PNNL-16940 Final

Pacific Northwest National Laboratory
Operated by Battelle for the U.S. Department of Energy

Hanford Soil Inventory Model (SIM) Rev. 2 Software Documentation – Requirements, Design, and Limitations

M. J. Anderson
B. C. Simpson
R. A. Corbin
C. T. Kincaid

October 2007

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

This document was printed on recycled paper.

PNNL-16940 Final

Hanford Soil Inventory Model (SIM) Rev. 2 Software Documentation – Requirements, Design, and Limitations

M. J. Anderson[(a)]
B. C. Simpson[(a)]
R. A. Corbin[(b)]
C. T. Kincaid

October 2007

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

(a) Vivid Learning Systems, Inc., Pasco, Washington
(b) Columbia Energy and Environmental Services, Inc., Richland, Washington

Foreword

This document reports the requirements, conceptual model, simulation methodology, testing, and quality assurance associated with Revision 2 of the Hanford Soil Inventory Model (SIM) at a summary level.

Terms/Acronyms

AMD	advanced micro devices
CB	Crystal Ball
CFL	correction factors – liquid
CFS	correction factors – solid
Ci	curie
COTS	commercial off-the-shelf
DTS	data transformation services
HDW	Hanford Defined Waste (Model)
g	gram
GB	gigabyte
GHz	gigahertz
GUI	graphical user interface
LHS	Latin hypercube sampling
MB	megabyte
μCi	microcurie
μg	microgram
ML	megaliter
mL	milliliter
OCB	Open Crystal Ball
PNNL	Pacific Northwest National Laboratory
QA	quality assurance
RAM	random access memory
RPD	relative percent difference
RSD	relative standard deviation
SAC	System Assessment Capability
SIM	Soil Inventory Model
UPS	uninterruptible power supply
USB	universal serial bus
VBA	Visual Basic for Applications

Contents

Figures

Tables

1.0 Overview

This document reports the requirements, conceptual model, simulation methodology, testing, and quality assurance associated with Revision 2 of the Hanford Soil Inventory Model (SIM) at a summary level. It supersedes the documentation included in Simpson et al. (2006b). The software life-cycle documentation associated with Hanford SIM Rev. 2, which provides greater detail on each of these topics, is part of the project file and, therefore part of the project record.

Because of the extensive use of commercial off-the-shelf (COTS) software products, the software development effort as part of this application was limited; however, adaptation efforts were substantial. Significant testing was completed of the various COTS features and capabilities used and the means by which they were integrated into Hanford SIM Rev. 2. Another outcome of the fiscal year 2007 effort is essentially the duplication of the Corbin et al. (2005) Hanford SIM Rev. 1 results using a version of the Hanford SIM created under DOE Order 414.1C for safety software.

The Hanford SIM Rev. 2 application computed waste discharges composed of 75 analytes at 377 waste sites (liquid disposal, unplanned releases, and tank farm leaks) over an operational period of approximately 50 years. The development and application of Hanford SIM Rev. 2 was an effort to develop a probabilistic approach to estimate comprehensive, mass-balance–based contaminant inventories for the Hanford Site post-closure setting. A computer model capable of calculating inventories and the associated uncertainties as a function of time was identified to address the needs of the Hanford Site.

To estimate mass-balanced contaminant inventories and their uncertainties for the Hanford Site post-closure setting, a stochastic simulation method (a Monte Carlo-type calculation) was selected to provide estimates of inventory and uncertainty. In this approach, several options were considered for model development, and the Open Crystal Ball (OCB) statistical package was selected in 2002.

A COTS limitation of the Hanford SIM Rev. 1 effort was that the core OCB and Crystal Ball (CB) applications were updated, and, hence, those used are now unavailable in the marketplace. Thus, the Hanford SIM Rev. 1 software components used and documented in Corbin et al. (2005) were either not compatible or no longer available. The Hanford SIM Rev. 2 portfolio of applications has been updated with commercially available software.

The design of Hanford SIM Rev. 2 is highly modular, with separate data input (Microsoft Excel) and calculation engine (OCB.dll) files administered through an interface application that acquires the inputs, manages the data reporting, and creates the output files. This design method allowed for concurrent development of the individual model elements, increasing project efficiency. Each data input is considered an independent variable; therefore, the waste stream composition/properties and waste stream discharge histories for the waste disposal sites could be examined and developed using a variety of source data (e.g., historical process data, tank waste modeling) and assumptions without influencing other variables.

Substantial efforts have been made to ensure that users could operate the software with a reasonable amount of training. However, the modeling software portfolio is very demanding on the user with regard to developing or modifying input data. Becoming proficient with software operation requires significant practice.

The application of Hanford SIM Rev. 2 has several limitations. Principal among them is that history matching the model results with reference data was not a goal of Hanford SIM Rev. 1, prior to publication of model results (e.g., Corbin et al. 2005). Instead, history matching is an ongoing effort that relies on the interpretation of field characterization data and is part of the error management, configuration control, and maintenance activities. Intensive history matching is a proposed future activity. Additionally, because the fiscal year 2007 effort was focused on creating a portfolio of software tools to be distributed, rather than updating the model results, no corrections or additions were made to the simulation inputs or outputs.

Other general limitations of the software and its results are associated largely with the reliability and availability of input data. The inputs used to generate the inventories for the Hanford SIM Rev. 2 architecture are controlled by various independent organizations (Pacific Northwest National Laboratory; CH2M HILL Hanford Group, Inc.; and Fluor Hanford, Inc.). However, the consistency and appropriateness of the various physical and chemical assumptions used in quantifying model behavior between user organizations has not been defined. There is a strong reliance on conventional Hanford operations and tank farms chemical processing assumptions in developing waste stream compositions and uncertainties.

While an intensive history matching effort has not been undertaken, some limited comparisons between model results and historical data have been performed. Accordingly, despite its limitations, the Hanford SIM Rev. 1 results reported by Corbin et al. (2005) are the best available information on contaminant releases to the waste sites simulated. Testing of Hanford SIM Rev. 2 has substantiated those earlier published results with only minor and specific differences (see Section 5).

2.0 Project Requirements

The principal project requirement for the Hanford Soil Inventory Model (SIM) Rev. 2 was to duplicate the capabilities and replicate the results from the Hanford SIM Rev. 1 effort (Corbin et al. 2005) and make those capabilities broadly available using commercial software. The principal project requirement of Hanford SIM Rev. 1 was to provide comprehensive quantitative estimates of contaminant inventory and its uncertainty for the various liquid waste sites, unplanned releases, and past tank farm leaks as a function of time and location at Hanford. As a result of the Hanford SIM Rev. 1 effort, a computer model was developed capable of performing these calculations and providing satisfactory quantitative output representing a robust description of its inventory and uncertainty for use in other subsequent models. Other requirements were identified from the initial project guidance (DOE-RL 1999):

- Use process chemistry models, historical records, and currently available field data to develop radionuclide and chemical inventories.

- Focus on the priority list established through the System Assessment Capability (SAC); the solution should be able to increase the list of contaminants of interest as individual project needs and SAC requirements evolve.

- Use probabilistic modeling to describe uncertainty.

- Report inventory with standard deviations for the Hanford 100, 200, and 300 Areas.

- Report inventory cases that represent maximum inventories associated with each specific waste type.

- Provide the ability to reconcile field data and model predictions.

This document summarizes software requirements, design, and limitations for Hanford SIM Rev. 2. This document is meant to also provide technical support of the results presented in Corbin et al. (2005). The application described in Corbin et al. (2005) has now been essentially duplicated using a version of Hanford SIM created under DOE Order 414.1C for safety software. A companion report (Simpson et al. 2007) provides a user's guide for Hanford SIM Rev 2.

Neither the purpose nor the goal of this model required specific history matching for site inventories as part of this effort (i.e., model results are not fitted to published data). History matching is a proposed future effort. Furthermore, the fiscal year 2007 task was not to provide corrections or additions to Hanford SIM Rev. 2 inputs or outputs. Disagreement of the model with reference values or inconsistent behavior between historically similar sites in the model or observed in the reference values are causes for further investigation with respect to the model system bases and the reference data. Evaluation of these disagreements is part of the error recovery guidance process. Maintaining and revising the model and its results is part of the configuration management and change control process.

2.1 Functional Requirement Description and Evaluation for Hanford SIM

The ability to use familiar, commercially available software on high-performance personal computers for data input, modeling, and analysis, rather than custom software on a workstation or mainframe computer for modeling, was preferred. The proof-of-principle task documented in Simpson et al. (2001) led to the development and application effort described in Corbin et al. (2005). Because of changes in the core software applications and evolving user requirements, the tasks in fiscal year 2007 were to update the Hanford SIM Rev. 1 portfolio of applications (i.e., employ available COTS software), make the modeling tools available to the Hanford technical community, and document the development of the software.

Several quantitative tools/programs such as sensitivity analysis, multivariate statistical models, and stochastic simulation, were considered to represent the disposal situation at the Hanford Site and provide analysis of the contaminant inventories discharged to ground. Each modeling method that was evaluated had its advantages and disadvantages, and each method was evaluated in context for this particular application (Corbin et al. 2005).

Stochastic simulation was chosen because the modeling parameters for this calculation did not have satisfactory closed-form definitions to approach the problem from a purely mathematical standpoint; the available waste stream/site data were not sufficiently comprehensive to apply regression analysis; and the desire for a comprehensive description of uncertainty eliminated sensitivity analysis as potential methods for analysis. Stochastic simulation is a broadly accepted modeling technique that meets the requirements of the task. Furthermore, substantial resources are available for its application in practice; therefore, this method was used in developing Hanford SIM Rev. 2. Because the objective of this task was to provide an approach to estimate mass-balanced inventories and their uncertainties for the Hanford Site post-closure setting, and appreciating the limitations of the other methods under consideration, a stochastic Monte Carlo simulation technique was selected to provide estimates of inventory and its uncertainty.

2.2 Supplemental Project Implementation Decisions

Several stochastic simulation options were considered for model development, and the Open Crystal Ball (OCB) statistical package (Decisioneering 2002) was selected. The OCB software provided an appropriate development platform with which to construct a model that could accommodate the scope and requirements associated with this task (i.e., compute the annual inventories and uncertainties for several hundred waste sites for 75 analytes over a 50-year timeframe, using approximately 200 waste streams to describe the various discharges that occurred). Updated versions of Crystal Ball and OCB became available in 2006 (Decisioneering 2006a, 2006b). Using these latest software packages and reusing as much of the previous work as possible to efficiently satisfy the project objectives were principal decisions regarding software requirements and software design.

Because there was no a priori method to determine a sufficient number of iterations for this model to ensure statistically valid and repeatable results, an experiment using a test file with a variety of distributions was developed and tested for convergence behavior. This process allowed for determining a sufficiently rigorous number of iterations necessary for the Hanford SIM Rev. 2 results to be stable and repeatable. A separate discussion regarding software testing, verification, and validation processes is presented in Section 5.

2.3 Hardware Requirements

Because of the desire to use conventional personal computers for this task, several hardware-based challenges impede the execution of the model. These challenges are associated with reading and writing the input data, performing large numbers of computations, and managing the output data. Thus, operation of Hanford SIM Rev. 2 is limited predominantly by the amount of available random access memory (RAM) provided in the computer. However, because the Windows XP Professional operating system (Version 2002, Service Pack 2) constrains RAM use to 1.3 gigabytes (GB), more RAM above this limit does not enhance performance. Table 2.1 details the minimum and recommended hardware necessary to operate Hanford SIM Rev. 2. The recommended hardware configuration was used for both development and execution of Hanford SIM Rev. 2.

Table 2.1. Hardware Requirements for Hanford SIM Rev. 2 Operation

Minimum Required Hardware to Operate/Execute Hanford SIM Rev. 2	Recommended Hardware to Operate/Execute Hanford SIM Rev. 2
Intel Pentium 4 system, with a clock speed of at least 2.53 GHz (Hanford SIM Rev. 2 has not been tested on the AMD platform)	Several Intel-based Pentium 4 systems with clock speeds greater than 3.0 GHz
1 GB of RAM	2 GB of RAM
1 GB of free hard drive space	5 GB of free hard drive space
Smaller than a 20-in. monitor	Greater than a 20-in. monitor
PC case operating with the original equipment manufacturer's installed fan	PC case with room for operating at least four case fans
	UPS
DVD-R	DVD-RW
	USB mass storage drive

AMD = Advanced micro devices.
DVD = Digital video disc.
GB = Gigabytes.
GHz = Gigahertz.
PC = Personal computer.
RAM = Random access memory.
SIM = Soil Inventory Model.
UPS = Uninterruptible power supply.
USB = Universal serial bus.

Computers meeting the minimum requirements can be used to run Hanford SIM Rev. 2, but the run times for the simulations become exceedingly long. A complete converged model run (assuming a typical 2005 model configuration) using the recommended hardware configuration distributed over four computers requires more than 100 hours of chronological time or more than 400 machine-hours of computing time. Therefore, using a single machine to execute a simulation as defined would require nearly three weeks of continuous operation to complete.

The amount of time necessary to complete a simulation varies as a function of the number of trials, sites, and analytes being evaluated. However, other than for relatively simple troubleshooting situations, these models are very demanding with regard to the amount of time they require to perform an analysis.

2.4 Software Requirements

The minimum off-the-shelf software requirements for performing calculations using the current Hanford SIM Rev. 2 and its associated infrastructure are as follows:

- Windows XP Professional,[1] Version 2002, Service Pack 2, provides the operating system for the computer. Appendix A contains a list of the software maintenance patches incorporated as part of the development and production environments.

- Microsoft.NET 1.1[1] provides the application environment. The software maintenance patches incorporated as part of the development and production environments are listed in Appendix A.

- Crystal Ball[2] v7.2 (Professional Edition) provides the ability to evaluate scenarios using macros as part of the quality assurance infrastructure and is key for the OCB runtime license. Hanford SIM Rev. 2 will not install or operate without a licensed copy of Crystal Ball 7 Professional Edition installed.

- Open Crystal Ball[2] v2.0 (OCB.dll) provides the computational engine to perform the stochastic calculations.

- C# interface (**OCBHanford[3]**) administers the simulation by managing inputs and outputs through OCB.

- Microsoft Excel[1] 2003 is the user interface for data input/output and analysis. Appendix A contains a list of the software maintenance patches incorporated as part of the development and production environments.

- Microsoft Visual Source Safe maintains source code configuration control.

During fiscal year 2007, while Hanford SIM Rev. 2 was being created, Microsoft issued two new operating systems in which Microsoft Excel[1] is embedded—Microsoft Vista[1] and Microsoft Office 2007.[1] This product, Hanford SIM Rev. 2, has been designed and tested to function properly using Microsoft Office 2003[1] and, specifically, Microsoft Excel 2003.[1]

---

[1]Software product of Microsoft Corporation, Redmond, Washington.
[2]Software product of Decisioneering, Denver, Colorado.
[3]**OCBHanford** is part of the Hanford SIM model and not a vendor product.

2.5 Project Requirements Translated to Software Requirements

The software requirements for Hanford SIM Rev. 2 at a high level are summarized as follows:

1. Use a Monte Carlo approach to achieve a probabilistic model.
2. Compute annual inventories, volumes, and waste concentrations.
3. Simulate 75 analytes including chemicals and radionuclides.
4. Simulate the period between Hanford Site startup and present day, more than 50 years of operation.
5. Simulate several hundred waste sites and unplanned releases using approximately 200 waste streams.
6. Be able to simulate 25,000 realizations.
7. Be able to accept the following input parameter distributions: normal, triangular, lognormal, exponential, beta, gamma, Weibull, zero, and unity.
8. Provide results decay corrected to 1 January 2001.
9. Simulate and report the mean, minimum, maximum, standard deviation, median, and the following percentiles: 0.5%, 5%, 10%, 15%, 20%, 25%…85%, 90%, 95%, and 99.5%.
10. Compute the mass or activity of the liquid and entrained solids using the general equation

$$I = \rho * C * V * CF \tag{2.1}$$

where
- I = inventory, in mass or activity (typically kilograms or curies)
- ρ = density, as mass/volume (typically in grams per milliliter)
- C = concentration, as analyte mass/total mass (typically micrograms/gram)
- V = volume (typically in megaliters)
- CF = correction factor, dimensionless, used to scale units.

11. Employ a highly modular architecture with three principal elements—a C# user and application interface code, an Open Crystal Ball calculation engine, and data input/output interface provided by Microsoft Excel.

The software life-cycle documentation associated with Hanford SIM Rev. 2 provides greater detail on software requirements.[4]

[4]Anderson MJ, BC Simpson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Software Requirements Specification*. VIV07-33573-SRS-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.

3.0 Mathematical Framework and Model Design

Stochastic simulation (or Monte Carlo) models typically use random number generators to draw samples from probability distributions and perform calculations. The objective of this simulation method is to quantify the uncertainties of the dependent variables based on the assumed uncertainties of a set of independent variables, when the relationships between the dependent and independent variables are too complex for an analytical solution. This method was considered to be the most appropriate for the task presented, but this method also has its limitations:

- The independent variables identified in the analysis may not actually be independent.
- The probability distributions assumed for the independent variables often are subjectively assigned and may not reliably describe historical actions.
- The number of iterations necessary to provide statistically valid results for the simulation is usually not known a priori; therefore, an evaluation of the results to demonstrate model repeatability and stability is necessary.

The theory underlying the Monte Carlo method of stochastic simulation used in Hanford SIM Rev. 2 is briefly addressed in the following text. Monte Carlo is the method of approximating an expectation by the sample mean of a function of simulated random variables. It is about invoking laws of large numbers to approximate expectations. In mathematical terms, consider a random variable X having probability mass function or probability function $f_X(x)$, which is greater than zero on a set of values χ. Then, the expected value of a function g of X is

$$\mathbb{E}(g(X)) = \sum_{x \in \mathcal{X}} g(x) f_X(x) \tag{3.1}$$

if X is discrete and

$$\mathbb{E}(g(X)) = \int_{x \in \mathcal{X}} g(x) f_X(x) dx \tag{3.2}$$

if X is continuous. Now, if an *n*-sample of independently generated Xs, (e.g., individual outcomes of the random variable: x_1, x_2, x_3, …x_n) and the mean of g(x) is computed over the sample, then that would result in the Monte Carlo estimate

$$\tilde{g_n}(x) = \frac{1}{n} \sum_{i=1}^{n} g(x_i) \tag{3.3}$$

of $\mathbb{E}(g(X))$.

Alternately, the random variable,

$$\tilde{g_n}(X) = \frac{1}{n} \sum_{i=1}^{n} g(X) \tag{3.4}$$

can be considered the Monte Carlo estimator of $\mathbb{E}(g(X))$.

If $\mathbb{E}(g(X))$ exists, then the weak law of large numbers indicates that for any arbitrarily small ϵ,

$$\lim_{n \to \infty} P\left(|\tilde{g_n}(X) - \mathbb{E}(g(X))| \geq \epsilon\right) = 0. \tag{3.5}$$

Equation (3.5) indicates that as *n* becomes large, there is a small probability that $\tilde{g_n}(X)$ deviates much from $\mathbb{E}(g(X))$. For this task, the weak law of large numbers says that so long as *n* is large enough, $\tilde{g_n}(X)$ arising from the Monte Carlo calculation shall be as close to $\mathbb{E}(g(X))$ as desired. For further detail regarding Monte Carlo methods, a principal reference cited in the Crystal Ball documentation is Hammersley and Handscomb (1964).

There are several variations of stochastic calculations. In this case, OCB has two methods of simulation, Monte Carlo and Latin hypercube sampling (LHS; Iman and Conover 1982). By selecting Monte Carlo, the calculation will proceed using a simple random sampling method. The random behavior in games of chance is similar to how the Monte Carlo simulation selects variable values at random throughout the selected probability distribution to simulate a model.

The LHS variation of this calculation works by segmenting the assumed probability distribution into a number of nonoverlapping intervals, each having equal probability, as shown in Figure 3.1. Thus, LHS

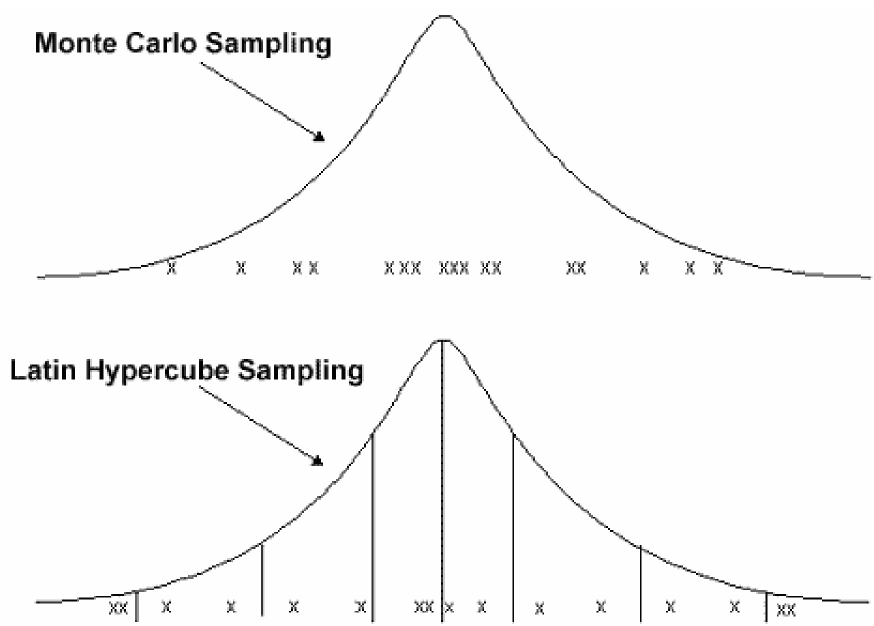


Figure 3.1. Monte Carlo Sampling vs. Latin Hypercube Sampling

simulation can provide convergence to a theoretical result faster than the simple random sampling Monte Carlo simulation for a given number of trials. However, demands on computing resources are higher for LHS (memory usage is higher and run-time performance is slower in the LHS simulation), and there is no guarantee of improvement (i.e., faster convergence).

3.1 Hanford SIM Rev. 2 Conceptual Model

The Hanford SIM Rev. 2 conceptual model is relatively straightforward:

1. Review and select source data and model boundaries.
2. Develop, configure, and test model inputs.
3. Develop, administer, and perform model calculation/simulation.
4. Report model calculation results.
5. Perform model quality assurance and error correction on model.
6. Refine model elements as necessary.

Evaluation and execution of the conceptual model elements often was performed concurrently and iteratively during the development of Hanford SIM Rev. 2. The following narrative sections present how implementation of the various portions of the concept increased in sophistication as the model evolved and not necessarily the sequential progression of the model's development.

3.2 Inventory Equation Descriptions

Hanford SIM Rev. 2 executes a linear equation that computes the mass or activity of a particular constituent. The general form of this equation is

$$I = \rho * C * V * CF \qquad (3.6)$$

Inventory (I) = density*concentration*volume*correction factor

Because in some cases entrained solids are included as part of the overall inventory, both liquid and solid phases of a waste stream must be computed, resulting in a slightly more complicated version of the equation:

$$I = \rho_l * C_l * V_t * (1 - V_s) * CFL + \rho_s * C_s * V_t * V_s * CFS \qquad (3.7)$$

Inventory = density (liquid)*concentration (liquid)*total volume*(1-volume percent solids)*correction factor (liquid) + density (solids)*concentration (solids)*volume percent solids*correction factor (solid).

- Inventory is the calculated output and is reported in kilograms (kg) or curies (Ci).
- Density (ρ_l and ρ_s) is the bulk density used to describe each waste stream phase (liquid or solid) reported in grams per milliliter (g/mL).
- Concentration (C_l and C_s) is the analyte amount per unit mass in each waste stream phase reported in micrograms per gram (µg/g) or microcuries per gram (µCi/g).

- Volume (V_t) is the total discharged amount reported in megaliters (ML).

- Volumetric solids content (V_s) are the estimated or assumed contribution of a solid to a particular waste stream (i.e., V_s = volume of solids/V_t) represented as a decimal (dimensionless).

- Correction factors applied to the liquid and solid phases (CFL and CFS) are the scalar multipliers used to provide inventory unit consistency. In this case, the units in calculating the amounts for the chemicals result in kilograms. Thus the correction factor is 1; the calculation for the radionuclides must be multiplied by 1E-06 to discount the inflation factor used to compensate for certain small radionuclide concentrations and provide output in curies.

This form of the equation was selected based on the observed prevalence of the units associated with the analytical data, and these parameters are presented also in the Hanford Defined Waste (HDW) Model waste stream descriptions (Higley et al. 2004).

3.3 Constraint Conditions of Hanford SIM Rev. 2

The following list summarizes the definitions, assumptions, and constraint conditions used by the Hanford SIM Rev. 2 system for modeling bases, data integrity, and uncertainty development. These modeling elements and their development are described in more detail in Corbin et al. (2005). Because of the flexible architecture of Hanford SIM Rev. 2, these constraint conditions usually can be modified or relaxed to accommodate specific situations or different environments as needed.

The application of a minimum basis set of waste streams is assumed to be appropriate and sufficient to describe disposal site inventories. A minimum basis set of waste streams is assumed to be a Hanford SIM Rev. 2 modeling boundary condition. This assumption has a two-fold purpose: it 1) keeps the model from getting unwieldy in size and 2) forces critical evaluation of the waste stream-disposal site environment. A model is not useful and does not explain much if there is no common behavior to exploit consistently and quantitatively to describe various observations. Hanford SIM Rev. 2 disagreements with reference values or inconsistent behavior between historically similar sites in the model are cause for further investigation with respect to the model system basis set of waste streams and the reference data.

Waste management procedures and operating conditions are (or have been) reasonably consistent throughout Hanford Site processes. Processes represented in Hanford SIM Rev. 2 are assumed to be well defined, and Hanford Site personnel are assumed to have conducted waste management operations within control specifications, ensuring consistency in model treatment between geographically and chronologically separated disposal sites receiving similar wastes (e.g., coupling between sites receiving the same waste stream).

Comprehensive waste stream compositions, such as HDW Model Rev. 5 waste stream definitions (Higley et al. 2004), were used where possible, and analyte correlations were maintained. The comprehensive methodology of the modeling process and calculation is documented in Agnew (1997), with the latest modifications and refinements presented in Higley et al. (2004). Alignment with Hanford tank farms regarding HDW Model stream compositions and chemistry assumptions is maintained, using contemporary sampling data sparingly.

In addition to being comprehensive in description, enforcing consistent solubility behavior, and minimizing circularity in Hanford SIM Rev. 2, the HDW Model waste streams have several desirable features: they are internally mass- and charge-balanced, the individual solubilities of the various analytes are specified (and can be modified as data dictate), radionuclides have been decayed to a common date (January 1, 2001), and the overall system inputs are mass-balanced. This condition enforces internal analyte chemical and radionuclide correlations. Therefore, with regard to HDW Model waste streams, mean site inventories calculated with Hanford SIM Rev. 2 are strictly non-negative and mass- and charge-balanced. Waste streams with compositions derived from the HDW Model maintain this characteristic. Waste streams obtained from surveillance data are used largely as-is, with no incorporation of unquantified analytes, unless there is a physical rationale and means to do so, such as secular equilibrium or maintaining isotopic ratios. Mass and charge balance are characteristics of the inputs and are propagated through Hanford SIM Rev. 2.

In addition, numerous waste streams associated with disposal to the vadose zone were never discharged to the tanks and, therefore, were never defined in the HDW Model. Because of the lack of comprehensive data in the available waste stream surveillance information in those cases, missing values often are derived from the HDW Model using various waste stream compositions as a partial basis. Losses of less than 1% of total inventory usually are assumed for the source amounts as reasonable estimates for contamination (Hanthorn 1957) of lightly contaminated waste streams (e.g., cooling water and steam condensates).

Simplicity in describing waste stream–waste site input allocations/contributions was maintained throughout model development, within known physical/chemical limits. Hanford SIM Rev. 2 uses the available reference/surveillance data to establish waste stream assignments, waste stream volumes, and the potential for entrained solids, depending on the inventories reported. In general, the simplest description of the site disposal operations that best aligned with the qualitative site information, inventory description, and physical/chemical boundary condition is the one selected, especially with regard to entrained solids.

Most of the model distribution parameters do not have any intrinsic behavior that is highly extreme (e.g., asymptotically or discontinuously approaching zero or infinity). Where some of the lognormal input distributions used extend over several orders of magnitude, a truncation rule using a fitted line was used to constrain their behavior in the right-hand tail.

Contamination control measures and physical constraints in place generally prevented the loss of solids from the tank–canyon system (Corbin et al. 2005). Very few waste streams disposed to the past-practice waste sites are considered to possess solids because of the waste management and surveillance practices employed during production operations and the general physical constraints of the system with regard to particulate entrainment (radiation monitors, settling tanks, no agitation, passive filtration, and so on in the tank–canyon system).

Waste stream compositions are as independent as practicable and minimize direct circularity in applying reference data values to modeling inputs. Use of the HDW Model Rev. 5 definitions for composition information partially addresses the difficulty of circularity with respect to Hanford SIM Rev. 2 inputs/output and the reference data. Because the HDW Model was developed in a manner far removed from Hanford SIM Rev. 2 (even though they share some common references), the HDW Model is considered an independent source of composition information.

The various input variables are assumed to be independent mathematically. Because of the methods used to obtain the HDW concentration inputs, no further correlation corrections are imposed in Hanford SIM Rev. 2, and each parameter/analyte is considered to be an independent variable in the modeling calculations. This assumption is key in the mathematics of the Monte Carlo calculation.

In the case of the inventory calculation executed in Hanford SIM Rev. 2, there are no direct dependencies or relationships between the principal variables (e.g., density, concentration, volume). However, a substantial limiting assumption for Hanford SIM Rev. 2 is that the liquid waste compositions of the HDW Model waste streams are at saturation. The application of the current uncertainty definitions can carry the upper values for many analytes higher than the solubility limit. However, in evaluating the context and environment of the waste streams, especially because of the potential variations in temperature, interactions of solution equilibria, and microprecipitate entrainment, these potential ranges in composition resulting from the application of the uncertainty definitions are considered reasonable, and the assumption of independence for these variables holds.

The upper range of uncertainty values for the solids can result in extreme concentration values for certain analytes because speciation is not assumed and a wide variety of species could be present. With the relatively low number of sites where solids are present, this assumption regarding solids composition behavior is necessary and reasonable and does not appear to degrade the results significantly.

Alignment with available surveillance data with regard to waste stream–disposal site volume assignments and inventory values is maintained where possible. Extensive data from the various plant technical manuals, numerous process engineering memoranda, and surveillance data are used to derive and/or assign waste stream compositions, define waste site operations, and assign waste streams and associated volumes for a particular site. The references in Corbin et al. (2005) enumerate the various sources of technical and operational data used to establish and define the variables used to calculate inventories at the various waste sites, and Hanford SIM Rev. 2 maintains this alignment with the available data where appropriate.

The specified campaign subdivisions for the ORIGEN2 reactor production data (Watrous et al. 2002) are assumed to be appropriate groupings for defining uncertainty behavior as a function of time for the various radionuclides in each separation process and their associated discharges. Process phasing and changes in operating philosophy are clearly evident as a function of time when reviewing the data for developing Hanford SIM Rev. 2 inputs (e.g., the timing of production and introduction of different operating procedures and fuel affected the amounts of specific analytes sent to the ground).

ORIGEN2 production data are grouped together on the same basis as the HDW Model Rev. 5 separations. Thus, the Hanford SIM Rev. 2 input structure is designed to reduce potential cross-contamination or cross-talk between processing regimes and aid in enforcing the overall and individual mass balance boundary conditions. This structure is dictated by how production and waste management operations were conducted at the Hanford Site—waste management practices segregated wastes in predictable ways, and the development and definition of the inputs mirrored those practices.

The uncertainties defined for the radionuclide concentrations are assumed to be well described by the ORIGEN2 distribution curve fits (radionuclides) and are not substantially confounded by solubility behavior. The uncertainty definitions assigned from curve fits of the ORIGEN2 production data often have substantial ranges, and these definitions were assumed to encompass the broad range of behavior observed for these analytes. However, the production variability is acknowledged to be potentially confounded with the chemical behavior (solubility) of the various species.

The interaction of chemical behavior, thermodynamic properties, and the dynamic chemical conditions in the tanks results in very large uncertainties for most radioactive species in these waste streams. Limited literature data on the behavior of these species under the waste stream/tank storage conditions (alkaline, with moderate to high-ionic strength, multicomponent solutions) could be used to define an independent set of uncertainty distributions. Furthermore, the Crystal Ball data-fitting treatment of the derived distributions could be considered the most conservative quantification interpretation for this distribution because the lower limit in this treatment always includes zero, resulting in broader uncertainties. Therefore, the derived distributions are considered appropriate uncertainty representations for these analytes.

The inter-batch variability for a specific waste stream is assumed to be encompassed by the selected uncertainty definition. The separation processes are assumed to have been operated within specifications, and abrupt changes in waste stream compositions and/or uncertainties are represented by new waste streams. Although there is evidence of modest process evolution, most of these changes do not result in practical changes to waste composition during the selected campaign timeframe, and the batch-to-batch variability is assumed to be encompassed by the assigned uncertainty. Thus, the waste streams are assumed to not change rapidly over time, and the mean waste stream analyte concentrations are present in fixed ratios to each other within a specific uncertainty regime.

3.4 Model Input Data Requirements

The data requirements dictate that the inputs used in Hanford SIM Rev. 2 (e.g., site volumes, waste stream compositions, densities) can be appropriately assigned and quantitatively described using the available distributions—and are technically defensible. An assumption used in the Hanford SIM Rev. 2 input files that is entered for use in OCB is the type of distribution and its corresponding quantitative description (e.g., expected value and range) for each variable (e.g., volume, volumetric solids content, density, concentration).

Each data input is considered an independent variable; therefore, the waste stream composition/ properties and waste stream discharge histories for the waste disposal sites could be examined and developed using a variety of source data and assumptions that do not necessarily have impacts on the other modeling variables. The references cited in Corbin et al. (2005) provide a broad spectrum of process engineering, modeling, and historical waste management data that were used in developing the inputs and represent a reasonable example of populating a model of this type.

Distributions for modeling parameters and their quantitative descriptions are assigned by a variety of methods. The distributions are interpreted by the OCB.dll by the "distribution type" index and the associated parameters, as seen in Table 3.1 (parameter 1, parameter 2, parameter 3, and parameter 4, which are different, depending on the distribution). All input cells in the Microsoft Excel spreadsheet must be filled with the appropriate values to define a distribution, or if a distribution does not use four parameters, zero (0) must be entered in the remaining cells to allow the simulation calculations to proceed. Table 3.2 provides a simple illustration of the distribution and the typical conditions or variables represented in modeling situations.

Table 3.1. Available Distribution Parameter Definitions in Hanford SIM Rev. 2

Distribution Type Index	Distribution Name	Parameter 1	Parameter 2	Parameter 3	Parameter 4
0	Normal	Mean	Standard deviation	0 = unconstrained; value = low cut-off	0 = none
1	Triangular	Minimum	Mode	Maximum	0 = none
4	Lognormal	Mean	Standard deviation	0 = unconstrained; value = high cut-off	0 = none
6	Exponential	Rate	0 = none	0 = none	0 = none
8	Weibull	Location	Scale	Shape	0 = none
9	Beta	Alpha	Beta	Maximum	Minimum
12	Gamma	Location	Scale	Shape	0 = none
17	Zero	0	0 = none	0 = none	0 = none
18	1 (unity)	1	0 = none	0 = none	0 = none

3.5 Hanford SIM Rev. 2 Input Data Structure

Hanford SIM Rev. 2 has an input data structure represented as a series of matrices. These matrices use Microsoft Excel worksheets as a user interface for data input. This section provides guidance on where the variables reside in the user interface. Four worksheets are used to collect and organize input data in the Hanford SIM Rev. 2 production workbook (example: *SIMInput_Base*). They are named *SiteInput, AnalyteInput, DensityInput,* and *CorrFactors* and are listed in Table 3.3 to the right of the input matrices. Table 3.3 describes the structure and location of the data. These worksheets contain the quantitative information describing the input values and the corresponding distribution definitions used in the model. They represent the basis for calculation.

Table 3.2. Hanford SIM Rev. 2 Continuous Distributions

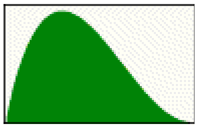
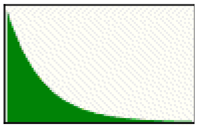
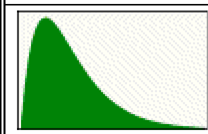
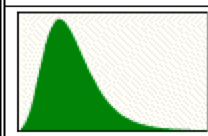
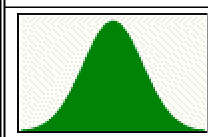
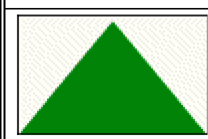
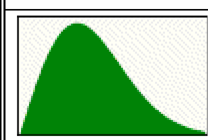
Distribution	Technical Summary
 Beta	The beta distribution is a very flexible distribution commonly used to represent variability over a fixed range. It can represent uncertainty in the probability of occurrence of an event. It is used also to describe empirical data and predict the random behavior of percentages and fractions (Decisioneering 2007, p. 274).
 Exponential	The exponential distribution is widely used to describe events recurring at random points in time or space, such as the time between failures of electronic equipment, the time between arrivals at a service booth, or repairs needed on a certain stretch of highway. It is related to the Poisson distribution, which describes the number of occurrences of an event in a given interval of time or space (Decisioneering 2007, p. 278).
 Gamma	The gamma distribution applies to a wide range of physical quantities and is related to other distributions: lognormal, exponential, Pascal, Erlang, Poisson, and chi-squared. It is used in meteorological processes to represent pollutant concentrations and precipitation quantities. The gamma distribution is also used to measure the time between the occurrence of events when the event process is not completely random. Other applications of the gamma distribution include inventory control, economics theory, and insurance risk theory (Decisioneering 2007, p. 280).
 Lognormal	The lognormal distribution is widely used in situations where values are positively skewed, for example in financial analysis for security valuation or in real estate for property valuation (Decisioneering 2007, p. 285).
 Normal	The normal distribution is the most important distribution in probability theory because it describes many natural phenomena, such as people's IQs or heights. Decision-makers can use the normal distribution to describe uncertain variables such as the inflation rate or the future price of gasoline (Decisioneering 2007, p. 290).
 Triangular	The triangular distribution describes a situation in which the minimum, maximum, and most likely values to occur are known. For example, the number of cars sold per week when past sales show the minimum, maximum, and usual number of cars sold could be used to describe the anticipated behavior (Decisioneering 2007, p. 294).
 Weibull	The Weibull distribution describes data resulting from life and fatigue tests. It is commonly used to describe failure time in reliability studies and the breaking strengths of materials in reliability and quality control tests. Weibull distributions also are used to represent various physical quantities, such as wind speed (Decisioneering 2007, p. 299)
Images and text used with permission of Oracle, Inc. For a greater understanding of each type of distribution and its definition, refer to the Crystal Ball User Manual installed as part of the software, available at http://www.crystalball.com.	

Table 3.3. Input Data Structure and Location in Hanford SIM Rev. 2

Model Input Matrix	Workbook Location
$\mathbf{CL}_{(i,k)}$: concentration liquid matrix (µg/g or µCi/g)	*AnalyteInput* worksheet
$\mathbf{CS}_{(i,k)}$: concentration solid matrix (µg/g or µCi/g)	*AnalyteInput* worksheet
$\mathbf{TV}_{(j,k,l)}$: total volume matrix (ML)	*SiteInput* worksheet
$\mathbf{VP}_{(j,k,l)}$: volumetric solids matrix (dimensionless)	*SiteInput* worksheet
$\mathbf{CFL}_{(i)}$: correction factor liquid matrix	*CorrFactor* worksheet
$\mathbf{CFS}_{(i)}$: correction factor solid matrix	*CorrFactor* worksheet
$\mathbf{DL}_{(i,k)}$: density liquid matrix (g/mL)	*DensityInput* worksheet
$\mathbf{DS}_{(i,k)}$: density solid matrix (g/mL)	*DensityInput* worksheet

i = number of chemicals or radionuclides	$I = 1, i_{max}$	i_{max} = 75 analytes
j = number of sites	$J = 1, j_{max}$	j_{max} = 377 total sites
k = number of waste streams	$K = 1, k_{max}$	k_{max} = 196 waste streams
l = years of operation	$L = 1944, l_{max}$	l_{max} = 2001 calendar year

The inventory calculations follow the example below and illustrate the correspondence of how the matrices relate to Equation (3.7) as part of executing the Monte Carlo simulation. Each parameter has an input distribution for each i, j, k, and l that serve as inputs to the simulation. A random selection from each independent input distribution is then used to calculate inventory, and the resulting output matrices are computed.

$\mathbf{FL}_{(i,j,l,)}$: inventory forecast liquid matrix for a specific site–analyte–year calculated over a waste stream, *k* (kilograms or curies);

$$\mathbf{FL}_{(i,j,l)} = \mathbf{CL}_{(i,k)} * \mathbf{DL}_{(i,k)} * \mathbf{TV}_{(j,k,l)} * [1-\mathbf{VP}_{(j,k,l)}] * \mathbf{CFL}_{(i)} \quad (3.8)$$

$\mathbf{FS}_{(i,j,l)}$: inventory forecast solid matrix for a specific site–analyte–year calculated over a waste stream, *k* (kilograms or curies);

$$\mathbf{FS}_{(i,j,l)} = \mathbf{CS}_{(i,k)} * \mathbf{DS}_{(i,k)} * \mathbf{TV}_{(j,k,l)} * \mathbf{VP}_{(j,k,l)} * \mathbf{CFS}_{(i)} \quad (3.9)$$

$\mathbf{FT}_{(i,j,l)}$: inventory forecast total matrix for a specific site–analyte–year calculated over a waste stream, *k* for both phases (kilograms or curies);

$$\mathbf{FT}_{(i,j,l)} = \mathbf{CL}_{(i,k)} * \mathbf{DL}_{(i,k)} * \mathbf{TV}_{(j,k,l)} * [1-\mathbf{VP}_{(j,k,l)}] * \mathbf{CFL}_{(i)} + \mathbf{CS}_{(i,k)} * \mathbf{DS}_{(i,k)} * \mathbf{TV}_{(j,k,l)} * \mathbf{VP}_{(j,k,l)} * \mathbf{CFS}_{(i)} \quad (3.10)$$

Equations (3.11) through (3.16) illustrate the comprehensive inventory forecast calculations over all contributing waste streams for a site–analyte–year. The binning of the various outcomes to determine the forecasted results for each analyte, site, year, and operating history is described in more detail in Section 4.3.

Deterministically,

$$\mathbf{FL}_{i,j,l} = \mathbf{CFL}_i * (\Sigma_k \mathbf{CL}_{i,k} * \mathbf{DL}_{i,k} * \mathbf{TV}_{j,k,l} * [1\text{-}\mathbf{VP}_{j,k,l}]) \tag{3.11}$$

$$\mathbf{FS}_{i,j,l} = \mathbf{CFS}_i * (\Sigma_k \mathbf{CS}_{i,k} * \mathbf{DS}_{i,k} * \mathbf{TV}_{j,k,l} * \mathbf{VP}_{j,k,l}) \tag{3.12}$$

$$\mathbf{FT}_{i,j,l} = \mathbf{FL}_{i,j,l} + \mathbf{FS}_{i,j,l} \tag{3.13}$$

Stochastically,

$$\mathbf{FL}_{i,j,l,t} = \mathbf{CFL}_i * (\Sigma_k \mathbf{CL}_{i,k,t} * \mathbf{DL}_{i,k,t} * \mathbf{TV}_{j,k,l,t} * [1\text{-}\mathbf{VP}_{j,k,l,t}]) \tag{3.14}$$

$$\mathbf{FS}_{i,j,l,t} = \mathbf{CFS}_i * (\Sigma_k \mathbf{CS}_{i,k,t} * \mathbf{DS}_{i,k,t} * \mathbf{TV}_{j,k,l,t} * \mathbf{VP}_{j,k,l,t}) \tag{3.15}$$

$$\mathbf{FT}_{i,j,l,t} = \mathbf{FL}_{i,j,l,t} + \mathbf{FS}_{i,j,l,t} \tag{3.16}$$

where t = one trial.

3.5.1 Volume Definition and Parameterization

Volume input data were reviewed and modeling parameters developed (Corbin et al. 2005) for both total volume and volume percent solids. These definitions were converted into a standard electronic format, the *SiteInput* worksheet of the *source data* workbook. The volume assumptions are specific to the site, year, and waste stream that contributed to the inventory and vary between categories (e.g., different volume distribution assumptions are associated with liquid waste disposal volumes, unplanned releases, and tank farm leaks). The complete data record used in the model includes the site label, year, waste stream label, total volume, and volume percent solids, which are entered in the subsequent columns of this worksheet, respectively. Input volumes are provided in megaliters (ML).

The waste site and waste stream indices correspond to the identification number in the *Legend* worksheet. The comprehensive volume definition (total volume and volume percent solids, waste stream assignments, their mean values, and their respective distribution descriptions) has quantitative information about the amount and uncertainty associated with a particular waste stream for each site–year combination. Each site has a unique combination of waste stream and year descriptions assigned. Table 3.4 provides an example of the structure of the volume input matrix as it would appear in the *SiteInput* worksheet.

Table 3.4. *SiteInput* Worksheet Example, Hanford SIM Rev. 2

Legend #	Legend #					Total Volume (ML)					Vol % Solids			
Waste Site Index, s	Waste Stream Index, w	Site	Year	Waste Stream	Dist Type	Parm 1	Parm 2	Parm 3	Parm4	Dist Type	Parm 1	Parm 2	Parm 3	Parm 4
1	45	200-E-100	1945	BiPO4 (BT1) Cool Wtr-Stm Cond	1	0.00219	0.00438	0.00657	0	17	0	0	0	0
65	50	216-A-19	1955	PUREX (P1) Cold Start	1	0.825	1.10	1.38	0	1	0.045	0.09	0.125	0

3.5.2 Waste Stream Definition and Parameterization

After the sites for analysis in Hanford SIM Rev. 2 were selected, the waste streams necessary to compute inventory and uncertainty were defined. Each waste stream has its own qualitative and quantitative description derived from historical process engineering data, assumptions regarding the presence and behavior of various analytes, and the previously developed waste stream values in the HDW Model (Higley et al. 2004). When a waste stream is developed from surveillance data, engineering judgment, or other sources, the analyst must incorporate logical and consistent rules for enforcing mass and charge balance in all cases and isotopic ratios where possible.

The *AnalyteInput* worksheet in the Hanford SIM Rev. 2 production workbook defines the quantitative information about concentration and uncertainty behavior, in micrograms per gram or microcuries per gram, of a specific analyte (or radionuclide) within a waste stream. Table 3.5 presents an example of the structure of the *AnalyteInput* worksheet. All the radionuclide values in the *AnalyteInput* worksheet are inflated by a multiplicative factor of 1E+09 because the OCB.dll calculation engine cannot perform computations on values less than 1E-16; thus, this accommodation was made as part of the development process, and the correction factor for the radionuclides is specified accordingly.

Table 3.5. *AnalyteInput* Worksheet Example, Hanford SIM Rev. 2

Legend #	Legend #				Derivation Worksheet Liquids Input (µg/g or µCi/g; radionuclides *1.0E+9)					Derivation Worksheet Solids Input (µg/g or µCi/g; radionuclides *1.0E+9)			
Waste Stream Index, w	Analyte Index, A	Waste Stream_ Unc	Analyte	Dist - liquid	Parm 1	Parm 2	Parm 3	Parm 4	Dist – Solid	Parm 1	Parm 2	Parm 3	Parm 4
1	1	1C Evap (BT2)	Na	4	8.99E+04	1.40E+04	3.43E+05	0	4	1.78E+05	2.77E+04	6.78E+05	0

In Hanford SIM Rev. 2, the chemical uncertainties were generally parameterized using the HDW Model-derived uncertainties for the various process waste streams and assigned a lognormal distribution. For the radionuclides, regression analysis was used in the curve-fitting process involving the ORIGEN2 data (Watrous et al. 2002) to quantify the radionuclide uncertainty distributions/parameters as they changed over time. The curve-fit algorithm in Crystal Ball was used to quantify the uncertainty parameters in a consistent and technically defensible manner.

These distributions were applied to the appropriate waste stream compositions for use in the stochastic simulation. Beta distributions were assumed for most of the radionuclides to best represent the data from ORIGEN2 for several reasons. They provide non-negative values throughout the data range, and they avoid certain mathematically extreme conditions (i.e., infinities). The ORIGEN2 data do not fit any distribution well; therefore, the beta distribution is as good as any other. As part of the update to this version of the software, certain definitions of the beta distributions that were previously used were not allowed. In these cases, a truncated lognormal distribution was applied to the affected analytes/campaigns for each waste stream. Lists of the affected waste streams and analytes are presented in Section 5.2.5.

The described method was used for Hanford SIM Rev. 2 input to maintain consistency throughout the model. However, the uncertainties associated with the analytes/radionuclides with respect to the individual waste stream compositions can be crafted and applied to the specific inputs to change them as desired or as the available data may dictate.

3.5.3 Density Definition and Parameterization

The *DensityInput* worksheet of the Hanford SIM Rev. 2 workbook defines the density of a specific waste stream in grams per milliliter. In this case, the guiding assumption is that all analytes have the same density within a waste stream phase (e.g., a separate bulk density is assumed for solids and liquids in a waste stream); thus, the density will be defined only by the specific waste stream and phase. Sources for density information include the HDW Model (Higley et al. 2004), historical process engineering data, and subject-matter expertise.

The waste stream index corresponds to the identification number in the *Legend* worksheet with the waste stream label. The mean values and distribution definitions for the supernatants and the solids are defined in the subsequent columns. Table 3.6 presents an example of the *DensityInput* worksheet.

Table 3.6. *DensityInput* Worksheet Example, Hanford SIM Rev. 2

Legend #		Supernatants	Density (g/mL)				Solids	Density (g/mL)			
Waste Stream Index, w	Waste Streams—Current	Dist Type	Parm 1	Parm 2	Parm 3	Parm 4	Dist Type	Parm 1	Parm 2	Parm 3	Parm 4
1	1C Evap (BT2)	4	1.26	0.063		0	4	1.77	0.088	0	0

3.5.4 Correction Factors

The *CorrFactors* worksheet contains scalar values that are used to convert units of the analyte inventories calculated in Hanford SIM Rev. 2 to those desired for use in other models. The unit basis for the chemical analytes allows the correction factor to be 1 for results to be reported in kilograms. The unit basis for radionuclides dictates that the correction factor be 1E-06 to provide for reporting results in curies, after correcting for the 1E+09 inflation factor applied to the inputs. The definition and application of the correction factors are discussed in more detail in Corbin et al. (2005). Figure 3.3 presents an excerpted example.

a	analyte	supernatant	solids
1	Na	1	1
2	Al	1	1
3	Fe	1	1
30	C-14	0.000001	0.000001
31	Ni-59	0.000001	0.000001
32	Ni-63	0.000001	0.000001

Figure 3.2. *CorrFactors* Worksheet Example, Hanford SIM Rev. 2

3.6 Modeling Boundary Conditions and Software Performance Limitations

Several limitations are associated with the modeling assumptions and software performance incorporated in Hanford SIM Rev. 2. Specific history matching between Hanford SIM Rev. 2 results and documented reference values was not a goal of the modeling effort, although some history matching for certain site–analyte combinations was done as part of the Hanford SIM Rev. 1 development. The Hanford SIM Rev. 2 system also includes other limitations:

- Physical and mathematical simplifications of the various behaviors and boundary conditions are necessary to reasonably quantify the model in the software.
- The mass balance and charge balance characteristics of a Hanford SIM Rev. 2 result are a direct function of analyst input—i.e., the mass and charge balance characteristics of waste streams. There are no "rebalancing" mass or charge algorithms in the model.
- Errors in interpretation of the historical process chemistry or site descriptions arise as a result of obtaining discovery information that refutes or illuminates previously unclear or undocumented disposal situations.
- No contemporaneous radionuclide inventories are provided during simulation.
- There is no concurrent decay for radionuclides over time during simulation.
- Backdecay and in-growth corrections are valid only to January 1, 2001 (the ORIGEN-DKPRO decay date).
- No active chemistry model for changes over time is available for changes introduced by decay.
- The software components used (e.g., Microsoft Excel, Crystal Ball) have some intrinsic limitations.
- Precisely quantifying very small (less than 1E-16) numbers with large uncertainties is challenging. This limitation has been partially addressed using the inflation and correction factors discussed in Sections 3.5.2 and 3.5.4.

Furthermore, the Hanford SIM Rev. 2 software was not necessarily intended to correct discrepancies attributable to human error or historical inconsistencies in the reference data, although identification, analysis, and correction of errors is part of the review and quality assurance (QA) process as demonstrated in Table 6-32 of Corbin et al. (2005). These limitations are described more fully in Corbin et al. (2005).

3.7 Input Data Limitations and Boundary Conditions

This model and its results are significantly reliant on the use of previously gathered surveillance data and information derived from process models. This section summarizes the limitations and boundaries dictated by the information used. The following conditions and their consequences are discussed in more detail in Appendix B:

- In evaluating the data (both inputs and results), the magnitude of uncertainty associated with estimates is significant, spanning in some cases an order of magnitude.

- Because of the modeling assumptions, the resulting output distributions were relatively simple—i.e., skewed, non-negative, and monomodal, each with a well-defined central maximum. Model simplification and constraints associated with software coding to make the inputs tractable removed many of the irregularities observed in the source data to condition its use as input to Hanford SIM Rev. 2.

- The principal factors influencing the model output is the degree to which reliable quantitative descriptions could be provided for the inputs and acceptance of current technical conventions as part of the waste definition. The extensive use of the HDW Model (Higley et al. 2004) and latest ORIGEN2-DKPRO output (Watrous et al. 2002), provide a substantial technical foundation for this model, but it is a potential limitation on Hanford SIM Rev. 2 as well.

- The use of tank data or other historical process data in modeling was evaluated closely for appropriateness in each potential application. The effects of time and waste management operations have compromised much of the contemporary tank waste composition data for use in this modeling effort; thus, direct use of current tank sampling data is extremely limited.

- Significant limitations of Hanford SIM Rev. 2 are associated with the assumptions made regarding the presence or absence of an unquantified analyte for a non-HDW Model waste stream and the assumption regarding the separations processes being well defined and operated within specifications.

- Past-practice data collection and recording methods that established the baseline site inventory values are ambiguous. As documented in several references (Healy 1953; Ruppert and Heid 1954; Paas and Heid 1955; Abrams 1956), significant challenges existed in obtaining surveillance data.

- Some reference and surveillance data are restricted from open publication at this time. All the sources are intermittent as a function of time and not comprehensive in scope.

- The application of the mass-balance boundary is that the mean Hanford SIM Rev. 2 values for a specific analyte summed over all disposal sites must be less than or equal to the total losses of that analyte from the tank–canyon system. This action is currently performed by the user as a function of defining the inputs. There is no automated rebalancing function in the software portfolio if a change to a waste stream is made.

- Logical extensions of contemporary waste stream data for analogous (but data-sparse) situations in the absence of early Hanford Site surveillance information are used. When changes in reactor production behavior (e.g., changes to fuel cladding, fuel element design, or reactor operating power) are observed to occur as a function of time while the basic chemical process remains unchanged, new uncertainties based on the production behavior change are derived and assigned to that waste stream, without changing the base waste composition.

- Mis-assignment of waste streams with generic designations but widely varying compositions as a function of their generation and disposal during production is a potentially significant error requiring an independent data source to provide a means of correction. These discrepancies require judgment on the part of the analyst to resolve; further, in certain cases, the decision to accept one source over another may have introduced an error. Thus, there is the opportunity to introduce human error/differences in technical judgment as part of the data input process of Hanford SIM Rev. 2.

4.0 Model Architecture

The design of Hanford SIM Rev. 2 is highly modular, with separate data input and executable files. There are three principal elements to the Hanford SIM Rev. 2 system—**OCBHanford**, the OCB.dll, and the *source data* workbook. The **OCBHanford** C# interface code directs communication between the OCB.dll calculation engine and the user interface and data input provided by Microsoft Excel. A modular architecture was selected to allow for efficiencies in model development and evaluation. Additionally, because run-time performance is a major constraint for models of this type, several design approaches were examined to optimize the speed of the simulation with regard to the available computing resources. A distributed computing feature with the ability to add or remove sites and analytes was developed for use in production to reduce the amount of time necessary to test and generate results.

The Hanford SIM Rev. 2 computing user interface also has three distinct elements. The Mircrosoft Excel production workbook (*source data*) has two of them—the *Setup* worksheet and the *Legend* worksheet. These worksheets provide an interface for the user to define the boundaries and reporting requirements of the simulation. Microsoft Excel was used because there was the desire to use a familiar and broadly available interface for data input and analysis.

The other interface element is **OCBHanford**. It is accessed via a dialog box that activates the simulation. Once the parameters in the *source data* workbook are set, the specific workbook to be used must be opened using **OCBHanford** and the program will execute. The calculation will then proceed as directed and generate outputs until the simulation is completed or interrupted. Table 4.1 briefly describes the various elements, functions, and relationships of the Hanford SIM Rev. 2 system. The software design document[1] defines the use cases and variable classes at work in Hanford SIM Rev. 2.

4.1 Executable Modeling File

Hanford SIM Rev. 2 uses the **OCBHanford** interface application to generate the various output files**.** It is the executable file containing the C# code that interfaces with the *source data* workbook, which contains all of the inputs and the OCB.dll, which creates the probability distributions and performs the inventory calculations. **OCBHanford** creates the output workbooks and manages the data reporting. The **OCBHanford** dialog box also presents a series of diagnostic data regarding simulation time and computing resources demand that can be useful in gauging hardware suitability and model parameter settings. More detailed discussion of the operation of **OCBHanford** is in the Hanford SIM Rev. 2 user's guide (Simpson et al. 2007).

[1]Anderson MJ, BC Simpson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Software Design Document*. VIV07-33573-SDD-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.

Table 4.1. Relationships and Descriptions of Hanford SIM Rev. 2 Files

Source Workbook/Worksheet File/Sub-File (application)	Function	Reads from	Sends to/Read by
SIMInput_Base/SiteInput (Microsoft Excel)	Source data: waste site definition, waste stream identification, waste volumes, years of operation	None; fundamental input file	**OCBHanford** (Reads)
SIMInput_Base/AnalyteInput (Microsoft Excel)	Source data: waste stream composition definition, solids and liquids	None; fundamental input file	**OCBHanford** (Reads)
SIMInput_Base/DensityInput (Microsoft Excel)	Source data: waste stream density definition, solids and liquids	None; fundamental input file	**OCBHanford** (Reads)
SIMInput_Base/CorrFactors (Microsoft Excel)	Source data: scalar adjustment for unit correction	None; fundamental input file	**OCBHanford** (Reads)
SIMInput_Base/Setup (Microsoft Excel)	Source data: simulation control parameters with regard to program execution	None; fundamental input file	**OCBHanford** (Reads)
SIMInput_Base/Legend (Microsoft Excel)	Source data: simulation control parameters with regard to source data involved	None; fundamental input file	**OCBHanford** (Reads)
OCBHanford C# Interface	Executable file: simulation data management and program administration	*SIMInput_Base/ SiteInput, AnalyteInput, DensityInput, CorrFactor, Setup, Legend*	OCB.dll; C# interface directs OCB.dll to send results to various Operable Unit output files; reads inputs from *SIMInput_Base* and sends results to *SIMInput_Base/ SumFrc/SolFrc/LiqFrc*
OCB.dll	Dynamically linked library: computational engine performing probabilistic inventory and uncertainty calculations	**OCBHanford**	**OCBHanford** C# interface directs OCB.dll to send results to various Operable Unit output files

4.2 Model Approach for Computing and Reporting Output

As the simulation progresses and inventory results are generated, **OCBHanford** reads the various simulation administration parameters and input distribution definitions, creating numerous temporary bins in resident memory as a function of the number of percentiles being reported into which results for each site–year–analyte combination are allocated. Figure 4.1 illustrates the progression of the inventory computation and binning process.

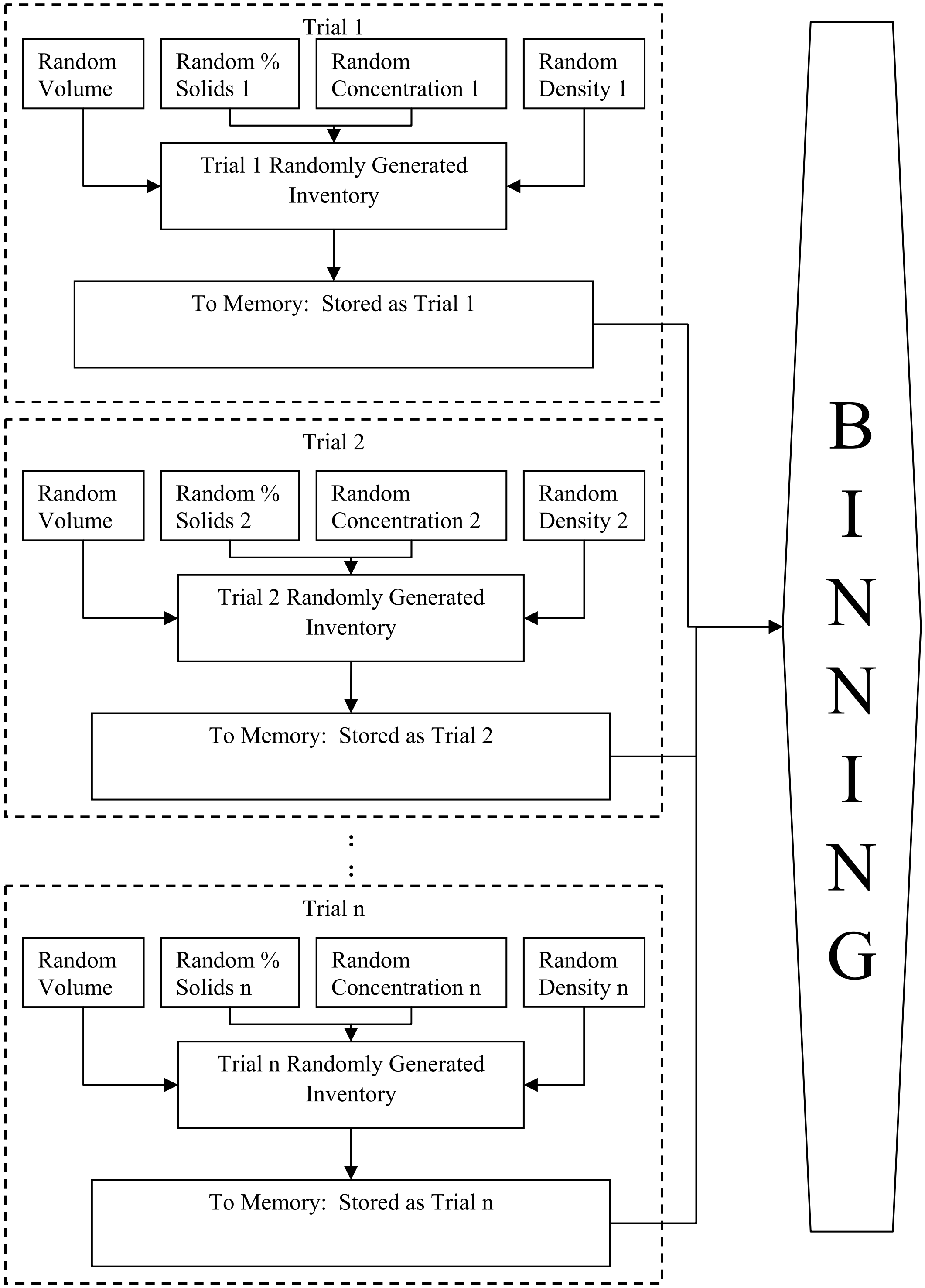


Figure 4.1. **OCBHanford** Calculation Sequence, Inventory Generation, and Binning Process at the Innermost Level

The output statistics for inventory are based on the results in these bins for a site as a function of time and for the level of resolution being reported. Each analyte-year combination is independently calculated and reported for a site; however, the analyte results for a site over time are also accumulated and quantified as well. Thus, Figure 4.1 demonstrates the innermost calculational sequence that proceeds for each waste site, year, and analyte for a given collection of waste sites and contributing waste streams. Additionally, the input distributions for each variable are not the same for each waste stream at all waste sites during the simulation. They are recreated as a function of their presence as part of how the calculational loops are performed during the progression of the simulation.

As part of the calculation of the output inventories, the ordered outcomes are maintained in these bins. This data management process is repeated and maintained at each level of resolution (each site over a number of years, for each site as a function of its operable unit membership, and for the overall system) for each analyte—hence, the need for significant memory and computing power. The resulting summary statistics and percentiles are obtained from the binned outcomes as a function of time, location, and model resolution. Table 4.2 exemplifies this organization and data management process.

Table 4.2. Model Trial Output Organization and Summary Statistical Bases

Site 216-X-001	Trial 1 Analyte Inventory Result	Trial 2 Analyte Inventory Result	Trial 3 Analyte Inventory Result	Trial 4 Analyte Inventory Result	Trial 5 Analyte Inventory Result	Results for Year Summation Bin
1961	a	h	o	v	ac	a,h,o,v,ac…
1962	b	i	p	w	ad	b,i,p,w,ad…
1963	c	j	q	x	ae	c,j,q,x,ae…
1964	d	k	r	y	af	d,k,r,y,af…
1965	e	l	s	z	ag	e,l,s,z,ag…
1966	f	m	t	aa	ah	f,m,t,aa,ah…
1967	g	n	u	ab	ai	g,n,u,ab,ai…
Results for Site Summation Bin	T_1 = a + b + c + d + e + f + g	T_2 = h + i + j + k + l + m + n	T_3 = o + p + q + r + s + t + u	T_4 = v + w + x + y + z + aa + ab	T_5 = ac + ad + ae + af + ag + ah + ai	T_1, T_2, T_3, T_4, T_5,…$T_{25,000}$

Thus, the percentile outcomes for a site over time (or the percentiles for a series of sites in a closure zone) cannot be simply summed and generate the resulting output distribution correctly, although summing the means over a series of years will provide the correct overall site mean. Each site–year–analyte outcome is analyzed over the number of selected trials and the resulting statistics generated. Furthermore, the distributive computing function prevents the creation of bins at the overall site level of resolution; thus, that series of comprehensive outputs can be created only by running a complete simulation on a single machine.

5.0 Software and Modeling Quality Assurance Testing

Testing software operates the software under controlled conditions with both nominal operating conditions and abnormal operating conditions, to 1) verify that the software behaves "as specified" (e.g., design or positive test cases); 2) to detect errors (in execution or user direction), and 3) to ensure that what has been specified as part of the software capabilities is what the user wanted. The abnormal operating conditions (e.g., fault or negative test cases) selected for testing are not unlimited but are restricted to reasonable, credible failures within the specified test conditions and environment. Fault values for input data and other detailed QA information are included in the Hanford SIM Rev. 2 validation and verification test report,[1] which is part of the project file and included in the project records. The software life-cycle documentation associated with Hanford SIM Rev. 2 provides greater detail on general quality assurance requirements.[2]

General testing of COTS software (such as Microsoft Windows OS, Microsoft Excel, Microsoft SQL Server 2000; Decisioneering's Crystal Ball and OCB; GoldSim[3]) was not performed. These commercial-grade items were previously qualified for use and are suitable for performing their general operational, mathematical, and database functions. Only their performance/function as it relates specifically to Hanford SIM Rev. 2 was tested.

5.1 Software Module Testing

This set of initial testing focused on the algorithmic and mathematical performance of individual modules with follow-on processing performed as needed to test the logical or customary progression of data through the simulation and diagnostic tools. This series of tests is described in further detail in the Hanford SIM Rev. 2 test plan,[4] which is part of the project file and included in project records. The modules evaluated during testing were as follows:

- **Volume Balance**—Performs global and year-by-year check on disposal site volumes for reconciliation with reference values. Eight design cases and four fault cases were tested. Each design case processed a number of site single-type volume definitions, having a wide variety of behavior in addition to a file having a blended (e.g., comprehensive) content with several different input definitions. The criteria for passing the tests for each design case were the acquisition of mean values from the input file, sorting the data into the identified categories, and summing the mean volumes for each category. The mean values and sums calculated for each design case were acquired and sorted correctly from the input file, and summing the mean values was arithmetically correct and consistent between categories for each design case. No difference was identified between the overall totals obtained in each design case for each category. The module passed its tests and is suitable for use. Fault cases were simulated to establish module performance. The

---

[1]Simpson BC, MJ Anderson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Validation and Verification Test Report*. VIV07-33573-VVTR-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.
[2]Anderson MJ, BC Simpson, and RA Corbin. 2007. *Hanford Soil Inventory Model Quality Assurance Plan*. VIV07-33573-QAP-01, Rev. 0, Vivid Learning Systems, Pasco, Washington.
[3]GoldSim, a highly graphical, object-oriented computer program for carrying our dynamic probabilistic simulations, is a product of the GoldSim Technology Group LLC, Issaquah, Washington.
[4]Simpson BC, MJ Anderson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Module Test Plan*. VIV07-33573-TP-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.

presence of fault values was found to cause the pivot table function used in this module to switch to identifying category counts instead of sums, indicating a user input error.

- **Site Evaluation**—Performs Crystal Ball simulations of specific site–inventory using the same input file parameters as OCB. This module is usually run during the development and testing of model inputs. Twelve design cases and zero fault cases were tested. The criteria for passing the tests for each design case were the acquisition of inventory equation variable values from the input file, sorting the data into the identified categories (liquid, solid, total), and performing the inventory calculation in Crystal Ball for each category. The values obtained and results calculated for each design case were acquired and sorted correctly from the input file, and the results were arithmetically correct in each category. The site evaluation macro performed properly in each case. Data were acquired from the various source worksheets, and the calculation correctly executed on sites with a variety of site conditions and operating histories. The module passed its tests and is suitable for use. The sites with long operating histories were found to take several minutes or more to process.

- **Split File**—Allows users to subdivide large simulations into smaller ones and run them on separate machines for speed or convenience. Three design cases and seven fault cases were tested. The criterion for success in the design cases was the identification of sites and operable units for removal on user direction. The remove selected operable units and split macro functions performed properly. All user-identified sites or operable units for removal were properly identified, and no additional site or operable units were identified. The criterion for success in the fault cases was user notification. A dialog appeared, notifying the user in the specified fault cases. The module passed its tests and is suitable for use. Furthermore, a file containing a single operable unit/group was shown to not split.

- **Test Distribution**—Allows users to check a file for input data conformance prior to performing a simulation. One design case and five fault cases were tested. The criterion for success in the design case was user notification of a successful initialization. This notification was achieved. The criterion for success in the fault cases was notification and identification of the presence of faults. This notification was achieved. The test distribution function was found to work as specified but is limited in certain respects. Two warning dialogs notify the user regarding the presence of faults that require correction during this process. The user can ignore these warnings and attempt to simulate data files with faults, but the program will fail to successfully complete because of the presence of uncorrected faults. Additionally, we found that when numerous faults are present in a tested input file, it is necessary to sequentially correct the faults and retest the file. However, attempting to simulate a file with numerous faults and inspecting the error log is allowed and useful as a troubleshooting method in these cases.

- **Graphical User Interface Functionality**—The C# and Microsoft Excel graphical user interface (GUI) forms control the administration of the software and its various features. Five design cases and seven fault cases were tested. The criterion for success in the design cases was activation of selected user interface features such as input cells, buttons and checkboxes. The GUI functionality works properly, passed its tests, and is suitable for use. The criterion for success in the specified fault cases was notification of the user to complete or correct an action. In the specified fault cases, a dialog appeared, notifying the user, and thus passed its test. There can be some latency between when the user action via mouse or keyboard is initiated and when the program responds, so patience on the part of the user is required. The fault cases represent a combination of plausible user oversights and out-of-specification inputs.

- **Inventory Calculation**—Is the principal function of Hanford SIM Rev. 2; it uses the input data to perform the stochastic calculation in OCB and report the results. Eight design cases and zero fault cases were tested. The design cases tested various options available to the user, such as using Latin hypercube sampling, excluding individual sites, analytes, and operable units, or processing subdivided input files. The criterion for success in the design cases was completion of the simulation calculation as defined for each design case. The inventory calculation and simulation administration functions worked properly for complete simulations and for selected exclusions for user-identified sites or analytes. Both simple random sampling and Latin hypercube sampling features activated and implemented properly. Processing split simulations provided usable results. The module successfully passed all its tests and is suitable for use. No fault cases were processed as part of this test because the error identification functions inform the user of the presence of any faults that require correction before simulation. This module is the core of the Hanford SIM Rev. 2 functionality.

- **Convergence—**Performs comprehensive trial-to-trial evaluation of Hanford SIM Rev. 2 results to determine model stability and repeatability. Three design cases and no fault cases were tested. The criteria for success for this module were to ensure that it computed the relative percent differences (RPDs) properly, saved the results, tallied the RPD results correctly, categorized them at the specified tolerances, and reported them to the worksheet interface. Processing identical file outputs demonstrated mathematical integrity of macro command (reported all zeros, as anticipated). Other cases demonstrated proper macro functions. Results of the tests were saved, and the RPDs were calculated correctly, categorized, and reported as designed. The module passed all its tests and is suitable for use. Convergence was found to take several minutes to perform, depending on the file size and number of files processed.

- **Merge—**Allows users to reconstitute split simulation results into a single comprehensive file. One design case and zero fault cases were tested. The criterion for success was that the macro command create a Top10 list with the correct order out of a subdivided file and consolidate the subfile results back into a single set of results. The merge macro performed properly, passing its test. Inspection of interim Top10 List files confirmed site–analyte order was maintained correctly when reconstituting, and all output files were populated with the user-defined results. The design case performed its tasks correctly and passed its test. The module is suitable for use.

- **Make SAC Output—**Extracts the Hanford SIM output results and reconfigures them into a format that the System Assessment Capability Inventory Module can use directly. One design case and no fault cases were tested. The success criteria are for it to create a directory for the output files, create workbook/worksheet files, and populate these files with the site–year–analyte results for each operable unit. On activation of the macro, the values were copied and pasted correctly to their designated worksheets and workbooks. The design case performed its tasks correctly and passed its test. The module is suitable for use.

- **BlackBoxtest1**—Performs a comprehensive check of the selected Top10 site–analyte list for calculation verification using the Site Evaluation macro. The results also serve as an additional check on the model convergence behavior. The results from this macro are reported in a worksheet in the identified source data workbook. One design case and no fault cases were tested. The criteria for success for this macro were to run to completion, executing the site–analyte combinations reported in the designated Top10List through the Site Evaluation macro, recording the results of that Crystal Ball simulation, copying the source OCB simulation results in the target worksheet, and counting the relative standard deviation (RSD) and median error results outside the design criteria. The RSD error design criterion is an RSD ratio of less than 0.95 between the two simulations, and the median error design criterion is an absolute median RPD of less than 0.05 between the two simulations. The macro ran to completion, performing all design operations successfully, and reported the results of its testing to the correct sheet. The values were copied, pasted, and calculated correctly. The start and end times were reported. The number and types of errors (median and RSD) were reported and quantified correctly. The design case performed its tasks successfully, and the macro is suitable for use. We found that this macro takes several hours to complete.

- **cCDIcompare**—Performs selected inventory comparisons between model results and reference values (Diediker 1999) for specific sites. This macro command performs a series of comparisons at various stages of resolution and reports the results in the *source data* workbook. One design case and no fault cases were tested. The success criteria for this macro are for it to run to completion, perform the identified site–analyte comparisons, and report the results. The macro ran to completion, performed its design functions, and reported the results of its testing to the correct sheet. The values were copied, pasted, and calculated correctly. The quantifications/ comparisons were performed correctly. The summary result reporting was done correctly. The design case performed its tasks successfully, and the macro is suitable for use.

- **0,1,2 Compare**—Performs selected inventory comparisons for sites and groups of sites against the sums of reference inventory values, correcting for the presence of "less-than" values after executing the **cCDI compare** command**.** One design case and no fault cases were tested. The success criteria for this macro are for it to run to completion, perform the identified site–analyte comparisons, corrections, and summations, and report the results. The macro ran to completion, performed its design functions, and reported the results of its testing to the correct sheet. The values were copied, pasted, and calculated correctly. The quantifications/comparisons were performed correctly. The summary result reporting was done correctly. The design case performed its tasks successfully, and the macro is suitable for use.

- **Database Transfer—**Maintains input data and output results under configuration control using SQL Server 2000. The data transformation services (DTS) function provided by SQL Server 2000 worked properly. These features represent functionality beyond that in Hanford SIM Rev. 1. Because these scripts are highly specific, they have only one design case and no fault cases. They will not work properly without a successfully completed simulation file or series of output files. The success criterion for this module is for it to run to completion, transferring data from Microsoft Excel files to SQL Server 2000 files. The data were transferred successfully. The design case performed its tasks successfully, and the module is suitable for use. Familiarity with databases and SQL Server 2000 is required to perform these functions.

- **Installation—**Transfers data and application files from storage media to properly configured target computer and hinders transfer to a nonconforming computer. Two design cases and one fault case were tested. The success criterion for this module was for the application and data files to transfer and operate on a properly configured machine. Installation and operation of OCB and GoldSim files were successful. Installation terminated as specified on a nonconforming machine (i.e., a target machine without all COTS software installed prior to installation of Hanford SIM Rev. 2). The design cases performed their tasks successfully, and the module is suitable for use.

These tests were conducted using files with a variety of inputs. Both design (positive) and fault (negative) test cases were conducted to establish the performance of the software portfolio. Inspections of the input files and simulation outputs were conducted during testing. The outcome of the testing is described more fully and documented in the Hanford SIM Rev. 2 module test report[5] and in the test logs that accompany the project file.

5.2 Verification and Validation Testing

A series of tests examined the performance of OCB and the associated distributions being used as part of Hanford SIM Rev. 2. Verification of software can be achieved by 1) comparison of code output and intermediate output against hand calculations, 2) intercomparison of codes in which the comparative code has been fully documented and verified, and 3) comparison of a numerical code against published analytical solutions. Verification in Hanford SIM Rev. 2 is achieved via intercomparison using GoldSim.

Validation of a computer code means that it provides a satisfactory representation of the actual process or system it is intended to reproduce. In general, the validation effort confirms that the conceptual model, the encoded algorithms, and its supporting data work together to produce an expected result, including the expected sensitivities of the result attributable to known variations in the model parameters. The comparisons with reference data during end-to-end portfolio testing, blackbox testing, convergence, and cross-validation convergence testing were used to validate the performance of Hanford SIM Rev. 2. This series of tests is described in further detail in the Hanford SIM Rev. 2 validation and verification test plan,[6] which is part of the project file and included in the project records.

5.2.1 Independent Software Calculation Verification Testing

A series of example calculations in a test file was defined to examine the performance of OCB and the associated distributions being used as part of Hanford SIM Rev. 2. The results of the OCB test cases were compared to a variety outputs using results created with GoldSim with mathematically congruent inputs as independent benchmark comparisons. These forecast outputs were compared to establish that the results generated by each software package are statistically indistinguishable (e.g., OCB results were compared with GoldSim results). This test provides assurance that the individual uncertainty components used in Hanford SIM Rev. 2 were being correctly quantified and propagated. One design case was run using eight sites and seven waste streams.

---

[5]Simpson BC, MJ Anderson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Module Test Report*. VIV07-33573-STR-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.
[6]Simpson BC, MJ Anderson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Validation and Verification Test Plan*. VIV07-33573-VVTP-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.

This verification test ensured that the OCB algorithms provide legitimate distribution definitions when compared to an independent third-party software package. All distributions available for use in Hanford SIM Rev. 2 were evaluated. Normal, lognormal, triangular, exponential, beta, Weibull, and gamma distributions of varying parameterizations, but mathematically congruent between the two software packages (GoldSim and OCB), were defined and forecasts created. These files are part of the installation package.

The success criteria for this test are for the locally run OCB simulation and GoldSim player files to provide consistent results and for the local OCB run to provide results within the convergence criteria when compared to the provided test data file. The OCB test file results and GoldSim player results can be inspected by the user as well as operated to confirm their performance. Visual inspection of the outcomes provided sufficient evidence that the OCB test file runs properly, the software applications are working, and that the GoldSim player provides reliable, statistically indistinguishable results.

5.2.2 Comprehensive Testing

This test represents a comprehensive in-use test of all the software components described in system testing (user interface, calculation, and diagnostics) under typical user conditions. The applications were tested initially using an appropriate test data file of known behavior to evaluate their in-use performance and their responses, before production files were simulated. Additionally, final end-to-end verification and validation testing using approved inputs was done. Two design cases were run. These results were evaluated using converged results for both design cases. The end-to-end test with proxy data ensured that all the modules worked together properly. Furthermore, most of the outcomes generated with production data did not present discrepancies.

In cases in which the end-to-end test with production data found that the new results (i.e., Hanford SIM Rev. 2) had some minor discrepancies when compared to the previous results (Hanford SIM Rev. 1), these discrepancies were anticipated because some of the comparisons performed initially were very close to the cCDI/0,1,2 comparison boundary ("on the fringe") for each of the analytes compared (reference values for strontium-90, cesium-137, uranium-238, plutonium-239) and had happened before. Thus, these same fringe effects from performing the cCDI/0,1,2 comparisons on the new Hanford SIM Rev. 2 results were observed. Seven site–analyte combinations that were previously inside the 0.5th percentile boundary fell out, and five site–analyte combinations that were previously outside the boundary came in, for a net loss of two positive comparisons. Table 5.1 documents the changes.

The fringe behavior observed was present in Hanford SIM Rev. 1 and is currently in Hanford SIM Rev. 2. No systematic change in behavior was observed, and no impact from changing distributions could be attributed to the observed behavior. The remaining site–analyte combination results remained the same. Thus, with a net change of less than one-half of one percent, the fringe behavior does not have a practical impact on the results.

Table 5.1. Comparison of Hanford SIM Rev. 1 and Hanford SIM Rev. 2 In- and Out-of-Range Results for Selected Site–Analyte Combinations

Site	Reference	Results Hanford SIM Rev. 1		Result (%)	Reference	Results Hanford SIM Rev. 2		Result (%)	Analyte; Change
	Values Inside Range	Reference Above Range	Reference Below Range		Values Inside Range	Reference Above Range	Reference Below Range		
216-A-7	3	0	1	75	4	0	0	100	Cs-137; out to in
216-B-23	3	0	1	75	2	0	2	50	Cs-137; in to out
216-B-36	3	0	1	75	4	0	0	100	Pu-239; out to in
216-B-52	2	0	2	50	1	0	3	25	Pu-239; in to out
216-B-9	2	1	1	50	1	1	2	25	Sr-90; in to out
216-S-7	2	0	2	50	3	0	1	75	Cs-137; out to in
216-T-18	3	1	0	75	2	1	1	50	Sr-90; in to out
216-T-24	4	0	0	100	3	0	1	75	Sr-90; in to out
216-T-36	3	0	1	75	4	0	0	100	Pu-239; out to in
216-Z-12	2	0	2	50	1	0	3	25	U-238; in to out
216-Z-1A	2	0	2	50	1	0	3	25	U-238; in to out
216-Z-4	3	0	1	75	4	0	0	100	Cs-137; out to in

5.2.3 Black Box Testing

In black box testing, an input test file was generated based upon selected cases involving specific Hanford SIM Rev. 2 analyte and site combinations. A Crystal Ball simulation then was run based on these parameterizations and run preferences identical to the OCB simulation. Statistical output from this Crystal Ball test was then compared against the OCB output for the same test analyte by site, determining internal consistency and calculation integrity.

In this case, a production-based Top10List was the input test file. One design case was run in which the criterion for success was no greater than 5% of the results found to be outside the established thresholds. The test was passed successfully. In addition to meeting the acceptance criterion, the results of the latest production black box test were improved somewhat from the previous series of results. This improvement is attributed to refinements in the software performance made by the vendor. The results are compared in Table 5.2.

Table 5.2. Comparison of Black Box Testing Results for Hanford SIM Rev. 1 and Hanford SIM Rev. 2

	Hanford SIM Rev. 1	Hanford SIM Rev. 2
RSD Error Count	24	15
Median Error Count	47	0

The RSD error threshold is where the ratio of CB RSD/OCB RSD is less than 0.95. The median convergence error threshold is where 1 − [absolute value (OCB result-CB result)/OCB result] = greater than >0.05. The performance of the macro demonstrates internal consistency and calculational integrity for the software. The full results of the black box test can be found in the Hanford SIM Rev. 2 validation and verification test report,[7] which is part of the project file and is included in the project records.

5.2.4 Convergence

Convergence testing was conducted to ensure that the model results are repeatable within the convergence tolerance definition. This test does not measure error in the conventional sense (divergence from an accepted or known value); it determines only if the model parameters and model environment provide for stability in the results. In this case, because the inventory calculation is linear, returning statistically indistinguishable results for any selected model output from simulation to simulation under the same conditions is the criterion. If this feature is observed, the number of iterations performed is considered to be satisfactory, the computing environment is stable, and the results are considered reliable. Discrepancies arising from comparing very small numbers are discounted from the evaluation.

Five production Hanford SIM Rev. 2 runs were completed. Then convergence testing was done using a macro-driven tool to obtain the following output for evaluation: mean, standard deviation, median (50th percentile), 0.5th percentile, 5th percentile, 95th percentile, and 99.5th percentile. The run-to-run deviations were quantified and compared to the acceptance tolerances. These tolerances are described in Corbin et al. (2005) and excerpted here. The 0.5th and 99.5th percentiles are included as practical minimum and maximum values but are not used to establish convergence because of the known simulation instability at these extreme values.

Evaluation of the central tendencies, standard deviation, and tails is done for each site–analyte–year combination. The convergence definition in use for this task is that no more than 5% of the results can have greater than 5% deviation within the 5th to 95th percentile ranges, inclusive (e.g., a 90th percentile output range), allowing for the observed bias of small numbers in the evaluation formula at the lower tail of the distribution. The small number bias threshold was 1E-12 (1 pCi). This level was selected as a practical compromise that allowed for reducing difficulties involving detection limits of the instruments/ analytical methods and background interference. One design case was run using all new outputs.

Although there have been changes and updates to the Crystal Ball and OCB software, the mathematical behavior of the simulations appears to be practically the same as with the previous version, Hanford SIM Rev. 1. Trial-to-trial variability remains highly stable and thus meets the acceptance criterion. The contribution of small number errors at the 5th percentile is approximately the same as well. There are 325,298 site–analyte–year data points in this analysis. The error count excluding those arising from small numbers at the 5th percentile is 6,521, or 2.00% of the total, confirming that a

[7] Simpson BC, MJ Anderson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Validation and Verification Test Report*. VIV07-33573-VVTR-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.

25,000-trial LHS simulation is satisfactory. The summary convergence test results are presented in the validation and verification test report,[8] which is part of the project file and is included in the project records.

5.2.5 Conformance with User Requirements – Cross-Validation Convergence

Cross-validation convergence testing was conducted to ensure that the model results from the current version of the Hanford SIM Rev. 2 inputs and software are statistically indistinguishable from the Corbin et al. (2005) results. Five production Hanford SIM Rev. 2 runs using different proportions of source data (i.e., previous and current simulation results) were tested with the same algorithm and tolerances described in Section 5.2.4. Two design cases were run (4 old, 1 new; 3 old, 2 new). Notable discrepancies arising from input changes dictated by implementing new software were identified.

These tests were conducted using files with test file and final design inputs. Only design (positive) test cases were run to establish the validation and verification performance of the software portfolio. Inspections of the input files and simulation outputs were conducted during testing.

Because of the changes and updates to the Crystal Ball and OCB software, especially with regard to disallowing previous distribution definitions that were mathematically suspect, the cross-validation convergence performance criterion was reconsidered.

The internal consistency of the latest results is as good as—or in some cases better than—the previous Hanford SIM Rev. 1 outcomes (observed in the convergence and black box tests). The results from the later series of simulations can be distinguished from the previous results. Tables 5.3 and 5.4 summarize the waste streams and analytes affected by this change. These changes affect analytes with behavior observed to be highly erratic (plutonium-241, plutonium-242, americium, and curium) in Hanford SIM Rev. 1 and are present in very small quantities. These intrinsic characteristics of the data (i.e., high variability and small quantities) did not change with the updates to the software; in fact, the software in Hanford SIM Rev. 2 provided more mathematically reliable results. The inventory changes to these analytes are considered to not have practical impacts on the results, and the deviations from the previous results are not considered a deficiency. Additional information, including a mapping of analytes to waste streams, can be found in the validation and verification test report,[8] which is part of the project file and is included in project records.

---

[8]Simpson BC, MJ Anderson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Validation and Verification Test Report*. VIV07-33573-VVTR-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.

Table 5.3. Waste Streams with Substituted Lognormal Distributions

Impacted Waste Streams	Number of Analytes Affected	Impacted Waste Streams	Number of Analytes Affected	Impacted Waste Streams	Number of Analytes Affected	Impacted Waste Streams	Number of Analytes Affected
1C1 (BT1)	5	5-6 (BT1)	5	PASF (P2)	7	Recuplex (Z2) aqu	15
209-E Reflector Wtr (BT1)	1	A1-SltCk(Z2)	15	PASF (P2')	11	Recuplex (Z2) org.	14
209-E Reflector Wtr (P2)	3	A2-SltSlr(Z2)	15	PASF (P3)	10	REDOX (P2') Cool Wtr	11
209-E Reflector Wtr (P2')	5	BiPO4 (BT1) Cool Wtr-Stm Cond	5	PFeCN1 (BT1)	5	REDOX (P2') Stack Drain	11
209-E Reflector Wtr (P3)	2	BiPO4 (BT1) Stack Drain	5	PUREX (P2) Chem Sewer	7	REDOX (P2') Tank Farm Cond	10
222-S Lab Wst (P2)	7	Conc Misc UNH Streams (P2')	1	PUREX (P2') Chem Sewer	10	REDOX (P3) Stack Drain	11
222-S Lab Wst (P2')	11	CWP1 (CWP1)	2	PUREX (P2) Cool Wtr-Stm Cond	7	REDOX (R1) Cool Wtr	3
222-S Lab Wst (P3)	8	CWP2 (CWP2)	6	PUREX (P2') Cool Wtr-Stm Cond	10	REDOX (R1) Org	3
222-S Lab Wst Wtr (P2)	7	CWR1 (CWR1)	5	PUREX (P2) Org Wst	7	REDOX (R2) Cool Wtr	7
222-S Lab Wst Wtr (P2')	11	CWR2 (CWR2)	26	PUREX (P2') Org Wst	11	REDOX (R2) Org	7
222-S Lab Wst Wtr (P3)	8	CWZr1 (CWZr1)	5	PUREX (P2) Org Wst aqu_OWW1	7	REDOX (R2) Tank Farm Cond	2
222-S Lab Wst Wtr (R1)	3	CWZr2 (CWZr2)	11	PUREX (P2) Org Wst aqu_OWW2	7	REDOX D-1 (R1)	3
222-S Lab Wst Wtr (R2)	7	Decon Stack Drain (R2)	2	PUREX (P2') Org Wst aqu_OWW3	11	REDOX D-1 (R2)	7
224 (BT1)	5	Decon Wst (BT1)	5	PUREX (P2) Stack Drain	7	REDOX D-2 (R1)	3
231-Z Metal Lab (Z2)	15	Decon Wst (P2)	7	PUREX (P2') Stack Drain	11	REDOX D-2 (R2)	7
232-Z Inc (Z1)	1	Decon Wst (P2')	11	PUREX (P2) Tank Farm Cond	7	REDOX Stack Drain (R1)	3
232-Z Inc (Z2)	15	Decon Wst (P3)	11	PUREX (P2') Tank Farm Cond	10	REDOX Stack Drain (R2)	7
234-5Z (BT1) D-6	5	Decon Wst (R2)	2	PUREX (P3) Chem Sewer	7	RG Process (BT1)	5
234-5Z (Z1) D-6	15	Dil Misc UNH Streams (P2)	3	PUREX (P3) Cool Wtr-Stm Cond	7	RSLT (R2)	7
234-5Z (Z2) D-6	15	Dil Misc UNH Streams (P2')	5	PUREX (P3) Process Cond	9	S1-SltCk(P2')	11
242-A Cond (P2')	10	Dil Misc UNH Streams (P3)	2	PUREX (P3) Stack Drain	11	S2-SltSlr(P2')	11
242-A Cond (P3)	7	Dil Misc UNH Streams (R2)	2	PUREX (P3) Tank Farm Cond	7	Spent Nitric Acid (P2)	7
222-S Lab Wst Wtr (P2)	7	Laundry Wst Wtr (P2')	10	PUREX PL2 (P3)	11	Spent Nitric Acid (P3)	11
222-S Lab Wst Wtr (P2')	11	Laundry Wst Wtr (P3)	7	R1 (R1)	3	Spent Nitric Acid (R1)	3
222-S Lab Wst Wtr (P3)	8	Laundry Wst Wtr (R2)	6	R2 (R2)	7	Spent Nitric Acid (R2)	7
222-S Lab Wst Wtr (R1)	3	MW1 (BT1)	5	Recuplex (BT1) aqu	5	Sr-Cs Rec (P2') Chem Sewer	6
222-S Lab Wst Wtr (R2)	7	N Decon Wst (P2)	7	Recuplex (BT1) org.	5	Sr-Cs Rec (P2) Cool Wtr	7
224 (BT1)	5	P3AZ1(P3)	11	Recuplex (Z1) aqu	16	Sr-Cs Rec (P2') Cool Wtr	11
231-Z Metal Lab (Z2)	15	P3AZ2(P3)	11	Recuplex (Z1) org.	9	Sr-Cs Rec (P2) Stack Drain	3
232-Z Inc (Z1)	1	Decon Wst (P2')	11	PUREX (P2) Tank Farm Cond	7	REDOX Stack Drain (R1)	3

Table 5.3. (contd)

Impacted Waste Streams	Number of Analytes Affected	Impacted Waste Streams	Number of Analytes Affected	Impacted Waste Streams	Number of Analytes Affected	Impacted Waste Streams	Number of Analytes Affected
Sr-Cs Rec (P2') Stack Drain	5	Sr-Cs Rec (R2) Chem Sewer	4	T2-SltCk(P2')	11	Z Complex Lab Wst (Z2)	15
Sr-Cs Rec (P3) Chem Sewer	5	Sr-Cs Rec Org Wst (P2)_B	7	Z Complex Cool Wtr-Cond (BT1)	5	Z Complex Stack Drain (BT1)	5
Sr-Cs Rec (P3) Cool Wtr	11	Sr-Cs Rec Org Wst (P2')_CSR	11	Z Complex Cool Wtr-Cond (Z1)	9	Z Complex Stack Drain (Z1)	9
Sr-Cs Rec (P3) Stack Drain	2	Sr-Cs Rec Org Wst aqu (P2)_BL	7	Z Complex Cool Wtr-Cond (Z2)	14	Z Complex Stack Drain (Z2)	14
Sr-Cs Rec (R1) Chem Sewer	3	Sr-Cs Rec Org Wst aqu(P2')_AR	11	Z Complex Lab Wst (BT1)	5	Z Complex Stack Drain_NCT (Z2)	14
Sr-Cs Rec (R1) Cool Wtr	3	Sr-Cs Rec Wst (P2)_SRR	7	Z Complex Lab Wst (Z1)	15	Z(Z2)	5

Table 5.4. Analytes with Substituted Lognormal Distributions

Impacte d Analytes	Number of Waste Streams Affected	Impacte d Analytes	Number of Waste Streams Affected	Impacte d Analytes	Number of Waste Streams Affected
Am-241	57	Eu-155	4	Sm-151	1
Am-243	121	I-129	1	Sn-126	1
Ba-137m	1	Nb-93m	1	Sr-90	1
Cm-242	110	Ni-59	2	Tc-99	1
Cm-243	121	Ni-63	1	U-232	42
Cm-244	121	Pu-238	77	U-234	21
Co-60	3	Pu-241	83	U-235	21
Cs-134	14	Pu-242	125	U-238	21
Cs-137	1	Ru-106	52	U-Total	21
Eu-152	14	Sb-125	13	Y-90	1
Eu-154	4	Se-79	1	Zr-93	1

Because the input change and resulting calculation was to rectify a latent error that existed, this condition was viewed as an improvement/correction and not an error. The arithmetic means where changes to inventory were made as a result of changing the input definitions were well within the performance criteria for convergence, but the variability behavior observed in the tails was slightly greater overall (between 7% and 10% deviation) when quantified between the two sets of results. Because the previously calculated results were potentially flawed, and the magnitude of the inventories where changes occurred are very small (typically 1E-06 Ci or less), there is no practical change or impact on the results.

The outcome of the testing is described more fully and documented in the Hanford SIM Rev. 2 validation and verification test report,[9] which is part of the project file and included in the project records.

---

[9]Simpson BC, MJ Anderson, and RA Corbin. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2, Validation and Verification Test Report*. VIV07-33573-VVTR-001, Rev. 0, Vivid Learning Systems, Pasco, Washington.

6.0 Hanford SIM Rev. 2 Outputs

The individual category (Operable Unit) workbooks specified by the user in the *Legend* worksheet are created as defined in **OCBHanford**, with their respective sites, analytes, years, and categories. There can be as many output workbooks as there are sites being evaluated. However, that structure is cumbersome because the number of workbooks created can be unwieldy. In Hanford SIM Rev. 2, logical groupings exist and are used to reduce the overall number of output files produced.

Additionally, as a function of producing the individual site–year–analyte results, summary analyte results were developed for the sites, analytes, and groupings over their operating lives, resulting in a distribution for each for all the contributing years. Summary outputs for selected sites and each operable unit group are exported into their respective worksheets of the production *source data* workbook (*SumFrcTotal, SumFrcSolid, SumFrcLiquid)* and the *Top10List* worksheet, consolidating the highest-magnitude results for a site–analyte combination.

Additionally, there is a post-process macro command used to generate SAC input files, **Create SAC Output,** that creates a file that can be read and used directly by the SAC model. Other outputs created by the Hanford SIM Rev. 2 post-process macro commands involve producing the results from the QA testing process. The execution of these macro commands will create or refresh results from a simulation (or series of simulations) in the production workbook (such as activating the **VolumeBalance** macro or the **cCDI Analyte Comparision**), or in a separate file as a result of executing the **Convergence** test macro.

The simulation inputs and outputs for a selected series of results can be preserved using Microsoft SQL Server 2000 in tabular database files to ensure traceability of the data. This feature is not necessary to perform or analyze a simulation but is recommended to maintain configuration control. An additional software requirement to use this capability is the acquisition and installation of SQL 2000 software.

7.0 References

Abrams RB. 1956. *Metal Recovery Waste Scavenging Program*. HW-43066, General Electric Company, Richland, Washington.

Agnew SF. 1997. *Hanford Tank Chemical and Radionuclide Inventories: HDW Model Rev. 4*. LA-UR-96-3860, Los Alamos National Laboratory, Los Alamos, New Mexico.

Corbin RA, BC Simpson, MJ Anderson, WF Danielson III, JG Field, TE Jones, and CT Kincaid. 2005. *Hanford Soil Inventory Rev. 1*. RPP-26744, Rev. 0, CH2M HILL Hanford Group, Inc., Richland, Washington.

Decisioneering. 2002. Crystal Ball Software. Decisioneering, Inc., Denver, Colorado. Available at http://www.decisioneering.com (October 2007).

Decisioneering. 2006a. *Crystal Ball 7.2.2 User Manual*. Decisioneering, Inc., Denver, Colorado.

Decisioneering. 2006b. Open Crystal Ball. Decisioneering, Inc., Denver, Colorado. Available at http://www.decisioneering.com/ocb/index.html (October 2007).

Diediker LP. 1999. *Radionuclide Inventories of Liquid Waste Disposal Sites on the Hanford Site*. HNF-1744, Fluor Daniel Hanford, Inc., Richland, Washington.

DOE Order 414.1C. 2005. *Quality Assurance*. U.S. Department of Energy, Washington, D.C.

DOE-RL. 1999. *Groundwater/Vadose Zone Integration Project Science and Technology Summary Description*. DOE/RL-98-48, Vol. III, Rev. 0, U.S. Department of Energy, Richland Operations Office, Richland, Washington.

Hammersley JM and DC Handscomb. 1964. *Monte Carlo Methods*. Methuen & Co. Ltd., London, England.

Hanthorn HE. 1957. *Hanford History, Technology, Expansion, and Present Efforts*. HW-51188, General Electric Company, Richland, Washington.

Healy JW. 1953. *Release of Radioactive Wastes to Ground*. HW-28121, General Electric Company, Richland, Washington.

Higley BA, DE Place, RA Corbin, and BC Simpson. 2004. *Hanford Defined Waste Model, Rev. 5*. RPP-19822, Rev. 0, CH2M HILL Hanford Group, Inc., Richland, Washington.

Iman RL and WJ Conover. 1982. "A Distribution-Free Approach to Inducing Rank Correlations Among Input Variables." *Communications in Statistics* B11(3):311-334.

Paas HJ and KR Heid. 1955. *Radioactive Contamination in Liquid Wastes Discharged to Ground at Separation Facilities though June 1955*. HW-38562, General Electric Company, Richland, Washington.

Ruppert HG and KR Heid. 1954. *Summary of Liquid Radioactive Wastes Discharged to the Ground – 200 Areas July 1952 to June 1954*. HW-33591, General Electric Company, Richland, Washington.

Simpson BC, RA Corbin, and SF Agnew. 2001. *Hanford Soil Inventory Model*. BHI-01496, Rev. 0, Bechtel Hanford Inc., Richland, Washington.

Simpson BC, RA Corbin, MJ Anderson, and CT Kincaid. 2006a. *Hanford Soil Inventory Model (SIM) Rev. 1 User's Guide*. PNNL-16099, Pacific Northwest National Laboratory, Richland Washington.

Simpson BC, RA Corbin, MJ Anderson, and CT Kincaid. 2006b. *Hanford Soil Inventory Model (SIM) Rev. 1 Software Documentation – Requirements, Design, and Limitations*. PNNL-16098, Pacific Northwest National Laboratory, Richland, Washington.

Simpson BC, RA Corbin, MJ Anderson, and CT Kincaid. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2 User's Guide*. PNNL-16925, Pacific Northwest National Laboratory, Richland Washington.

Watrous RA, DW Wootan, and SF Finfrock. 2002. *Activity of Fuel Batches Processed Through Hanford Separations Plants, 1944 Through 1989*. RPP-13489, Rev. 0, CH2M HILL Hanford Group, Inc., Richland, Washington.

Appendix A

Software Environment Maintenance

Appendix A

Software Environment Maintenance

During the life of the project, operating system, application, and security patches were distributed for the various software packages used. This appendix provides a listing of the patches incorporated.

Microsoft Windows XP Professional, Version 2002, Service Pack 2

- Security Update for Windows XP (KB913433)
- Windows Media Player 10
- Hotfix for Windows XP (KB896344)
- Update for Windows XP (KB894391) installed on 8/29/2006
- Security Update for Windows XP (KB896358) installed on 8/29/2006
- Security Update for Windows XP (KB896423) installed on 8/29/2006
- Security Update for Windows XP (KB896424) installed on 8/29/2006
- Security Update for Windows XP (KB896428) installed on 8/29/2006
- Update for Windows XP (KB898461) installed on 8/29/2006
- Security Update for Windows XP (KB899587) installed on 8/29/2006
- Security Update for Windows XP (KB899589) installed on 8/29/2006
- Security Update for Windows XP (KB899591) installed on 8/29/2006
- Windows XP HotFix – (KB587333) installed on 8/29/2006
- Security Update for Windows XP (KB890046) installed on 8/29/2006
- Security Update for Windows XP (KB901214) installed on 8/29/2006
- Security Update for Windows XP (KB902400) installed on 8/29/2006
- Security Update for Windows XP (KB904706) installed on 8/29/2006
- Windows XP HotFix (KB888583) installed on 8/29/2006
- Security Update for Windows XP (KB905414) installed on 8/29/2006
- Security Update for Windows XP (KB905749) installed on 8/29/2006
- Security Update for Windows XP (KB908519) installed on 8/29/2006
- Security Update for Windows XP (KB893756) installed on 8/29/2006
- Update for Windows XP (KB910437) installed on 8/29/2006
- Update for Windows XP (KB911280) installed on 8/29/2006
- Security Update for Windows XP (KB911562) installed on 8/29/2006
- Security Update for Windows Media Player (KB911564) installed on 8/29/2006
- Security Update for Windows XP (KB911567) installed on 8/29/2006
- Security Update for Windows XP (KB911927) installed on 8/29/2006
- Security Update for Windows XP (KB912919) installed on 8/29/2006
- Windows XP HotFix – (KB889178) installed on 8/29/2006
- Security Update for Windows XP (KB913580) installed on 8/29/2006
- Security Update for Windows XP (KB914388) installed on 8/29/2006
- Security Update for Windows XP (KB914389) installed on 8/29/2006
- Windows XP HotFix – (KB888583) installed on 8/29/2006

- Update for Windows XP (KB916595) installed on 8/29/2006
- Security Update for Windows XP (KB917159) installed on 8/29/2006
- Security Update for Windows XP (KB917422) installed on 8/29/2006
- Windows XP HotFix – (KB888618) installed on 8/29/2006
- Security Update for Windows Media Player 9 (KB917734) installed on 8/29/2006
- Security Update for Windows XP (KB917953) installed on 8/29/2006
- Windows XP HotFix – (KB888747) installed on 8/29/2006
- Security Update for Windows XP (KB918439) installed on 8/29/2006
- Windows XP HotFix – (KB888830) installed on 6/29/2006
- Update for Windows XP (KB908531) installed on 8/29/2006
- Security Update for Windows XP (KB920214) installed on 8/29/2006
- Windows XP Hotfix – (KB889085) installed on 8/29/2006
- Security Update for Windows XP (KB920670) installed on 8/29/2006
- Security Update for Windows XP (KB920683) installed on 8/29/2006
- Windows Genuine Advantage Notifications (KB905474) installed on 8/29/2006
- Windows Genuine Advantage Validation Tool (KB892130) installed on 8/29/2006
- Security Update for Windows XP (KB921398) installed on 8/29/2006
- Security Update for Windows XP (KB921883) installed on 8/29/2006
- Security Update for Windows XP (KB922616) installed on 8/29/2006
- Security Update for Windows XP (KB901017) installed on 8/29/2006
- Security Update for Windows Media Player 10 (KB917734) installed on 9/22/2006
- Security Update for Windows XP (KB923414) installed on 10/19/2006
- Security Update for Windows XP (KB919007) installed on 10/19/2006
- Security Update for Windows XP (KB920685) installed on 10/19/2006
- Security Update for Windows XP (KB923191) installed on 10/19/2006
- Security Update for Windows XP (KB922819) installed on 10/19/2006
- Security Update for Windows XP (KB924191) installed on 10/19/2006
- Update for Windows XP (KB931836) installed on 3/12/2007
- Security Update for Windows XP (KB924667) installed on 5/26/2007
- Security Update for Windows Media Player 6.4 (KB925398) installed on 5/26/2007
- Update for Windows XP (KB900485) installed on 5/26/2007
- Security Update for Windows XP (KB923980) installed on 5/26/2007
- Security Update for Windows XP (KB925902) installed on 5/26/2007
- Security Update for Windows XP (KB926255) installed on 5/26/2007
- Security Update for Windows XP (KB923694) installed on 5/26/2007
- Security Update for Windows XP (KB927779) installed on 5/26/2007
- Security Update for Windows XP (KB927802) installed on 5/26/2007
- Update for Windows XP (KB927891) installed on 5/26/2007
- Security Update for Windows XP (KB928255) installed on 5/26/2007
- Security Update for Windows XP (KB928843) installed on 5/26/2007
- Update for Windows XP (KB922582) installed on 5/26/2007
- Security Update for Windows XP (KB930 178) installed on 5/26/2007
- Update for Windows XP (KB930916) installed on 5/26/2007

- Security Update for Windows XP (KB931261) installed on 5/26/2007
- Security Update for Windows XP (KB931784) installed on 5/26/2007
- Security Update for Windows XP (KB924270) installed on 5/26/2007
- Security Update for Windows XP (KB918118) installed on 5/26/2007
- Security Update for Windows XP (KB923689) installed on 5/26/2007
- Security Update for Windows XP (KB920213) installed on 5/26/2007
- Security Update for Windows XP (KB926436) installed on 5/26/2007
- Update for Windows XP (KB920872) installed on 5/29/2007
- Update for Windows XP (KB925720) installed on 5/29/2007
- Windows Imaging Component installed on 5/29/2007
- Update for Windows XP (KB920342) installed on 5/29/2007
- Update for Windows XP (KB904942) installed on 5/29/2007
- Hotfix for Windows XP (KB914440) installed on 5/29/2007
- Update for Windows XP (KB920342) installed on 5/29/2007
- Update for Windows XP (KB904942) installed on 5/29/2007
- Update for Windows XP (KB936357) installed on 7/11/2007
- Update for Windows XP (KB925876) installed on 5/29/2007
- Hotfix for Windows XP (KB914440) installed on 5/29/2007
- Security Update for Windows XP (KB932168) installed on 5/29/2007
- Security Update for Windows XP (KB935840) installed on 6/13/2007
- Security Update for Windows XP (KB929123) installed on 6/13/2007
- Security Update for Windows XP (KB935839) installed on 6/13/2007

Microsoft.NET 1.1 provides the application environment. These are the updates used as part of the development and production environment.

- Microsoft .NET Framework 1.1 Hotfix (KB928366)
- Microsoft .NET Framework 2.0
- Security Update for Microsoft .NET Framework 2.0 (KB928365)

Microsoft Excel 2003; user interface for data input/output and analysis as part of Microsoft Office Professional 2003 with updates used as part of the development and production environment.

- Update for Office 2003: LCCWIZ
- Office 2003 Service Pack 2
- Security Update for Office 2003: RICHED20
- Security Update for Office 2003: MSXML5
- Update for Office 2003: OTKLOADR
- Update for Office 2003: STLIST
- Security Update for Office 2003: GPFILTIff
- Security Update for Excel 2003: EXCEL
- Security Update for Office 2003: MSO

Appendix B

Impacts of Input and Boundary Condition Development

Appendix B

Impacts of Input and Boundary Condition Development

In evaluating the data (both inputs and results), the magnitude of uncertainty associated with estimates is significant, spanning in some cases an order of magnitude. This condition does not necessarily represent a deficiency in the data; all that can be inferred is that the system has a substantial amount of intrinsic uncertainty and that any decisions made must consider this feature. Efforts to minimize uncertainty by removing selected data points without strong technical rationale can introduce bias into the model. In fact, extreme features observed in the data often were deliberate and cannot be discounted. Uncertainties of this magnitude have been observed in sample data obtained from other sources at Hanford; thus, the behavior is not unexpected.

Because of the modeling assumptions, the resulting output distributions were relatively simple—skewed, non-negative, and monomodal, each with a well-defined central maximum. More complicated behavior is possible for these model inputs (and outputs); however, the assumptions for the contaminants used in Hanford SIM Rev. 2 are considered a practical compromise in appropriately describing uncertainties associated with Hanford Site waste management and disposal activities at a reasonable level of resolution.

Model simplification and constraints associated with software coding to make the inputs tractable removed many of the irregularities observed in the source data to condition its use as input to Hanford SIM Rev. 2. This smoothing effect is a modest source of bias when evaluating the Hanford SIM Rev. 2 results versus surveillance data because the source data are inherently spiky. The output concentration values estimated by Hanford SIM Rev. 2 for a waste site receiving multiple streams during its process history over a year represent an "instantaneous concentration" and do not necessarily account for highly variable contaminant contributions.

The assumption of ideal mixing tied to the lack of accounting for progressive losses throughout the disposal sites make the size and location of discharge volumes potentially important when evaluating losses to the environment. The interaction between the ideal mixing assumption, the differences in representing the discharges and inventories as point sources rather than spread over an area with regard to infiltration rates, and the boundary conditions associated with the introduction of waste into the system can bias contamination distribution using the Hanford SIM Rev. 2 assumptions.

In many cases, the observed discrepancies between Hanford SIM Rev. 2 estimates and the reference values (Diediker 1999) result as a function of the smoothing assumptions used to develop general waste stream descriptions (composition and uncertainties) or may be attributable to one or more actions as a function of processing or human error within operations (e.g., dilution, rework, mixing, documentation error, process excursion, or separation and removal of analytes). The site- or batch-specific nature of these actions clouds attempts at highly specific, history-matching efforts. As an example, averaging potentially bimodal or multimodal concentration behavior likely will result in overall inventory agreement with reference values at a consolidated level of resolution for a group of sites but will usually result in poor individual site comparisons. Because of the statistical nature used in describing the waste streams, site inventories, and granularity of the data, and the assumptions used in the model, a number of these inventory comparisons may not be within the quantified uncertainty for a series of site–analyte combinations.

Contaminant inventory distribution as a result of infiltration likely does not simply or evenly scale with volume but is highly complex and contingent on the behavior of the analytes and the environment. Volume disposed to a specific site can be seen to play a significant role with regard to inventory; however, for several sites and analytes, other factors are at work. Particularly, certain specific events (e.g., 241-BX-102), solubility considerations (e.g., helium-3), disposal timing (e.g., technetium-99), reprocessing/recycling activity (e.g., cesium-137, strontium-90, plutonium-239) or characteristics of the waste type (e.g., ferrocyanide scavenged waste) can be shown to affect specific contributions to vadose zone inventory more significantly than total volume. The physical and chemical assumptions regarding analyte concentrations having the greatest impact on inventories could be considered in four different categories:

- Analytes that could be considered process water impurities and are principally influenced by overall water volume, such as calcium or sulfate, scale almost directly with volume and are not necessarily related to the chemical processes at the Hanford Site.

- Highly soluble and pervasive analytes such as sodium, nitrate, and chloride have significantly greater losses and are influenced by a combination of volume, composition, and solubility.

- Certain process-specific waste streams were highly enriched in specific analytes (e.g., Z Plant waste for plutonium and carbon tetrachloride). Although the volume of these wastes was relatively low, they represent the majority of the source term observed for these analytes. This feature is also evident when evaluating tank leak and unplanned release losses.

- Wastes with solids or the potential to form solids because of transient increases in solubility can disproportionately influence inventory. The default condition for Hanford SIM Rev. 2 is to not incorporate entrained solids. Where the potential exists for affecting the inventory of a specific site because of the presence of entrained solids, calculations are performed, using data-derived solubility limits, to estimate the expected range of values for volume percent solids. This method of estimating solubility used significant amounts of contemporary data, which is a potential source of bias—especially where the solubility conditions in the tank or process were highly unstable.

The principal factor influencing the model output is the degree to which reliable quantitative descriptions could be provided for the inputs. In the case of the radionuclides, a significant risk is that the uncertainty definitions are merely a reflection of the ORIGEN2 results. However, the physics and mathematics guiding the nuclear reactions used in plutonium production are known very well, and the ORIGEN2 results for Hanford are considered to be reliable.

The extensive use of the HDW Model (Higley et al. 2004) and latest ORIGEN2-DKPRO output (Watrous et al. 2002) provides a substantial technical foundation for this model, but it is a potential limitation on Hanford SIM Rev. 2 as well. Although the sources, methods, and assumptions used in these contributing models and their outputs are conventionally accepted and have gone through extensive review and analysis, there are inherent limitations on their performance. These limitations are discussed in detail in Higley et al. (2004).

The use of tank data or other historical process data in modeling was evaluated closely for appropriateness in each potential application. The effects of time and waste management operations have compromised much of the contemporary tank waste composition data for use in this modeling effort; for this reason, direct use of current tank sampling data is extremely limited. Thus, very little contemporary sampling data (1989 to present) are used to develop waste stream descriptions or uncertainty definitions; further, no best-basis inventory data are used to quantify inventory. Furthermore, many of the waste streams disposed to the ground were never introduced to the tanks.

Sample data are used indirectly as part of Hanford SIM Rev. 2 to quantify the solubility behavior of a variety of analytes more closely in the tank waste/disposal environment. This contemporary information is used to calibrate the solubility subroutine in the HDW Model (Higley et al. 2004). This assumption is also acknowledged to be a potential source of bias. However, because of the highly complex solubility environment in the process plants and in the waste tanks, this approach is considered an acceptable, practical compromise in modeling the system.

As documented in several references (Healy 1953; Ruppert and Heid 1954; Paas and Heid 1955; Abrams 1956), significant challenges existed in obtaining surveillance data. There is often sufficient evidence to call the reference data values into question, either from the inconsistencies/discontinuities observed in the record or in the calculation of the reference value (e.g., arithmetic error, changes in bases, or refinement in judgment regarding the physical/chemical behavior of analytes in the tank–canyon environment). Some of these observed reasons include the following:

- inconsistent data-gathering during the early Hanford Site production era, resulting in annual inventory values that may have varied by as much as two to four times the presented value
- changes in analytical procedures that resulted in changes in observed values by a factor of 100
- use of detection limit values in quantifying inventory, but no indication given that the presented value represents an upper bound
- the uncertain impact of colloidal particles on inventory in certain disposal sites where the conditions for their potential presence are intermittent and unpredictable
- consolidation of the waste tank–disposal site inventory into a single reported value instead of reporting individual site inventories.

Where these particular instances influence the evaluation of a site, the disagreement with the reference value is acknowledged and the technical reasoning and references are included in an exceptions table in the user's guide for Hanford SIM Rev. 2 (Simpson et al. 2007). The exceptions table appears also in Corbin et al. (2005) as Table 6-32. If the error is correctable, the correction is incorporated and the site remains as part of the evaluation process. If the error or basis change does not allow for a clear evaluation of the model estimate versus the reference data, it is excluded from the evaluation. At this time, if the reference data are in conflict with currently accepted convention regarding process behavior or tank chemistry, the accepted convention represents the baseline.

Significant limitations of Hanford SIM Rev. 2 are associated with the assumptions made regarding the presence or absence of an unquantified analyte for a non-HDW Model waste stream and the assumption regarding the separations processes being well defined and operated within specifications. The estimation of trace analytes using “less-than” values, assuming contamination and dilution levels for non-HDW Model waste streams, or assuming nonquantified analytes equal zero is an extremely speculative exercise that resists rigorous quantification. Furthermore, the batch-to-batch processing variability and the different degrees of effectiveness with which solids were entrained and lost to the ground or retained in the tank–canyon system can be highly specific to a site or timeframe and not be representative of the total campaign associated with an operation or disposal site when abstracted to an overall model-level assumption, thus introducing a potential source of bias.

Additional problems associated with the reference and surveillance data are that some are restricted from open publication at this time. The sources all are intermittent as a function of time and not comprehensive in scope. Obtaining an equivalent contemporary field data set for Hanford SIM Rev. 2 or even a select number of locations is unlikely because of the costs involved. Field samples have significant limitations, and a standard convention for interpreting those data and deriving a waste site inventory does not currently exist. Because of the inherent spatial variability associated with field data and its high costs and relative scarcity, a corresponding, independent sample data set providing inventory estimates for waste sites does not exist and cannot be used to validate Hanford SIM Rev. 2.

The application of the mass balance boundary is that the mean Hanford SIM Rev. 2 values for a specific analyte summed over all disposal sites must be less than or equal to the total losses of that analyte from the tank–canyon system. This action is currently performed by the user as a function of defining the inputs. There is no automated “rebalancing” function in the software portfolio if a change to a waste stream is made.

Because the probability distribution functions applied to the inputs are currently left unconstrained and each analyte is treated as an independent variable, summing extreme values (such as using all 90th percentile values) for a specific analyte over all disposal sites to derive a global inventory will result in an unrealistic soil inventory estimate for that analyte. Thus, as a result of the constrained source terms (both overall and for each separation process), if an extremely high inventory value is selected for a site or a series of sites, the availability of that analyte is diminished for the remaining sites, and there needs to be a concomitant number of sites with extremely low inventories of that analyte to maintain the mass conservation boundary.

Another part of the mass balance assumption for Hanford SIM Rev. 2 is on the tank–canyon–disposal-site system. As a simplification, no losses are assumed to the atmosphere. This interpretation of mass balance boundary condition is an assumption with significant consequence. The impact of this assumption is that for purposes of comparison to reference data, the tank–canyon–disposal-site system is considered “closed,” even though in actuality there are likely unquantified losses to the environment.

The role of atmospheric releases for volatile analytes such as helium-3 and iodine-129 can significantly impact soil inventory estimates if these losses can be better quantified and validated. Thus, whatever the initial, decayed, production values from ORIGEN2 are, the sum of the mean amounts in the tanks, canyons, and lost to the ground for purposes of evaluation must be equal to that initial production amount. These conditions assist in evaluating the results and help ensure that the soil inventory results maintain mass balance within the documented waste volumes disposed and analyte masses produced or used at the Hanford Site.

Although Hanford SIM Rev. 2 waste streams lose some of the previously mentioned mass/charge balance character imposed by the HDW Model because the analytes being tracked in Hanford SIM Rev. 2 do not include water (a principal mass contributor) and hydroxide (used to enforce charge balance), the simplifications used in representing the waste stream do not appear to degrade the results significantly. This feature constrains the potential contributions of most analytes considerably. The remaining possible variations allowed by the largely unconstrained behavior of the component uncertainties are considered reasonable under the broad variety of potential disposal conditions.

Logical extensions of contemporary waste stream data for analogous (but data-sparse) situations in the absence of early Hanford Site surveillance information are used. When changes in reactor production behavior (e.g., changes to fuel cladding, fuel element design changes, or reactor operating power) are observed to occur as a function of time but the basic chemical process remains unchanged, new uncertainties based on that change are derived and assigned to that waste stream, without changing the base waste composition. Additionally, in cases where later waste production conditions existed that could be assumed analogous to earlier Hanford processing conditions and the surveillance data were collected from these later data, these data are assumed to be suitable representations of those earlier process conditions and used in Hanford SIM Rev. 2.

Other available process and surveillance data were used in defining and assigning waste streams that were not included in the HDW. However, the nomenclature used in the various references describing the wastes being disposed is often ambiguous with regard to waste assignment to the disposal site with little supporting analytical data, and, thus, it is open to interpretation.

Mis-assignment of waste streams with generic designations but widely varying compositions as a function of their generation and disposal during production (e.g., a label of PUREX condensate could potentially be considered either PUREX process condensate—this waste results from the direct reduction of high-level waste and likely has significant radionuclide content; or PUREX steam condensate—this waste is the result of non-contact cooling water or process water that has little if any contamination) is a potentially significant error requiring an independent data source to provide a means of correction. These discrepancies require judgment on the part of the analyst to resolve, and in certain cases, the decision to accept one source over another may have introduced an error. Thus, there is the opportunity to introduce human error/differences in technical judgment as part of the data input process of Hanford SIM Rev. 2.

References

Abrams RB. 1956. *Metal Recovery Waste Scavenging Program*. HW-43066, General Electric Company, Richland, Washington.

Corbin RA, BC Simpson, MJ Anderson, WF Danielson III, JG Field, TE Jones, and CT Kincaid. 2005. *Hanford Soil Inventory Rev. 1*. RPP-26744, Rev. 0, CH2M HILL Hanford Group, Inc., Richland, Washington.

Diediker LP. 1999. *Radionuclide Inventories of Liquid Waste Disposal Sites on the Hanford Site*. HNF-1744, Fluor Daniel Hanford, Inc., Richland, Washington.

Healy JW. 1953. *Release of Radioactive Wastes to Ground*. HW-28121, General Electric Company, Richland, Washington.

Higley BA, DE Place, RA Corbin, and BC Simpson. 2004. *Hanford Defined Waste Model, Rev. 5*. RPP-19822, Rev. 0, CH2M HILL Hanford Group, Inc., Richland, Washington.

Paas HJ and KR Heid. 1955. *Radioactive Contamination in Liquid Wastes Discharged to Ground at Separation Facilities though June 1955*. HW-38562, General Electric Company, Richland, Washington.

Ruppert HG and KR Heid. 1954. *Summary of Liquid Radioactive Wastes Discharged to the Ground – 200 Areas July 1952 to June 1954*. HW-33591, General Electric Company, Richland, Washington.

Simpson BC, RA Corbin, MJ Anderson, and CT Kincaid. 2007. *Hanford Soil Inventory Model (SIM) Rev. 2 User's Guide*. PNNL-16925, Pacific Northwest National Laboratory, Richland Washington.

Watrous RA, DW Wootan, and SF Finfrock. 2002. *Activity of Fuel Batches Processed Through Hanford Separations Plants, 1944 Through 1989*. RPP-13489, Rev. 0, CH2M HILL Hanford Group, Inc., Richland, Washington.

Distribution

No. of
Copies

OFFSITE

B. C. Simpson
Vivid Learning Systems, Inc.
5728 Bedford Street
Pasco, WA 99301

M. J. Anderson
Vivid Learning Systems, Inc.
5728 Bedford Street
Pasco, WA 99301

R. A. Corbin
Columbia Energy and Environmental Services, Inc.
1806 Terminal Drive
Richland, WA 99352

No. of
Copies

ONSITE

20 **DOE Richland Operations Office**

J. G. Morse (20)	A6-38

29 **Pacific Northwest National Laboratory**

R. W. Bryce	E6-35
P. W. Eslinger	K6-52
M. D. Freshley	K9-33
C. T. Kincaid (20)	K9-33
G. V. Last	K6-81
W. E. Nichols	K9-33
B. E. Opitz	K6-75
M. B. Triplett	K6-52
Hanford Technical Library (2)	P8-55