

The requirement to evaluate uncertainty in the performance assessment is contained in 10 CFR 63.114 (b):

> Any performance assessment used to demonstrate compliance with 63.113 must:
>
> (b) account for uncertainties and variabilities in parameter values and provide for the technical basis for parameter ranges, probability distributions or bounding values used in the performance assessment.

2.2 KEY TECHNICAL ISSUE AGREEMENT

The following statement of Key Technical Issue agreement TSPAI 3.37, is documented in the summary highlights of the NRC and DOE technical exchange and management meeting on TSPA and integration that occurred on August 6-10, 2001 (Reamer 2001):

> Provide a quantitative analysis that the sampling method including the correlations to NP [sic] used by the TSPA code to abstract the GENII-S process model code adequately represents the uncertainty and variability and correlations for the biosphere process model (DOSE3.4.1) [sic].
>
> DOE will provide a quantitative analysis that the sampling method including the correlations between BDCFs utilized by the TSPA code to abstract the GENII-S process model data adequately represents the uncertainty and variability and correlations for the biosphere process model. This will be documented in Nominal Performance Biosphere Dose Conversion Factor Analysis AMR (ANL-MGR-MD-000009), Disruptive Event Biosphere Dose Conversion Factor Analysis (ANL-MGR-MD-000003) or other document expected to be available to NRC in FY 2003. Results of these analyses will be documented in the TSPA for any potential license application expected to be available to NRC in FY 2003.

3. BASIS FOR REGULATORY COMPLIANCE STATEMENT

The information in Section 3.1 provides background information concerning the BDCFs and their uncertainty and variability as represented by probability distribution functions. Section 3.2 contains a quantitative evaluation of the effect of using unity for the correlation coefficient when sampling from a BDCF distribution for other radionuclides and from the BDCF distribution for neptunium-237.

3.1 BACKGROUND

The BDCF distributions for the TSPA-SR model are reported in *Nominal Performance Biosphere Dose Conversion Factor Analysis* (CRWMS M&O 2001a) and *Disruptive Event Biosphere Dose Conversion Factor Analysis* (CRWMS M&O 2001b). The GENII-S code was

used in each analysis to generate a set of BDCF realizations for each radionuclide by randomly sampling over multiple inputs defined by distributions to reflect their individual uncertainty and variability. Using the same predefined seed for each radionuclide, i, ensured that the distribution of the j^{th} parameter for that radionuclide, p_{ij}, was sampled at the same percentile point as all the other radionuclides considered. Thus for parameters that are common to each radionuclide (such as rainfall, eating habits, time spent outdoors) correlation between BDCFs are captured. For example, if for a particular realization for the first radionuclide the code selected a random value of drinking water consumption rate of 423.5 liters per year, then each additional radionuclide would have the drinking water rate sampled as 423.5 liters per year for that realization. Thus, as far as the variability and uncertainty in parameters used to define the receptor are concerned, the resultant BDCF sets are correlated as expected. If these ordered BDCF data were provided to TSPA model, they could be sampled by order, and the results would exhibit the correlation as reflected by the input uncertainty and variability. However, as discussed later, the GENII-S code cannot consider the loss of contaminated soil by erosion. This process has a significant effect on the magnitude of the BDCFs for radionuclides with a large sorption coefficient on soils. The effect of soil loss was included in the BDCFs by postprocessing, which would have raised questions about the validity of any correlation inferred from the input data.

There are some input parameters that may reasonably be expected to be correlated. Examples include the correlation between the radionuclide transfer factors from soil to plants for radionuclides considered in the TSPA analysis and the correlation between the partition coefficients used to determine the radionuclide leaching rate from soil. These parameters for one radionuclide could be correlated with the corresponding parameters for another radionuclide. The version of the GENII-S code used does not allow these correlations to be incorporated into the analysis. Although the correlation due to a common factor (across all radionuclides) in areas such as the habits and diets of the receptor can be addressed by GENII-S, there are other potential correlation effects between input parameters that cannot be considered by GENII-S or that are simply unknown. Examples of unknowns include the correlation between transfer coefficient for different produce classes and the correlation between the sorption coefficients between various elements.

The approach outlined above could be used to consider some but not all of the correlation in the BDCFs. In addition, the approach adopted in the TSPA-SR model required post processing. This processing did not preserve the correlation. In the case of the groundwater release scenario, the effect of soil erosion is taken into account externally to GENII-S. The approach adopted to quantify the soil removal effect required that the GENII-S output be approximated by an analytical statistical distribution. Statistical methods were used to demonstrate that the fitted distribution was acceptable to represent the GENII-S data (CRWMS M&O 2001a; CRWMS M&O 2001b). This process required the BDCFs to be put into a number of bins. This binning process, while incorporating the underlying input conditions that were inherent in the original data, did not preserve the correlations between the outputs. In the TSPA-SR model these distributions were used as abstractions for the BDCFs in the TSPA code (CRWMS M&O 2000).

For the volcanic release BDCFs, there was no need to fit a statistical distribution. The BDCF distribution was captured by using the ordered data (i.e., the data were placed in numerical ascending order during which the correlation was lost) to generate a table of percentile points

(minimum, maximum, and every fifth percentile). The TSPA-SR model used the tables of BDCF percentiles from which random samples were drawn (CRWMS M&O 2000).

To address correlation in TSPA analysis when sampling the BDCFs for a particular scenario, an assumption was made in the TSPA-SR model to assign a correlation coefficient of unity between the BDCFs for all radionuclides in a data set. This assumption was based on the fact that the mean value of the sum of multiple sets of random variables (irrespective of correlation) is the sum of the mean values of the sets. Correlation manifests itself when the variance of the sum is considered. The variance is at its maximum when the correlation is unity and falls as the correlation is decreased to zero. This can be illustrated in the case of fully correlated variables: When a large value of one variable is selected, a large value of all variables in the sum is also selected. When there is no correlation, all variables are sampled at random, thereby reducing the variance. This fact is developed in detail in Section 3.2.

As discussed during a telephone conference call on August 28, 2002, the NRC noted that, while the approach used in TSPA-SR appears to be conservative it is not a realistic representation of the physical phenomenon in the biosphere. The DOE acknowledged the NRC comments and restated that the approach is conservative. In addition, DOE noted that this conservatism yields pessimistic results with regard to demonstrating compliance with the individual protection requirements of 10 CFR 63 and the Yucca Mountain Review Plan acceptance criteria. Therefore, the Project selected a conservative approach with full knowledge that it does not reflect realism. The DOE intends to continue to use some conservative calculations for the License Application.

3.2 TECHNICAL BASIS FOR USING A CORRELATION COEFFICIENT OF UNITY BETWEEN THE BDCF FOR NEPTUNIUM-237 AND THE BDCF FOR ALL OTHER RADIONUCLDIES WHEN SAMPLING IN THE TSPA-SR CODE

3.2.1 Fixed Concentrations of Radionuclides

As discussed in Section 3.1, if a single radionuclide dominates dose, then the question of correlation is moot. For the purpose of this analysis of correlation, the assumption is a time in the TSPA dose history where two radionuclides make approximately equal contributions to dose. This situation occurs a few times in a TSPA calculation and is the period of time when a radionuclide that had been dominant in dose determination is in the process of being superceded in dominance by another radionuclide.

The two radionuclides contributing to dose are assumed to be X and Y, and the true deterministic (not stochastic) concentrations of X and Y in the groundwater (the analysis of the volcanic release case is similar) are assumed to be C_X and C_Y, respectively. If the i^{th} random sample of the BDCFs for radionuclides X and Y are x_i and y_i, then the expected dose D_i is:

$$D_i = C_X x_i + C_Y y_i \qquad \text{Eq. 1}$$

The expected dose is then estimated by:

$$\overline{D} = (1/n)\sum_{1}^{n}(C_X x_i + C_Y y_i)$$

$$= (1/n)\left(\sum_{1}^{n} C_X x_i + \sum_{1}^{n} C_Y y_i\right)$$

$$= (1/n)\left(C_X \sum_{1}^{n} x_i + C_Y \sum_{1}^{n} y_i\right)$$

$$= C_X \overline{x} + C_Y \overline{y} \qquad \text{Eq. 2}$$

Equation 2 holds whatever correlation exists between x_i and y_i. For given concentrations of two contaminants, each having its own distribution of BDCFs (correlated to any degree), the expected combined dose is dependent only upon the mean values of the BDCFs. The expected dose is not dependent on the higher moments (shape) of the distributions or correlation between the distributions.

To evaluate whether the assumption used in the TSPA model to assume that BDCFs were all related by a correlation coefficient of unity was adequate in determining dose, two cases of correlation are considered. The first case considers that all the BDCFs for the individual radionuclides are (statistically) independent. In such a case, the BDCFs will have a correlation coefficient of zero. The other case considers the TSPA approach where the BDCFs are totally correlated with a coefficient of unity. Although, as shown above, the compliance metric of expected dose is independent of correlation, the two cases considered (r = 0 and 1) do produce different distributions of predicted dose. The difference between these two cases will demonstrate that the approach used in TSPA does not introduce a non-conservative bias into the demonstration of compliance.

The approach requires the development of the expected variance of the dose discussed above and establishes the conditions for a correlation coefficient of unity between the BDCFs.

The sample variance (s_D^2) of n dose realizations (D_i) can be calculated by the standard formula:

$$s_D^2 = \frac{1}{(n-1)}\sum_{1}^{n}(D_i - \overline{D})^2$$

Substituting for D_i from Equation 1 and for $\overline{D}$ from Equation 2 gives:

$$s_D^2 = \frac{1}{(n-1)}\sum_{1}^{n}((C_X x_i + C_Y y_i) - (C_X \overline{x} + C_Y \overline{y}))^2$$

$$= \frac{1}{(n-1)}\sum_{1}^{n} C_X^2 (x_i - \overline{x})^2 + \frac{2}{(n-1)}\sum_{1}^{n} C_X C_Y (x_i - \overline{x})(y_i - \overline{y}) + \frac{1}{(n-1)}\sum_{1}^{n} C_Y^2 (y_i - \overline{y})^2$$

$$= \frac{C_X^2}{n-1}\sum_{1}^{n}(x_i - \overline{x})^2 + \frac{2C_X C_Y}{n-1}\sum_{1}^{n}(x_i - \overline{x})(y_i - \overline{y}) + \frac{C_Y^2}{n-1}\sum_{1}^{n}(y_i - \overline{y})^2$$

$$= C_X^2 s_X^2 + 2C_X C_Y s_{XY} + C_Y^2 s_Y^2 \qquad \text{Eq. 3}$$

To evaluate the effect of correlation on the TSPA results, the correlation coefficient (r) is defined by:

$$r = \frac{\sum_1^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_1^n (x_i - \bar{x})^2 \sum_1^n (y_i - \bar{y})^2}} \qquad \text{Eq. 4}$$

Assuming the case of a simple linear relationship between the BDCF data (x_i, y_i) such that

$$y_i = a + bx_i \qquad \text{Eq. 5}$$

by summing over all i and dividing by the number of samples (n), it follows that:

$$\bar{y} = a + b\bar{x} \qquad \text{Eq. 6}$$

Substituting the relationships in Equations 5 and 6 into Equation 4 allows the correlation coefficient for these data to be calculated:

$$r = \frac{\sum_1^n (x_i - \bar{x})(a + bx_i - (a + b\bar{x}))}{\sqrt{\sum_1^n (x_i - \bar{x})^2 \sum_1^n (a + bx_i - (a + b\bar{x}))^2}}$$

$$= \frac{\sum_1^n (x_i - \bar{x})(bx_i - b\bar{x})}{\sqrt{\sum_1^n (x_i - \bar{x})^2 \sum_1^n (bx_i - b\bar{x}))^2}}$$

$$= \frac{b\sum_1^n (x_i - \bar{x})^2}{\sqrt{b^2 \left(\sum_1^n (x_i - \bar{x})^2 \right)^2}}$$

$$= \frac{b\sum_1^n (x_i - \bar{x})^2}{abs(b)\sum_1^n (x_i - \bar{x})^2}$$

$$= sign(b) \times 1$$

Thus, when the random variable pairs are related by a linear relationship, the data sets have a correlation coefficient of magnitude unity. The sign of the correlation coefficient is the sign of the slope (b) in linear relationship between the variable pairs. In this case, in Equation 4 the numerator must be equal to the denominator, thus:

$$\sum_{1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \pm\sqrt{\sum_{1}^{n}(x_i-\bar{x})^2\sum_{1}^{n}(y_i-\bar{y})^2}$$

$$= \pm(n-1)\sqrt{\frac{\sum_{1}^{n}(x_i-\bar{x})^2}{(n-1)}\frac{\sum_{1}^{n}(y_i-\bar{y})^2}{(n-1)}}$$

$$= \pm(n-1)s_x s_y \qquad \text{Eq. 7}$$

For the case where the correlation coefficient is zero, the numerator in Equation 4 has to be zero. As the variables are independent, the terms in the summation can be evaluated separately, and both terms are identically equal to zero from the definition of the mean:

$$\sum_{1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = 0 \qquad \text{Eq. 8}$$

Using Equation 3 and substituting for the term defined above in Equation 8 gives for the variance case of no correlation (the cross product term is zero):

$$s_D^2(r=0) = C_X^2 s_X^2 + C_Y^2 s_Y^2$$

Using a correlation coefficient of unity, Equations 3 and 7 give:

$$s_D^2(r=\pm 1) = C_X^2 s_X^2 \pm 2C_X C_Y s_X s_Y + C_Y^2 s_Y^2$$
$$= \pm(C_X s_X + C_Y s_Y)^2$$

From these two results it can be seen for a given data set that:

$$s_D^2(r=1) \geq s_D^2(r=0)$$

Intermediate values of r give rise to intermediate values of s_D^2. Equality only holds when either the variance of one of the BDCFs is zero or the concentration of one of the radionuclides is zero. In the former case, the BDCF distribution has zero variance (i.e., zero width), implying there is no uncertainty in the BDCF values; therefore, BDCFs can be represented as a simple deterministic scalar. In the latter case, there is no radionuclide present. In both cases, the question of correlation is of no significance.

When only two radionuclides are important, it is conceivable that the BDCFs could exhibit negative correlation. The result in this case is:

$$s_D^2(r=0) \geq s_D^2(r=-1)$$

This and other TSPAs calculate dose by conducting a set of realizations using input parameters selected at random over their uncertainty and variability distribution to generate a set of dose values as a function of time ($D_i(t)$). These multiple realizations are then used to generate the expected (mean) dose as a function of time, $\overline{D}(t)$ (some TSPA sensitivity analyses use other measures of dose such as the 95th percentile). Equation 1 indicates that the expected dose does not depend upon any correlation between BDCFs. However, because of the finite sample size, when using the TSPA model, the average value calculated from the set of dose calculations, $D_i(t)$, is only an estimate of the true expected value. As the sample size increases, the average tends (asymptotically) to the true expected value. Standard statistical methods as described in this report can be used to provide estimates of the uncertainty associated with the finite sample size. The uncertainty bounds derived from this approach are proportional to variance of the distribution.

If the true value of the mean dose of the TSPA calculation at a fixed time is μ_D (this value can only be known if an infinite number of realizations are made for the particular scenario under study), standard statistical texts tell us that the statistic t defined in Equation 9 has the Student t distribution with *(n-1)* degrees of freedom. As n increases, the Student t distribution tends to the normal distribution. For the number of realizations generated in TSPA (typically several hundred), the normal approximation is adequate:

$$t = \frac{\overline{D} - \mu_D}{s_D / \sqrt{n}} \qquad \text{Eq. 9}$$

If TSPA generates $\overline{D}$ and s_D, from n realizations, then an upper limit can be derived for the true mean dose, μ_D, at a given confidence level. At a confidence level of 95%, μ_D is given by:

$$\mu_D < \overline{D} + \frac{1.64 s_D}{\sqrt{n}}$$

As s_D reaches its largest value when the BDCFs are positively correlated, the use of r=1 in BDCF sampling within the TSPA code leads to a higher upper bound for the estimate of the true mean of the distribution. Thus, this sampling procedure is conservative in establishing an upper bound for the expected dose.

3.2.2 Stochastic Concentrations of Radionuclides

In the case where radionuclide concentrations are stochastic variables generated within TSPA, the dose for realization i is given by Equation 10:

$$D_i = C_{Xi} x_i + C_{Yi} y_i \qquad \text{Eq. 10}$$

A quantitative analysis of the effects of correlation on Equation 10 is more difficult to evaluate than for the simple case presented in Section 3.2.1. Not only is there the possibility of

correlation between the BDCFs (x_i and y_i) addressed by this report, but there is also the potential for correlation to exist between the concentrations (C_{xi} and C_{yi}) and between the BDCF and concentration for each radionuclide. (This latter correlation could readily be caused by changes in precipitation during climate evolution.) The matter is further complicated by the necessity in any TSPA analysis to demonstrate convergence of the expected value (and, for this analysis, the second moment used in Equation 9) of D from a limited number of realizations.

The total number of realizations considered in the TSPA is assumed to be large (n). Dividing the range of each of the stochastic concentrations (C_x and C_y) into concentration bins (j and k) thereby forms a two-dimensional array that covers all concentration combinations for the two contributing radionuclides. For each realization i, the appropriate two-dimensional concentration bin (j, k) should be located (i.e., the intersection of the C_{xj} bin and the C_{yk} bin), and the stochastically sampled BDCFs (x_i and y_i) should be inserted. This operation should be repeated for each i of the set of n.

An estimate of the contribution to dose from the samples in bin i, j is generated by using the midpoints of the concentration bin (i.e., $\overline{C_{xj}}$ and $\overline{C_{yk}}$), rather than the actual concentration. This operation does introduce an error, but by increasing the sample size n and decreasing the bin widths, any error can be reduced to inconsequential levels.

For each concentration bin, the two applicable stochastic concentrations have been replaced with fixed (midpoint) values. Thus, for each bin, the logic developed in Section 3.2.1 applies. The contribution to dose from the bin is unaffected by any correlation between the BDCFs. The variance of predicted dose within each bin increases as the correlation between BDCFs is increased from zero to unity.

After summation over all bins, the expected dose is unchanged by BDCF correlation, while the variance of the total sample is maximized when the correlation is unity.

3.3 SUMMARY

The expected dose at a given time from a set of stochastic TSPA calculations is independent of the degree of correlation between the BDCFs. Use of a correlation coefficient of unity between the BDCFs, as was done in the TPSA-SR model, gives rise to a wider distribution of dose over the individual realizations than would be obtained for a lower degree of correlation. Therefore, when the TSPA-SR model uses a high percentile value of the predicted dose distribution (D_i) to measure performance, the estimate is more conservative than would have been obtained using a lower correlation. The correlation in the BDCF data used in the TSPA-SR model, as generated by ordered vector sampling due to GENII-S, primarily arises from the sampling of the characteristics of the receptor. The sampling of these uncertainties to generate the BDCFs gives rise to some correlation that can be taken into account by using the BDCF vectors. However, there are other processes not addressed by GENII-S, where relationships between additional parametric uncertainties may introduce additional correlation. Such a case could arise if soil pH could affect transfer factors from soil to plants as well as radionuclide build-up factors in soils. If GENII-S could stochastically sample over the range of uncertainty of pH, it is possible the correlation between BDCFs could potentially be greatly increased. The approach described by the NRC (at the August 6 to 10, 2001 technical exchange) of using the limited capability of the

GENII-S without justification to demonstrate that other processes do not influence the correlation between BDCFs could result in nonconservative compliance conclusions being drawn.

To ensure that the TSPA-SR model did not systematically underestimate annual dose, the conservative assumption that all BDCFs data sets were correlated with the correlation coefficient of unity was used. This approach adequately represented the uncertainty, variability, and correlation for the biosphere process model in the TSPA-SR model by using a conservative assumption that is documented in this report.

4. REFERENCES

4.1 DOCUMENTS CITED

BSC (Bechtel SAIC Company) 2001. *Analysis of Resolution Status Key Technical Issue: Total System Performance Assessment and Integration Subissue 3: Model Abstraction.* Las Vegas, Nevada: Bechtel SAIC Company. ACC: MOL.20010921.0129.

CRWMS M&O 2000. *Total System Performance Assessment for the Site Recommendation.* TDR-WIS-PA-000001 REV 00 ICN 01. Las Vegas, Nevada: CRWMS M&O. ACC: MOL.20001220.0045.

CRWMS (Civilian Radioactive Waste Management System) M&O (Management and Operating Contractor) 2001a. *Nominal Performance Biosphere Dose Conversion Factor Analysis.* ANL-MGR-MD-000009 REV 01. Las Vegas, Nevada: CRWMS M&O. ACC: MOL.20010123.0123.

CRWMS M&O 2001b. *Disruptive Event Biosphere Dose Conversion Factor Analysis.* ANL-MGR-MD-000003 REV 01. Las Vegas, Nevada: CRWMS M&O. ACC: MOL.20010125.0233.

Reamer, C.W. 2001. "U.S. Nuclear Regulatory Commission/U.S. Department of Energy Technical Exchange and Management Meeting on Total System Performance Assessment and Integration (August 6 through 10, 2001)." Letter from C.W. Reamer (NRC) to S. Brocoum (DOE/YMSCO), August 23, 2001, with enclosure. ACC: MOL.20011029.0281.

4.2 CODES, STANDARDS, REGULATIONS, AND PROCEDURES

10 CFR 63. Energy: Disposal of High-level Radioactive Wastes in a Geologic Repository at Yucca Mountain, Nevada. Readily available.