PNWD-2156 HEDR
UC-000

HEDR Model Validation Plan

Hanford Environmental
Dose Reconstruction Project

B. A. Napier
R. O. Gilbert
J. C. Simpson
J. V. Ramsdell, Jr.
M. E. Thiede
W. H. Walters

June 1993

Letter report prepared for
the Technical Steering Panel and
the Centers for Disease Control and Prevention
under Contract 200-92-0503(CDC)/18620(BNW)

Battelle
Pacific Northwest Laboratories
Richland, Washington 99352

Preface

The Hanford Environmental Dose Reconstruction (HEDR) Project was prompted by mounting concern about possible health effects to the public from more than 40 years of nuclear operations at the U.S. Department of Energy's (DOE) Hanford Site near Richland, Washington. The primary objective of the HEDR Project is to estimate the radiation dose (with descriptions of the uncertainties inherent in such estimates) that individuals could have received as a result of radionuclide emissions since 1944 from the Hanford Site. An independent Technical Steering Panel (TSP) directs the work on the project which is conducted by Battelle, Pacific Northwest Laboratories (BNW) under contract with the Centers for Disease Control and Prevention.

The HEDR Project work is conducted under several technical and administrative tasks, among which are the Technical Integration and Statistics Tasks. The staff on these tasks provide the technical guidance, coordination, and communication among other technical tasks, as well as ensuring the appropriate use of statistics and statistical methods.

These efforts include the model reliability analysis, a shared activity between the two tasks. Validation of the model outputs is part of the analysis of the various models. Results of the activities outlined in this plan will be included in the HEDR Model Reliability Report, Milestone 0803B. This report fulfills Milestone 0205E.

Summary

The Hanford Environmental Dose Reconstruction (HEDR) Project has developed a set of computational "tools" for estimating the possible radiation dose that individuals may have received from past Hanford Site operations. This document describes the planned activities to "validate" these tools. In the sense of the HEDR Project, "validation" is a process carried out by comparing computational model predictions with field observations and experimental measurements that are independent of those used to develop the model.

Scope

A high level of coverage of the various space/time/pathway combinations used in the primary dose calculations would lead to the most rigorously defensible validation. Data are not available to support such an ambitious validation program. Contemporaneous data do not address all the necessary pathways, over space or over time, needed to provide a complete validation. The data sets that have been selected for validation were chosen to provide the best examples of coverage of the domain in time, in space, and for as many pathways as possible.

Approach

The general philosophy of this plan is to compare the calculated values of dose, or of the surrogate measurement closest to dose available (e.g., concentrations of radioiodine in sagebrush), with the measurements. The purpose is primarily to characterize the comparisons with respect to the spatial, temporal, or pathway elements of the available data. The approach is to provide descriptive statistics, rather than formal "hypothesis testing" statistics.

Results

This report discusses the data sets available to address the general validation goals. Fourteen specific procedures are provided to be used to prepare the existing data sets, estimate environmental contamination levels or doses, and compare the observed and predicted data. Comparison of the HEDR models is described through benchmarking with other international models in the International Atomic Energy Agency's Coordinated Research Program on Validation of Radionuclide Transport Models for Terrestrial, Urban, and Aquatic Environment. A generic description of deliverables of each individual validation exercise is provided. Future uses of the validation results are addressed. Projected schedule and costs of performing these analyses are provided.

Contents

Figures

Tables

1.0 Introduction

The Hanford Environmental Dose Reconstruction (HEDR) Project has developed a set of computational "tools" for estimating the possible radiation dose that individuals may have received from past Hanford Site operations. This document describes the planned activities to "validate" these tools. In the sense of the HEDR Project, "validation" is a process carried out by comparing computational model predictions with field observations and experimental measurements that are independent of those used to develop the model. A model cannot be considered validated until sufficient comparisons have been performed to ensure an acceptable level of predictive accuracy. These comparisons must be performed over the range of conditions in which the model may be applied. The acceptable level of accuracy is subject to judgment and will vary depending on the specific problems or questions being addressed by the model (IAEA 1989).

Thus, validation is one part of the overall model "reliability" analysis being performed for the HEDR models. The full reliability analysis, including the code testing and verification, validation process, and uncertainty and sensitivity analyses, will be documented in the HEDR Model Reliability Report.

1.1 The HEDR Dose Estimation Toolbox

An entire suite of computerized models has been prepared to support the HEDR dose analysis capability. The validation exercise includes each of the following components, which perform the indicated functions:

- Reactor Model (RM)
  - Radioiodine content in discharges from B, D, and F reactors, 1944-1949
- Atmospheric Source Term Release Model (STRM)
  - Radioiodine release rates from B and T separations plants, 1944-1949
- Atmospheric transport model (RATCHET)
  - Radionuclide time integrated air concentrations in the HEDR spatial domain, 1944-1972
  - Radionuclide surface deposition rates in the HEDR spatial domain, 1944-1972
- Environmental accumulation model (DESCARTES or PILOT)
  - Accumulation of radioiodine in different plant products at numerous space/time combinations
  - Accumulation of radioiodine in different animal products at numerous space/time combinations
- Individual dose model (CIDER or PILOT)
  - Radionuclide body burdens in humans of various ages and sexes from ingestion and inhalation pathways

- River source term release model
  - Emission rates of 6 radionuclides released to the Columbia River, 1944-1972

- River transport model (CHARIMA)
  - Monthly concentrations in Columbia River water of 6 radionuclides at 12 locations

- River dose model (CRD)
  - Monthly concentrations in Columbia River fish for multiple radionuclides at 12 locations
  - Annual concentrations in Willapa Bay oysters for 2 radionuclides
  - Annual concentrations in Columbia River salmon for 2 radionuclides.

- Air pathways scoping spreadsheets (2)
  - Concentrations in air for up to 12 radionuclides
  - Concentrations in plant and animal products for up to 12 radionuclides
  - Body burdens in humans for up to 12 radionuclides.

1.2 Model Validation Strategy

Complete validation can be said to consist of four steps: 1) peer review of the models as they are being developed, 2) verification of the computer implementations as the codes are developed, 3) verification of the assumptions and parameters going into the codes, and 4) comparisons of the results to actual measurements. The HEDR models have been subjected to numerous reviews by the TSP and others (e.g., the CDC/TSP review of the RATCHET code, extensive discussions with the TSP during the development of the surface water modeling effort). Independent testing of the various codes has been underway to assure correct implementation of the models. The assumptions and parameters have been independently published and continue to undergo scrutiny (e.g., the commercial milk distribution model initially prepared by Beck et al. (1992) and the environmental accumulation and dose model parameter report (Snyder et al. 1992)). This plan describes the comparison of calculated results to historical measurements. Each of these steps will be formally documented in the HEDR Model Reliability Report scheduled for Fiscal Year 1994.

The HEDR models are used to describe the potential for radiation dose to individuals living in a large spatial area, over long periods of time, by a number of potentially-important exposure pathways. It would be highly desirable to validate the various models at points throughout the spatial domain, in areas of high deposition, light deposition, and sporadic or minimal deposition. It would be desirable to observe the variation in time of radionuclide concentrations in each of the pathways at these various locations. A high level of coverage of the various space/time/pathway combinations used in the primary dose calculations would lead to the most rigorously defensible validation. As described in the following sections, data are not available to support such an ambitious validation program. Contemporaneous data do not address all the necessary pathways, over space or over time, needed to provide a complete validation. The data sets that have been selected for validation were chosen to provide the best examples of coverage of the domain in time, in space, and for as many pathways as possible. We believe that the tests defined in this plan provide a reasonable set for the needs of the project, and that sufficient coverage of the spatial, temporal, and pathway variables is achieved for the demonstration of the adequacy of the HEDR approach and implementation.

Evaluation of the results of the validation tests is a necessary component of the validation. The general philosophy of this plan is to compare the calculated values of dose, or of the surrogate measurement closest to dose available (e.g., concentrations of radioiodine in sagebrush), with the measurements. The purpose is to understand the differences between the calculated doses and the measurements. Thus, the statistical methods that will be used are aimed at describing these differences so that the causes can be understood and recommendations for improvements can be made.

1.3 Potential Areas For Model Validation

The models and tools described in Section 1.1 each provide intermediate information used in calculating the radiation dose to particular individuals. It is not possible to validate individual doses, because these were not measured, and no database of individual dose exists. However, radionuclide concentrations were measured at various times and in various media by environmental monitoring programs operated at Hanford and elsewhere. Although insufficient for estimating doses directly, these measurements do provide the possibility of validating portions of the HEDR toolbox for particular times and/or occurrences. It is expected that compilation of a sufficient number of these component validations will demonstrate the general reliability of the HEDR dose estimation toolbox.

The progression of the calculations required to estimate dose via the atmospheric pathway is illustrated in Figure 1.1. The progression for estimation of doses via the Columbia River pathway is shown in Figure 1.2. The individual computer models are represented as boxes; intermediate data transfers are listed. Ideally, it would be desirable to validate the outputs of each model at each of the intermediate points. However, historical information is not available for all intermediate areas. Because the intermediate information near the end of the computations is based on information calculated at points earlier, in several instances it is possible to provide an "inferred" validation, where, if a later step can be validated, the validity of the earlier information is implied.

1.4 Data Quality Objectives

Data Quality Objectives (DQOs) for the overall validation effort are not stated in the HEDR workplan. However, the Task 02 DQOs for accuracy ("...incorporate physical phenomena as accurately as any other nationally known code for similar use") and comparability ("models developed by the project should compare within factors of two to three of international models") can be construed to apply (Shipler 1993, pp. 3.8-3.9). In addition, several of the individual models have DQOs associated with their development. The air and water source term models require that "accuracy will be verified by comparison to historical measurements" and for comparability that "cross-comparisons will be made with multiple historical reference sources where possible" (ibid., pp 4.6, 4.11-4.12). For the atmospheric transport model RATCHET, the DQO for accuracy "...is that bias in monthly average air concentrations be less than a factor of three. Statistical evaluation of the stochastic realizations will be performed and compared to monitoring data for selected locations" (ibid., p. 5.3), and that "model accuracy will be evaluated....by direct comparisons of model output with available monitoring data" (ibid., p. 5.23). For the surface-water transport model CHARIMA, the DQO for comparability "...is for the results to be comparable to existing environmental monitoring reports,

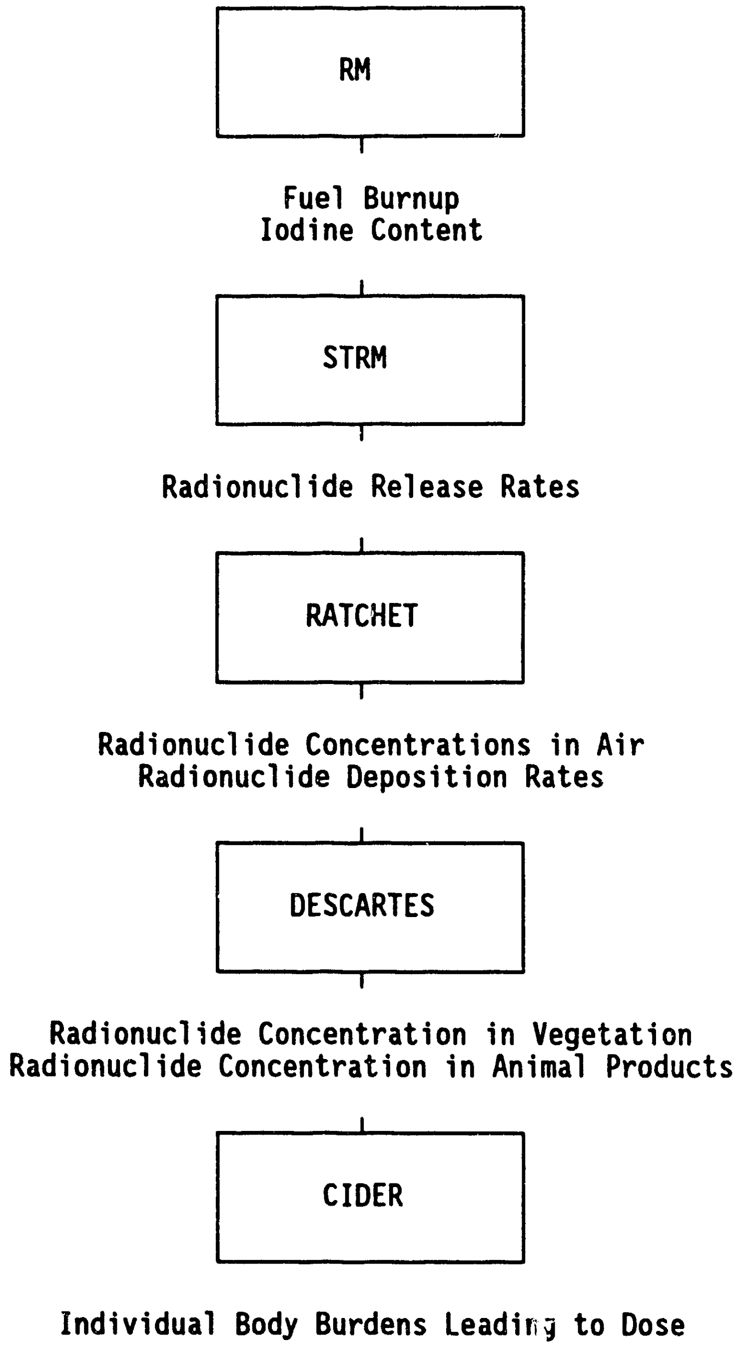

Figure 1.1. Computer Codes and Intermediate Calculated Data Used to Estimate Individual Dose from the Atmospheric Pathway

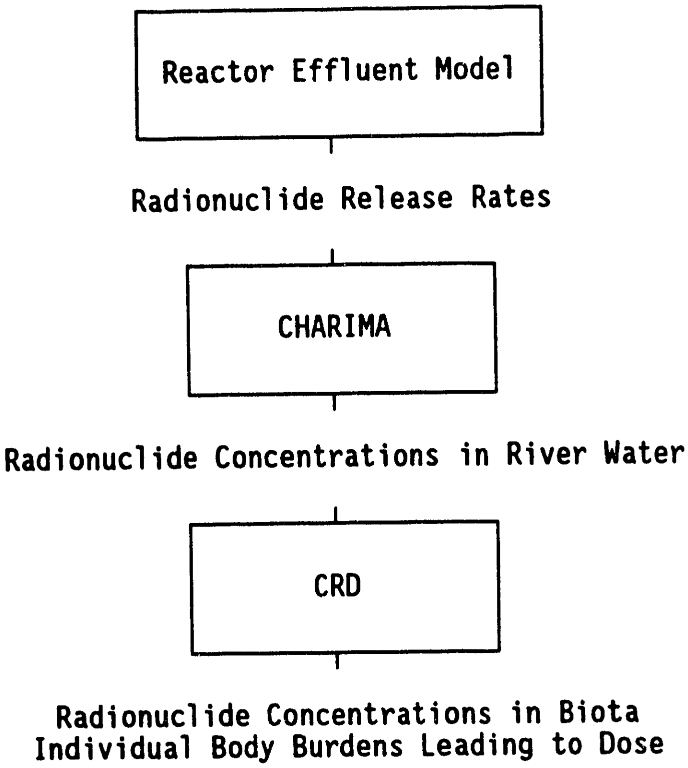

Figure 1.2. Computer Codes and Intermediate Calculated Data Used to Estimate Individual Dose from the Columbia River Pathway

publicly available dose estimates, and current understanding by cognizant agencies of river hydraulics" (ibid., p. 5.15). For the environmental accumulation models, the DQO for comparability is "...that the model results are 'reasonable' when compared with available monitoring and dosimetry data. ... Other model output is generally expected to be within a factor of three of available validation data. Model validation studies, including direct comparison of model results to historical monitored data, will be used to measure this objective." (ibid., p. 8.7).

It is apparent that implementation of this model validation plan will address the DQOs of most of the technical tasks on the HEDR Project. The IAEA recommends (IAEA 1989, p.22) that performance standards be established prior to validation testing. The collected DQOs described previously essentially set that standard. It is apparent that comparisons of predicted versus monitored historical data must be made, and that the general objective is that the overall bias of the results be less than a factor of three. If this standard is not met, then additional investigations are needed to determine whether to conduct an increased validation effort, improve the model, and/or adjust the desired performance standard.

A matrix will be prepared listing all of the HEDR models, the DQOs applicable to each, and the particular validation exercises that demonstrate the fulfillment of the DQOs.

1.5 Document Overview

Section 2.0 of this report discusses the data sets available to address the general validation goals described in Section 1.2. Section 3.0 provides 14 specific procedures to be used to prepare the existing data sets, estimate environmental contamination levels or doses, and compare the observed and predicted data.

The comparison of the HEDR models through benchmarking with other international models is described in Section 4.0. The IAEA's Coordinated Research Program on Validation of Radionuclide Transport Models for Terrestrial, Urban, and Aquatic Environments (VAMP) is discussed in that section.

The generic description of deliverables of each individual validation exercise is provided in Section 5.0. Future uses of the validation results are addressed in Section 6.0. Projected schedule and costs of performing these analyses are provided in Section 7.0.

2.0 General Validation Methods

The overall approach to validation is to compare calculated radionuclide concentrations in environmental media with historical monitoring data. As discussed in Section 1.2, a high level of coverage of the various space/time/pathway combinations used in the primary dose calculations would lead to the most rigorously defensible validation. Data are not available to support such an ambitious validation program. Contemporaneous data do not address all the necessary pathways, over space or over time, needed to provide a complete validation. The data sets that have been selected for validation were chosen to provide the best examples of coverage of the domain in time, in space, and for as many pathways as possible.

2.1 Available Data Sets

Several applicable sets of historical monitoring data have been identified. Some of these are Hanford Site monitoring data, as summarized and reported in HEDR Phase I (e.g., Denham et al. 1993) or prepared but not published yet by the project (e.g., "Scenario H" for VAMP). Some consist of continuing Hanford Site occupational records, such as the Site whole-body counting records (e.g., Swanberg 1962). Others have been discussed previously on the project, such as the historical "thyroid count data" (Ikenberry 1991). A list of potential data sets was provided to the Technical Steering Panel.[a] The subset of that list selected for detailed investigation is presented in Table 2.1.

As is indicated in Table 2.1, some of the data sets contain information sufficient to contribute directly to the validation of individual components of the HEDR computational system. An example is the data collected during the 1963 accidental release from the PUREX plant. This data set includes direct measurements of the releases from the PUREX stack, monitored air concentrations of iodine-131, monitored vegetation concentrations, monitored iodine concentrations in milk from cows grazing in the area, and actual thyroid counts of children drinking the locally produced milk (Soldat 1965). Each component of this data set can be used directly for comparisons. In contrast, the iodine-131 thyroid counts of workers taken in the late 1940s are only directly useful for comparison with the CIDER output. However, if the CIDER calculations are correct, they imply that the entire system from radionuclide release through transport and accumulation to exposure is reasonable. Validation exercises are thus appropriate at several levels. The selected procedures described in the following sections are designed to incorporate these various levels.

2.2 Calculational Considerations

The historical monitoring information may be used in a number of ways. In some cases, individual monitoring values must be used to represent contamination over a large area, or over a

(a) Letter, "HEDR Project Model Validation Strategy," from D. B. Shipler, BNW, to J. E. Till, TSP, and M. R. Donnelly, CDC, March 5, 1993.

Table 2.1. Data Sets Appropriate for Validation, with Applicable Portions of the HEDR Dose Calculation Process

Iodine-131 concentrations in air, vegetation, milk, and children, 1963 uncontrolled release
- RATCHET- dispersion, deposition
- DESCARTES- environmental media concentration
- CIDER- body burden

Iodine-131 concentrations in air, vegetation, 1949 controlled release (Green Run)
- RM- content in reactor discharges
- STRM (air)- hourly releases
- RATCHET- dispersion, deposition
- DESCARTES- environmental media concentrations

Iodine-131 concentrations in vegetation, 1940s
- STRM (air)-RATCHET-DESCARTES integrated vegetation concentrations

Iodine-131 worker thyroid counts, 1940s
- STRM (air)-RATCHET-DESCARTES-CIDER- Thyroid burdens

Krypton-85 concentrations in air, 1984-1988
- RATCHET monthly average air concentrations

Zinc-65, Sodium-24 (river nuclides) worker whole body counts, 1959-1972
- STRM (river)-CHARIMA-CRD-body burdens

Nuclides in River, 1960s
- CHARIMA- concentrations in water
- CRD- concentrations in fish and shellfish

VAMP data sets and scenarios
- DESCARTES-CIDER- environmental media concentrations, body burdens

long period. In other cases, several data points collected over a small area may be combined to provide a measured distribution of iodine deposition for a certain time, or several data points collected at several places over a brief time period may be combined to provide a distribution of a deposition pattern in space. Some of the HEDR computational models provide output distributions of estimated values to account for uncertainty. Others (such as the river dose model) provide only single values. Therefore, the possible comparisons of monitored and calculated values include distributions against point values, point values against point values, and distributions against distributions. The individual validation exercises must account for these permutations and provide appropriate statistical interpretations.

3.0 Specific Procedures

This section describes the specific validation exercises to be undertaken by the HEDR Project. Each procedure describes the assessment question being addressed, that is, the intermediate data being calculated and their ultimate relationship to the individual dose. It describes the available data set being used, and its limitations. It describes the models to be evaluated and the development of the input data required to run them. The type of data (i.e., single point, distribution) available from the historical data and calculated by the pertinent model(s) is indicated. A description of the statistical evaluation techniques is provided.

For the atmospheric pathway, Section 3.1 addresses limited validation of the Reactor Model (RM). Sections 3.2 through 3.4 address the validation of the atmospheric source term-transport-environmental accumulation model in time sequence at selected locations. Sections 3.5 and 3.6 address the validation of the source-through-accumulation models across space at selected times. Sections 3.7 and 3.8 address the validation of the system from source through individual uptake. Section 3.9 addresses the use of current data to validate the atmospheric dispersion model.

The validation of Columbia River pathway environmental media is addressed in Sections 3.10 through 3.13. Human exposures via river pathways are addressed in Section 3.14.

3.1 RM Burnup Calculations

The main design output from the reactor model is the content of iodine in the fuel at discharge. The Hanford reactors were built for the production of plutonium using neutrons from the fission of uranium; the iodine is a product of the uranium fission. The content of iodine in the fuel was never measured directly, therefore, direct validation is not possible. However, other information was recorded so that calculations based on known reactor physics can provide highly reliable estimates of the iodine content of the discharged fuel. The primary recorded data were the average reactor power levels and the burnup of the fuel at discharge. If the reactor model can accurately predict the burnup of discharged fuel, then the iodine content of the fuel can be known with little uncertainty and the model is valid for use.

The reactor model is described by Heeb (1993, Appendix A). This reference also provides a summary of comparisons of the calculated discharge burnup with recorded burnups from Hanford P-Department Reports (General Electric Company 1947). The comparisons are depicted in Heeb's Figures A.2, A.3, and A.4 (1993). As stated by Heeb (1993, p. A.1):

> "In general, the computed and measured exposure matches fairly well. Of the 226 discharges, 92.5% have estimated exposures within 10% of the reported exposures and 99.9% have estimated exposures within 15% of the reported exposures. Maximum discrepancies are within the ±15% range."

For the purposes of validating the reactor model, this information is considered to be sufficient, and no additional work is proposed.

3.2 Time-Sequence Data For Richland, 1946

A substantial number of vegetation samples was collected regularly at standardized locations in the City of Richland, Washington, beginning in late 1945. These provide the possibility of validating a series of calculations over time for this location.

3.2.1 Assessment Question

The assessment question addressed is, "What is the time history of daily concentration of iodine-131 on vegetation (sagebrush) in Richland (HEDR atmospheric dispersion grid node number 878) during 1946?"

3.2.2 Available Richland Data

Hanford historical monitoring data are available to the HEDR Project for the period beginning in mid-1945 through the present. These data are being made available to the public in Denham, et al. (1993) and another forthcoming report. The developing Hanford monitoring of that period focussed on vegetation (usually sagebrush). The publications provide the original data and recently developed modifications to account for biases in the measurements that were contemporaneously unknown. These data are essentially all that are available for these time periods that are of high interest to the HEDR Project. The data are uneven in coverage of space (most monitored locations are either on or close to the Hanford Site) and in coverage of time (the monitoring, with a few notable exceptions, was not routinely performed at repeated locations). Thus, while there are over 3500 samples reported for the year 1946, Richland has a complete history for each month of 1946 consisting of a total of about 550 values.

To illustrate the relative richness of this data source, consider that in 1946, around 3500 samples were taken. Of these, about 1450 were taken on the Hanford Site, and about 1400 were taken in the Tri-Cities/Benton City area (see Sections 3.3 and 3.4). Another 175 or so were taken at various points along Van Giesen Road, west from Richland. This leaves fewer than 500 samples taken throughout the remainder of the Pacific Northwest over the course of the year. Ten were taken in The Dalles, Oregon at various times. A total of 7 were taken in Walla Walla, 13 in Spokane over a 2-day period, 3 in Pendleton, Oregon, 6 in Weston, Oregon, 5 in Toppenish, etc. No location had more than 15 samples taken, usually in only one to two sampling expeditions. The remainder of the data, outside the immediate vicinity of Hanford, are essentially worthless for long-term trend analysis.

3.2.3 Models to be Evaluated

Vegetation-monitoring data for Richland in 1946 may be compared directly with output of the DESCARTES accumulation model. This will provide indirect validation of the RM and STRM source terms, and the RATCHET dispersion model, as well.

3.2.4 Evaluation to be Performed

Input will be provided to the DESCARTES code from the STRM/RATCHET output database. Output from DESCARTES will be the 100 daily and monthly realization values for sagebrush concentration in Richland, node 878. The resultant daily distributions will be compared with the available daily point values (in some instances, there may be multiple values for each day--these will be treated as points, not as distributions); this will be a distribution-to-point comparison. The resultant monthly distributions will be compared with composited monthly monitoring values. This will be a distribution-to-distribution comparison.

The comparison of measured and modeled values will be conducted using graphical descriptive methods. The daily Richland measurements will be plotted for each day of 1946 on a time-series plot. The plot will also display the box plot of each day's 100 realizations of daily concentrations to permit visual comparisons of the measured and predicted values. This time-series plot will be used to look for patterns of model-prediction bias over time. For example, the plot will help determine if key deposition events, as indicated by the measured values, are in agreement with the predicted values. An example of the type of plot to be generated is shown as Figure 3.1.

In addition to the time-series plot, the ratio of the median predicted value to the median measured value will be computed for each day. These daily ratios will be plotted on a second time-series plot to examine visually whether the daily ratio remains within acceptable bounds, for example, a factor of three. An example plot of this nature is shown as Figure 3.2. A box plot that describes the distribution of the daily ratios will provide a summary of the results. The proportion of ratios that exceeds various values will also be reported to aid in summarizing the performance of the model at the Richland node.

In production mode, DESCARTES will output only the average daily value for each month. Hence, there is interest in comparing these average values with the average of the measured daily values over the month. However, for some days, no measurements were made. So, the average daily computed value for the month will be adjusted to take into account the missing days. This adjustment will consist of computing the average daily values only for days for which measured data exist.

We note that temporal correlation among the measured values creates problems in the interpretation of statistical tests. Hence, we will rely primarily on descriptive and graphical methods to evaluate the model's performance.

3.3 Time Sequence Data for Kennewick/Pasco, 1946

A substantial number of vegetation samples was collected regularly at standardized locations in the Cities of Kennewick and Pasco, Washington, beginning in late 1945. Several locations within each city were routinely monitored. The most extensive data set for these locations was collected in the year 1946, the year of second-highest radioiodine emissions. Both cities, separated only by the Columbia River, fall within a single HEDR atmospheric dispersion grid node (number 838). The numerous data from the two towns combined provide the possibility of validating a time series of calculations for this location.

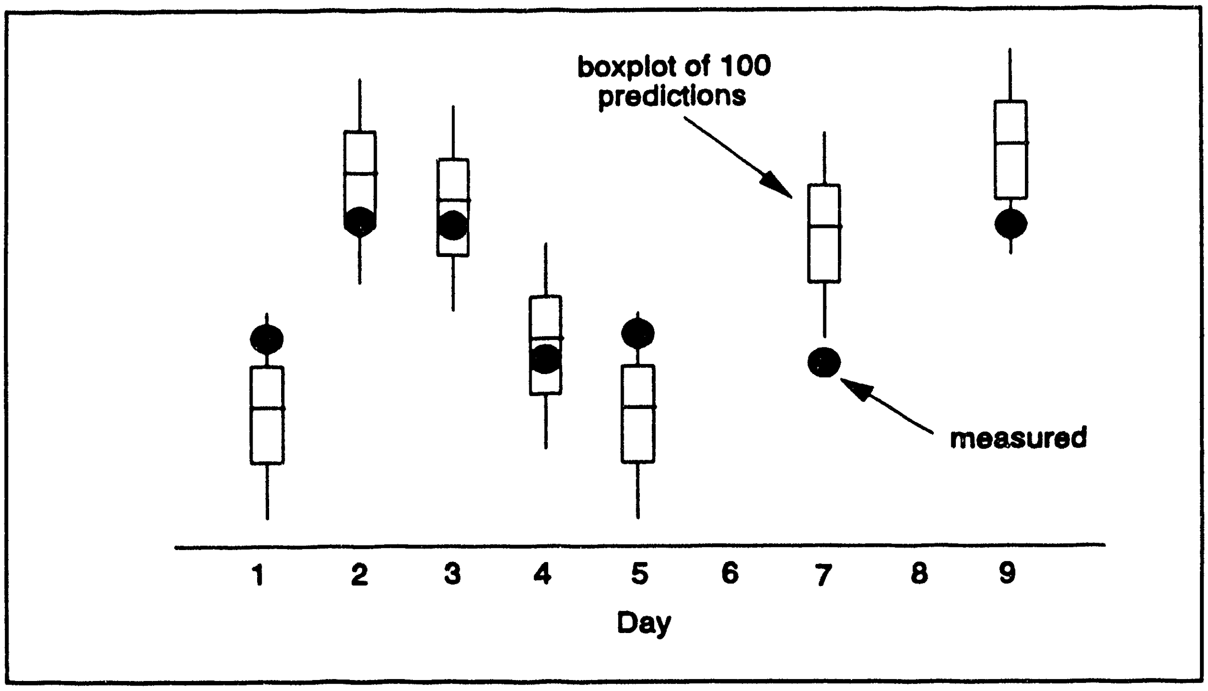

Figure 3.1. Example Time-Series Plot of Measured and Predicted Concentrations

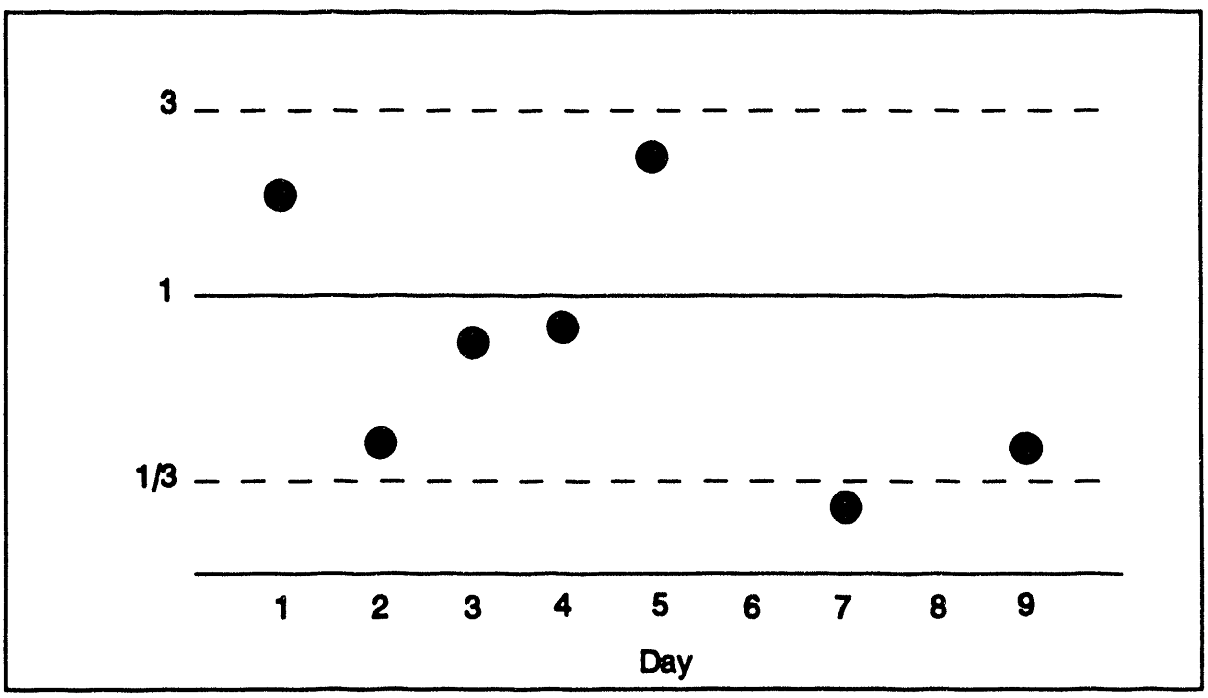

Figure 3.2. Example Time-Series Plot of Measured/Predicted Ratio

3.3.1 Assessment Question

The assessment question addressed is, "What is the time history of daily concentration of iodine-131 on vegetation (sagebrush) in Kennewick/Pasco (HEDR atmospheric dispersion grid node number 838) during 1946?"

3.3.2 Available Kennewick/Pasco Data

Hanford historical monitoring data are available to the HEDR Project for the period beginning in mid-1945 through the present. These data are being made available to the public in Denham, et al. (1993) and another forthcoming report. The developing Hanford monitoring of that period focussed on vegetation (usually sagebrush). The publications provide the original data and recently developed modifications to account for biases in the measurements that were contemporaneously unknown. These data are essentially all that are available for these time periods that are of high interest to the HEDR Project. The data are uneven in coverage of space (most monitored locations are either on or close to the Hanford Site) and in coverage of time (the monitoring, with a few notable exceptions, was not routinely performed at repeated locations). Thus, while there are over 3500 samples reported for the year 1946, for off-Site locations, Pasco and Kennewick combined have a complete history for each month of 1946 consisting of a total of about 645 values.

3.3.3 Models to be Evaluated

Vegetation monitoring data for Kennewick/Pasco, 1946, may be compared directly with output of the DESCARTES accumulation model. This will provide indirect validation of the RM and STRM source terms, and the RATCHET dispersion model, as well.

3.3.4 Evaluation to be Performed

Input will be provided to the DESCARTES code from the STRM/RATCHET output database. Output from DESCARTES will be the 100 daily and monthly realization values for sagebrush concentration in Kennewick/Pasco, node 838. The resultant daily distributions will be compared with the available daily point values (in some instances, there may be multiple values for each day - these will be treated as points, not as distributions); this will be a distribution-to-point comparison. The resultant monthly distributions will be compared with composited monthly monitoring values. This will be a distribution-to-distribution comparison.

The comparison of measured and modeled values will be conducted using graphical descriptive methods as well as more formal statistical tests. The daily Kennewick/Pasco measurements will be plotted for each day of 1946 on a time-series plot. The plot will also display the box plot of each day's 100 realizations of daily concentrations to permit visual comparisons of the measured and predicted values. This time-series plot will be used to look for patterns of model-prediction bias over time. For example, the plot will help determine if key deposition events, as indicated by the measured values, are in agreement with the predicted values.

In addition to the time-series plot, the ratio of the median predicted value to the median measured value will be computed for each day. These daily ratios will be plotted on a second time-series plot to examine visually whether the daily ratio remains within acceptable bounds, for example, a factor of three. A box plot that describes the distribution of the daily ratios will provide a summary of the results. The proportion of ratios that exceeds various values will also be reported to aid in summarizing the performance of the model at the Kennewick/Pasco node.

In production mode, DESCARTES will output only the average daily value for each month. Hence, there is interest in comparing these average values with the average of the measured daily values over the month. However, for some days, no measurements were made. So, the average daily computed value for the month will be adjusted to take into account the missing days. This adjustment will consist of computing the average daily values only for days for which measured data exist. We note that spatial correlation among the measured values creates problems in the interpretation of statistical tests. Hence, we will rely primarily on descriptive and graphical methods to evaluate the model's performance.

3.4 Time Sequence Data For Benton City, 1946

A substantial number of vegetation samples was collected regularly at standardized locations in the City of Benton City, Washington, beginning in late 1945. These provide the possibility of validating a time series of calculations for this location.

3.4.1 Assessment Question

The assessment question addressed is, "What is the time history of daily concentration of iodine-131 on vegetation (sagebrush) in Benton City (HEDR dispersion grid node number 835) during 1946?"

3.4.2 Available Benton City Data

Hanford historical monitoring data are available to the HEDR Project for the period beginning in mid-1945 through the present. These data are being made available to the public in Denham, et al. (1993) and another forthcoming report. The developing Hanford monitoring of that period focussed on vegetation (usually sagebrush). The publications provide the original data and recently developed modifications to account for biases in the measurements that were contemporaneously unknown. These data are essentially all that are available for these time periods that are of high interest to the HEDR Project. The data are uneven in coverage of space (most monitored locations are either on or close to the Hanford Site) and in coverage of time (the monitoring, with a few notable exceptions, was not routinely performed at repeated locations). Thus, while there are over 3500 samples reported for the year 1946, Benton City, Washington, has a complete history for each month of 1946 consisting of a total of about 200 values.

3.4.3 Models to be Evaluated

Vegetation monitoring data for Benton City, 1946, may be compared directly with output of the DESCARTES accumulation model. This will provide indirect validation of the RM and STRM source terms, and the RATCHET dispersion model, as well.

3.4.4 Evaluation to be Performed

Input will be provided to the DESCARTES code from the STRM/RATCHET output database. Output from DESCARTES will be the 100 daily and monthly realization values for sagebrush concentration in Benton City, node 835. The resultant daily distributions will be compared against the available daily point values (in some instances, there may be multiple values for each day - these will be treated as points, not as distributions); this will be a distribution-to-point comparison. The resultant monthly distributions will be compared against composited monthly monitoring values. This will be a distribution-to-distribution comparison.

The daily Benton City measurements will be plotted for each day of 1946 on a time-series plot. That plot will also display the box plot of each day's 100 realizations of daily concentrations to permit visual comparisons of the measured and predicted values. This time-series plot will be used to look for patterns of model-prediction bias over time. For example, the plot will help determine if key deposition events, as indicated by the measured values, are in agreement with the predicted values.

In addition to the time-series plot, the ratio of the median predicted value to the median measured value will be computed for each day. These daily ratios will be plotted on a second time-series plot to examine visually whether the daily ratio remains within acceptable bounds, for example, a factor of three. A box plot that describes the distribution of the daily ratios will provide a summary of the results. The proportion of ratios that exceed various values will also be reported to aid in summarizing the performance of the model at the Benton City node.

In production mode, DESCARTES will output only the average daily value for each month. Hence, there is interest in comparing these average values with the average of the measured daily values over the month. However, for some days, no measurements were made. Thus, the average daily computed value for the month will be adjusted to take into account the missing days. This adjustment will consist of computing the average daily values only for days for which measured data exist.

3.5 Dispersion/Deposition Footprint, April 1946

Data collected by the Hanford monitoring groups in the late 1940s and early 1950s tended to focus on the Hanford Site or adjacent areas. Most detailed monitoring occurred in a few preferred down-wind locations. Sweeps of outlying areas were performed at erratic intervals, the first in early 1946. About 82 samples in the directions of Walla Walla, Ellensburg-Ritzville, and Toppenish-The Dalles were taken on January 12, 1946. About 110 samples extending from Lewiston, Idaho, to Portland, Oregon, and from Moses Lake, Washington, to Pendleton, Oregon, were taken on February 9, 1946. About 83 samples from Ellensburg-Ritzville, Umatilla-The Dalles, and Sprague-Spokane were taken on April 13, 1946. About 114 samples were taken in the southeast-to-northeast quadrant on November 11, 1948. These were the major off-site investigations recorded in the Hanford records prior to the multitude of samples before and after the 1949 "Green Run" (Section 3.6). Only the April set was taken during "growing season" conditions.

3.5.1 Assessment Question

The question to be addressed is, "What was the spatial footprint of radioiodine deposition on April 13, 1946?" This assessment question will address the deposition over all measured nodes within the HEDR atmospheric domain. The April 1946 date is selected to be one within the growing season of that year.

3.5.2 Available Data

It appears from the spatial distribution of the samples taken on April 13, 1946, that three vehicles were sent out with instructions to sample vegetation at intervals on preselected routes. One vehicle made a loop up the Yakima Valley to Ellensburg, then east to Ritzville. Another went north to Ritzville and then east to Spokane. The third went south along the Oregon side of the Columbia River Gorge toward The Dalles. Concentration measurements from approximately 83 vegetation samples from throughout the region are available.

3.5.3 Models to be Evaluated

Vegetation monitoring data for the monitored areas on April 13, 1946, may be compared directly with output of the DESCARTES accumulation model. This will provide indirect validation of the RM and STRM source terms, and the RATCHET dispersion model, as well.

3.5.4 Evaluation to be Performed

Input will be provided to the DESCARTES code from the STRM/RATCHET output database. Output from DESCARTES will be the 100 daily realization values for sagebrush concentration at all nodes for which monitoring data are available. The resultant daily distributions will be compared with the available daily point values; this will be a distribution-to-point comparison. The comparison of measured and modeled values will be conducted using graphical and descriptive methods. The measured values will be displayed on a map of the area. The median predicted concentration for each node in the area will also be displayed on this map. Box plots that show the distribution of the 100 predicted values for each node will be prepared and a metric developed that indicates where in the box plot the measured values fall. This metric, e.g., a value from 1 to 6 (1 indicating below the 5th percentile, 2 indicating between the 5th and 25th percentiles, etc.), will then be plotted on the map at the location of each measurement. The resulting map of indices will be used to look for patterns of poor prediction over the area.

In addition to the spatial metric plot, the ratio of the median predicted value to the median measured value will be computed for each node where measured values exist. A box plot that shows the distribution of these ratios will provide a summary of the results. The proportion of the ratios that fall above 3 or below 1/3 will be determined to help evaluate attainment of DQOs. We note that spatial correlation among the measured values creates problems in the interpretation of statistical tests. Hence, we will rely primarily on descriptive and graphical methods to evaluate the model's performance.

3.6 1949 Green Run

As described by Robkin (1992), one of the singular events in the history of Hanford Site operations was the Green Run experiment that began on December 2, 1949. The experiment was part of the development of monitoring methods for intelligence efforts regarding the emerging Soviet nuclear program. A description of the experiment has been released (Jenne and Healy 1950).

3.6.1 Assessment Question

The assessment question to be addressed is, "What was the spatial distribution of radioiodine deposition following the Green Run release of December 1949?" This assessment question will address the deposition at each of the HEDR nodes for which data are available, based on the before-and-after measurements taken. Therefore, deposition prior to the beginning of the Green Run release must be accounted for.

3.6.2 Available Green Run Vegetation Data

It appears from the available monitoring data that preparations for the Green Run began about two weeks prior to the actual release. Between November 17, 1949, and the start of the experiment in December, about 234 off-site monitoring samples were taken--the most in a coordinated fashion since the start of the program in 1945. These samples were taken in loops up the Yakima Valley to Ellensburg, from Ritzville to Spokane, and down the Columbia River from Umatilla through the Gorge.

Sampling efforts intensified during and after the Green Run release. About 618 samples taken during the month of December 1949 are available from throughout the HEDR atmospheric dispersion domain. Singlevich (1950) and Parker (1950) both report that 1365 vegetation samples were taken--however, many of these were on-site. An additional 100 samples taken in January 1950 are available, but at a more restricted set of off-site locations; less apparent effort went into collecting these.

3.6.3 Models to be Evaluated

The source term will be prepared using the hourly data of the STRM model. Dispersion will be done with hourly inputs to RATCHET. The DESCARTES code will be used to obtain the daily deposition values.

3.6.4 Evaluation to be Performed

Input will be provided to the DESCARTES code from the STRM/RATCHET output database. Output from DESCARTES will be the 100 daily realization values for sagebrush concentration in the HEDR domain for the months of November and December 1949. For each node, the resultant daily distributions will be compared with the available daily point values (in some instances, there may be multiple values for each day--these will be treated as points, not as distributions); this will be a distribution-to-point comparison.

The comparison of measured and modeled values will be conducted using graphical and descriptive methods. For each day the measured values will be displayed on a map of the area. The median predicted concentration for each node in the area for that day will also be displayed on the map. Box plots that show the distribution of the 100 predicted values for each node will be prepared and a metric developed that indicates where in the box plot the measured values fall. This metric, e.g., a value from 1 to 6 (1 indicating below the 5th percentile, 2 indicating between the 5th and 25th percentiles, etc.), will then be plotted on the map at the location of each measurement. The resulting map of indices for each day will be used to look for patterns of poor prediction over the area for that day. In addition, the sequence of daily maps will be compared to detect patterns of poor spatial prediction over time.

In addition to the daily spatial metric plots, the ratio of the median predicted value to the median measured value will be computed for each node where measured values exist. A box plot that describes the distribution of these ratios will provide a summary of the results. The proportion of the ratios that fall above 3 or below 1/3 will be determined to help evaluate attainment of DQOs. This process will be conducted for each daily plot. We note that spatial correlation among the measured values creates problems in the interpretation of statistical tests. Hence, we may rely primarily on descriptive and graphical methods to evaluate the model's performance.

3.7 PUREX Release, September 1963

An event similar to the 1949 Green Run occurred at the PUREX facility in 1963, although by accident and with a much smaller concurrent release. An acute, inadvertent release of iodine from the 60-meter stack at PUREX occurred from September 2 to 5, 1963, as a result of inadvertently charging short-cooled fuel elements into the dissolver. Plant operations were shut down as soon as the abnormal release was detected. Steps were taken immediately to retain as much of the iodine as possible within the plant. Laboratory analyses of stack effluent samples were made. The routine program of environmental monitoring was augmented with additional sampling. Measurements of wind velocity and temperature were made routinely at the site meteorology tower. No significant rainfall occurred in the HEDR domain during the period. No protective measures were taken following the release. There were no significant atmospheric nuclear tests in the several months prior to the accident.

3.7.1 Assessment Questions

There are several questions that may be addressed for this case. The first is, "What was the spatial distribution of radioiodine deposition following the release?" The second is, "What was the concentration of radioiodine in pasture grass and alfalfa at Horn Rapids (node 878)." The third is "What was the integral concentration of radioiodine in milk in the vicinity of Horn Rapids (node 878) following the deposition?" Finally, "What were the thyroid burdens in a 4-year-old and an 8-year-old child drinking 1 gallon/day and 1 quart/day, respectively, from a single backyard cow at this location?"

3.7.2 Available Data

This particular incident was extensively studied. A public report of the results is available (Soldat 1965). Air measurements were taken at daily intervals (24-hour samples) at Benton City, Richland, and Kennewick, as well as at 18 on-site locations. Measurements of pasture grass were taken daily at two farms in node 878, and sporadically at numerous other locations. Milk was monitored at two farms in the Benton City area and at the local creameries. Thyroid counts were taken on two children who were consuming milk from a backyard cow at one of the farms.

3.7.3 Models to be Evaluated

The source term for this particular accident is known. Meteorological data are available. A special run of RATCHET for the local area will be required with daily inputs for the month of September 1963. This will be coupled with a special run of the PILOT code (equivalent to the DESCARTES code) for the node 878. (Using PILOT is equivalent to using DESCARTES, but without the multiple node dependencies.) PILOT will also be used to estimate the uptake by the two subject children.

3.7.4 Evaluations to be Performed

The RATCHET code will provide 100 realization values of the air concentration at node 878, to be compared with the monitoring data. The PILOT code will provide 100 realizations of the daily concentrations of iodine in pasture grass and individual-cow milk. These values will be compared with the daily monitoring values as distribution-to-point comparisons. PILOT will also provide estimated distributions of intake by the subject children. These will be post-processed with selected assumptions pertaining to the dose factor to arrive at body burdens, which will be compared with the actual measured body burdens of the subjects.

3.7.4.1 Assessment Question on Spatial Distribution Following the Release

The comparison of measured and modeled air values will be conducted using graphical and descriptive methods. For each day the measured values will be displayed on a map of the area. The median predicted concentration for each node in the area for that day will also be displayed on the map. Box plots that show the distribution of the 100 predicted values for each node will be prepared and a metric developed that indicates where in the box plot the measured values fall. This metric, e.g., a value from 1 to 6 (1 indicating below the 5th percentile, 2 indicating between the 5th and 25th percentiles, etc.), will then be plotted on the map at the location of each measurement. The resulting map of indices for each day will be used to look for patterns of poor prediction over the area for that day. The sequence of daily maps will also be compared to detect patterns of poor spatial prediction over time.

In addition to the daily spatial metric plots, the ratio of the median predicted value to the median measured value will be computed for each node where measured values exist. A box plot that shows the distribution of these ratios will provide a summary of the results. The proportion of the ratios that fall above 3 or below 1/3 will be determined to help evaluate attainment of DQOs. This process will be conducted for each daily plot. We note that spatial correlation among the measured values creates problems in the interpretation of statistical tests. Hence, we will rely primarily on descriptive and graphical methods to evaluate the model's performance.

3.7.4.2 Assessment Question on Pasture Grass and Alfalfa at Horn Rapids

The following description applies to both pasture grass and alfalfa. The comparison of measured and modeled values will be conducted using graphical and descriptive methods as well as more formal statistical tests. The daily Horn Rapids measurements will be plotted on a time-series plot. The plot will also display the box plot of each day's 100 realization values of daily concentrations to permit visual comparison of the measured and predicted values. This time-series plot will be used to look for patterns of model-prediction bias over time. For example, the plot will help determine if key deposition events, as indicated by the measured values, are in agreement with the predicted values.

In addition to the time-series plot, the ratio of the median predicted value to the median measured value will be computed for each day. These daily ratios will be plotted on a second time-series plot to examine visually whether the daily ratio remains within acceptable bounds, for example, a factor of three. A box plot that shows the distribution of the daily ratios will provide a summary of the results. The proportion of ratios that exceed various values will also be reported to aid in summarizing the performance of the model at the Horn Rapids node.

3.7.4.3 Assessment Question on Integral Concentration on Milk

The comparison of measured and modeled values will be conducted using graphical and descriptive methods. The 100 realization values of the integrated milk concentration over time at Horn Rapids will be compared with the integral of the measured values. This will be a comparison of a single measured point value to a box plot. Detailed statistical methods of performing the comparisons are not required--only one pair of points is to be compared. Acceptable results are represented by the predicted values matching the observed values to within a factor of three.

3.7.4.4 Assessment Question on Thyroid Burden

The comparison of measured and modeled values will be conducted using graphical and descriptive methods. The 100 realization values of the thyroid burdens for each individual will be compared with that individual's measured burden. This will compare the two measured points to their two corresponding box plots. Detailed statistical methods of performing the comparisons are not required--only two pairs of points are to be compared. Acceptable results are then represented by predicted values matching the observed values to within a factor of three.

3.8 Richland Worker Thyroid Counts, 1946-1947

Thousands of thyroid radioactivity measurements were made on Hanford workers employed in the nuclear fuel reprocessing facilities during the years 1944 through 1946. At that time, these measurements were used as a qualitative measurement of radioactive iodine in worker thyroid glands. Called "thyroid checks," the measurements were performed in the general plant environment using portable radiation detection instruments. The results of the thyroid checks were compared with a screening level for tolerable thyroid exposure to ensure that workers did not receive excessive iodine-131 exposures.

3.8.1 Assessment Question

The assessment question being addressed is "what was the average thyroid burden of iodine for adult male residents of Richland, Washington, for each month between June 1945 through August 1946?"

3.8.2 Available Thyroid Check Data

The total database of thyroid counts was prepared and reported by Ikenberry (1991). Nearly 7900 measurements between June 1945 and August 1946 are available and documented.

3.8.3 Models to be Evaluated

The source term will be prepared using the hourly data of the STRM model. Dispersion will be done with hourly inputs to RATCHET. The PILOT or DESCARTES code will be used to obtain the monthly deposition values for the Richland node (node 878). CIDER or PILOT will be used to generate distributions of intake for a reference adult male individual living in Richland. (Use of CIDER/DESCARTES or PILOT will lead to equivalent information.) This will provide indirect validation of the RM and STRM source terms, and the RATCHET dispersion model, as well.

3.8.4 Evaluations to be Performed

Input will be provided to the DESCARTES or PILOT code from the STRM/RATCHET output database. Output from DESCARTES or PILOT will be the 100 monthly realization values for food concentration in the Richland node. We will predict the thyroid burden for the average reference adult male living in Richland, working at Hanford, and eating in Hanford cafeterias. For each month, the resultant distributions of predicted average thyroid burdens will be compared with the average of measured thyroid check values for that month. This will be a distribution-to-point comparison.

The comparison of measured and modeled values will be conducted using graphical and descriptive methods as well as more formal statistical tests. Box plots will be prepared that show the distribution of the 100 predicted values for each month. The ratio of the median mean predicted value to the mean monthly measured value will be computed for each month. These monthly ratios will be plotted on a time-series plot to examine visually whether the monthly ratio remains within acceptable bounds, for example, a factor of three. A box plot that shows the distribution of the monthly ratios will provide a summary of the results. The proportion of ratios that exceeds various values will also be reported to aid in summarizing the performance of the model.

3.9 Krypton-85 Atmospheric Dispersion

Relatively recent Hanford Site data on atmospheric dispersion include a data set of coupled monthly source terms and environmental measurements of atmospheric krypton-85. These data were collected on a network established with the restart of the PUREX facility in late 1983 and through its campaign until 1988.

3.9.1 Assessment Question

The assessment question to be addressed is "what is the monthly average concentration of krypton-85 at the selected monitoring stations from 1983 through 1988?"

3.9.2 Available Krypton-85 Monitoring Data

The Hanford monitoring data consist of individual samples for 14 to 38 days duration collected at several locations within the Columbia Basin. The cryogenic monitoring network expanded over the period, so the number of locations sampled per year increases from 4 in 1984 to 14 in 1988, with up to 12 samples per year at each location (not all stations were available 12 months/year for various reasons and not all samplers were left at the same location throughout the period). A summary of the available data is presented in Table 3.1. Each point represents a "monthly" composite sample from a distinct location. The "On-Site" data will not be used.

3.9.3 Models to be Evaluated

Using the available monthly source terms for krypton-85, the RATCHET code will be used to estimate the distribution of krypton-85 concentrations in air for the relevant months between 1984 and 1988. This will be a direct validation of the transport portions of the RATCHET model; the deposition algorithms are not used for noble gases such as krypton.

Table 3.1. Summary of Krypton-85 Data Availability (Number of Samples)

Year	On-Site	Perimeter	Nearby	Distant
1983	0	6	0	5
1984	44	22	0	9
1985	39	41	0	12
1986	28	38	29	20
1987	20	36	33	23
1988	46	45	29	25

3.9.4 Evaluations to be Performed

The output realization values for each month for each location from RATCHET will be compared with the individual sampling point data. This will be a distribution-to-point comparison.

The code will be run to predict the spatial pattern of air concentrations for the specified period of time. Then values will be post-processed to provide monthly integrated values at the sampler locations for periods of sampler operations (the samples do not necessarily coincide with calendar months). These results will be compared with the measured concentrations.

The comparison of measured and modeled values will be conducted using graphical and descriptive methods. Box plots that show the distribution of the predicted values for each sampling station will be prepared and a metric developed that indicates where in the box plot the measured values fall. This metric, e.g., a value from 1 to 6 (1 indicating below the 5th percentile, 2 indicating between the 5th and 25th percentiles, etc.), will then be assigned to each sample. Then the spatial distribution of metric values will be examined for evidence of spatial transport bias and overall bias in predicted concentrations.

In addition to the spatial metric plots, the ratio of the median predicted value to the measured value will be computed for each sample location. A box plot that shows the distribution of these ratios will provide a summary of the results. The proportion of the ratios that fall above 3 or below 1/3 will be determined to help evaluate attainment of DQOs. Also, a scatter plot of the median predicted concentrations versus the measured values will be constructed to look for bias.

3.10 Columbia River Hydraulics

Modeling of transport of radionuclides released from Hanford production reactors by the Columbia River involves the use of a model that accurately reproduces the flow characteristics of the river, by which the radionuclides are diluted and in which they undergo radiological decay as they travel downstream. Inputs to the model are the water discharge of the Columbia River upstream of Hanford and discharges of the down-stream tributaries. Outputs depend on the quantity of water and on the length of time it takes to travel from one place to the next.

3.10.1 Assessment Question

The question being addressed is, "What are the discharge hydrograph and water surface elevations of the water in the Columbia River at specific down-stream locations?"

3.10.2 Available Columbia River Data

Measured water surface elevations and actual water discharge hydrographs from the U.S. Army Corps of Engineers at the various Columbia River gauging stations (e.g., Hanford Site locations, John Day, The Dalles, and other locations) are available from the Corps and from Hanford Site sources.

3.10.3 Models to be Evaluated

The hydrographs and radionuclide transport are being simulated with the CHARIMA computer model. The hydraulics of the river are independent of the radionuclide source term.

3.10.4 Evaluations to be Performed

We will compare the hourly point estimates of the water surface elevations and discharge hydrographs against the hourly point values of the measured data at several locations over a time period of several weeks, for periods with varying discharge. These will be time-series point-to-point comparisons.

For each variable and location, the measurements and model predictions will be plotted for each time increment on a time-series plot. The set of time-series plots will be used to look for patterns of model-prediction bias over time and space for the different variables. The objective is to look for systematic deviations from the predicted values over time.

In addition to the set of time-series plots, the ratios of the predicted to measured values will be computed for each time increment and location. These ratios will be plotted on a second set of time-series plots to examine visually whether the ratios remain within acceptable bounds, for example, a factor of three. Box plots that show the distribution of the daily ratios will provide a summary of the results. The proportion of ratios that exceed various values will also be reported to aid in summarizing the performance of the model.

3.11 Columbia River Water Concentrations, 1967

Concentrations of six radionuclides in Columbia River water are being estimated using the models in the HEDR toolbox.

3.11.1 Assessment Question

The assessment question being addressed is "what is the monthly average concentration of phosphorus-32, zinc-65 and chromium-51 in Columbia River water at Ringold, Richland, McNary/ Umatilla, and Bonneville in 1967?" The year 1967 is selected to be comparable with the validation of fish and ocean products discussed below in Sections 3.12 and 3.13.

3.11.2 Available Columbia River Water Concentration Data

Numerous measurements were taken of Columbia River water radionuclide concentrations during the 1960s. Monthly composite samples are available for phosphorus-32, zinc-65, and chromium-51 at several locations downstream of the Hanford release points.

3.11.3 Models to be Evaluated

Concentration of radionuclides in water depends on both the source term and transport calculations. The direct comparison will be with the CHARIMA outputs. Validation of those outputs will serve as indirect validation of the river source term release model.

3.11.4 Evaluations to be Performed

The calculation process involves the use of the 100 monthly realization values of the radionuclide release source term, transported using the CHARIMA model. This will provide 100 realization values of water concentration of each radionuclide at each location. These can be compared with the monthly composite water samples taken at the respective sampling locations. This will involve a series of distribution-to-point comparisons.

For each variable and location, the monthly measurements and model predictions will be plotted for each month on a time-series plot. The set of time-series plots will be used to look for patterns of model-prediction bias over time and space for the different variables. The objective is to look for systematic deviations from the predicted values over time.

In addition to the set of time-series plots, the ratios of the predicted-to-measured values will be computed for each variable, month, and location. These ratios will be plotted on a second set of time-series plots to examine visually whether the monthly ratios remain within acceptable bounds, for example, a factor of three. Box plots that show the distribution of the monthly ratios will provide a summary of the results. The proportion of ratios that exceeds various values will also be reported to aid in summarizing the performance of the model.

3.12 Columbia River Resident Fish Concentrations, 1967

Concentrations of five radionuclides in Columbia River fish are being estimated using bioaccumulation factors developed on the basis of ratios radionuclide concentrations in fish to those in river water. Data for the years 1960 through 1966 are being used to develop the bioaccumulation factors. The data for 1967 are being reserved for validation use.

3.12.1 Assessment Question

The assessment question being addressed is, "What is the monthly average concentration of phosphorus-32 and zinc-65 in Columbia River omnivores, first-order predators, and second-order predators at Ringold, Richland, and McNary/Umatilla in 1967?"

3.12.2 Available Columbia River Fish Concentration Data

Numerous measurements of Columbia River fish radionuclide concentrations were taken by Hanford monitoring groups. Samples are available for phosphorus-32 and zinc-65 at several locations near the Hanford release points. Few samples are available at locations below McNary Dam. As discussed by Walters, Dirkes, and Napier (1992, Sections 7.3 and 9.0), off-site agencies were interested in water and sediment, but not resident fish.

3.12.3 Models to be Evaluated

Concentration of radionuclides in fish depends on the source term and transport calculations, and on the bioaccumulation modeled in the Columbia River Dose model (CRD). The direct comparison will be with the CRD intermediate outputs. Validation of those outputs will serve as indirect validation of the river source term and transport models.

3.12.4 Evaluations to be Performed

The calculation process involves use of the single monthly realizations of the radionuclide concentrations in three types of fish prepared by the CRD model. These can be compared with the monthly composite fish samples taken at the respective sampling locations. This will involve a series of point-to-point comparisons.

For each variable and location, the monthly measurements and model predictions will be plotted for each month on a time-series plot. The set of time-series plots will be used to look for patterns of model-prediction bias over time and space for the different variables. The objective is to look for systematic deviations from the predicted values over time.

In addition to the set of time-series plots, the ratios of the predicted to measured values will be computed for each variable, month, and location. These ratios will be plotted on a second set of time-series plots to examine visually whether the monthly ratios remain within acceptable bounds, for example, a factor of three. Box plots that show the distribution of the monthly ratios will provide a summary of the results. The proportion of ratios that exceeds various values will also be reported to aid in summarizing the performance of the model.

3.13 Salmon and Oyster Concentrations, 1967

Concentrations of phosphorus-32 and zinc-65 are being prepared as annual averages for application to all locations, because the major source of the contamination in the fish is a chronic, dilute source in the Pacific Ocean. The concentrations are being based on annual cumulative source terms and the monitoring data from Pacific Ocean salmon and Willapa Bay oysters. Data are available for several years in the 1960s; most are being used to develop the functional relationships, but the 1967 data are being reserved for validation use.

3.13.1 Assessment Question

The assessment question being addressed is, "What is the 1967 annual average concentration of zinc-65 and phosphorus-32 in salmon and oysters resulting from Hanford operations?"

3.13.2 Available Salmon and Oyster Data

Relatively few data are available from Hanford-related or off-site monitoring of salmon. The few sources reported by Walters, Dirkes, and Napier (1992) have been enhanced with published and unpublished data from other outside sources. Much better data are available for oysters from Willapa Bay, Washington. These were routinely monitored from 1959 through 1977 (although contamination levels declined dramatically after reactor shutdown in 1971).

3.13.3 Models to be Evaluated

The salmon and oyster concentrations are to be estimated using relationships between cumulative annual release and monitored contamination based on the available monitoring data. Most of the data are being used to develop the functional relationships. The data for the year 1967 are being retained for validation use, because there was still a sufficient source term in that year to provide reliable measurements.

3.13.4 Evaluations to be Performed

The estimates of salmon and oyster concentrations of zinc-65 and phosphorus-32 will be compared with the average of the measurements composited over the year. This will be a point-to-point comparison.

Detailed statistical methods of performing the comparisons are not required--only two pairs of points are to be compared. Acceptable results are represented by the predicted values matching the observed values to within a factor of three.

3.14 Richland Whole-Body Counts, 1960-1970

Tens of thousands of whole-body radioactivity measurements were made on Hanford workers employed throughout the Hanford operations during the years 1959 through the present. These measurements were used as a measure of exposure to radioactive substances in the workplace. A large fraction of the Hanford workers lived in the Tri-Cities area and were routinely exposed to contaminants in Columbia River water through drinking and recreational activities. Almost all of the whole-body counts taken during the period of reactor operation indicate the presence of zinc-65, sodium-24, cesium-137 (from fallout), and naturally occurring potassium-40 (e.g., Swanberg 1962).

3.14.1 Assessment Question

The assessment question being addressed is, "What was the average body burden of sodium-24 and zinc-65 for adult male residents of Richland, Washington, for each month between January 1960 through December 1969?"

3.14.2 Available Whole-Body Count Data

The HEDR Project has acquired all of the routine, worker-related whole-body counts from 1959 through 1971. In all, the database consists of over 40,000 records.

3.14.3 Models to be Evaluated

The source term will be prepared using the monthly data of the river release model. Monthly river water and fish concentrations will be prepared. The river dose model CRD will be used to obtain the monthly intake values for the Richland location; intake for a reference adult male individual living in Richland will be used. A post-processor will be used to estimate resulting body burden. This will provide indirect validation of the source terms, and the CHARIMA transport model, as well.

3.14.4 Evaluations to be Performed

Output from the CRD model will be a single monthly realization value for body burden in the Richland location. For each month, the resultant point values of predicted body burden will be compared with the available monthly distributions of composited whole-body count values. This will be a point-to-distribution comparison.

The comparison of measured and modeled values will be conducted using graphical and descriptive methods as well as more formal statistical tests. Box plots that show the distribution of the measured values for each month will be prepared. The ratio of the predicted value to the median monthly measured value will be computed for each month. These monthly ratios will be plotted on a time-series plot to examine visually whether the monthly ratio remains within acceptable bounds, for example, a factor of three. A box plot that shows the distribution of the monthly ratios will provide a summary of the results. The proportion of ratios that exceeds various values will also be reported to aid in summarizing the performance of the model.

Formal statistical tests will also be conducted if the data are adequate. For example, we may use the paired t test on either the original data or the ranks of the data to test for differences between the predicted value and median of the measured values across months.

4.0 IAEA Vamp Project Coordination

In 1988 the International Atomic Energy Agency started a "coordinated research program" on the Validation of Models for the Transfer of Radionuclides in Terrestrial, Urban, and Aquatic Environments (IAEA 1990). Since 1989, the HEDR Project has been active in the so-called VAMP program through a no-cost research agreement (meaning no cost to the IAEA--funding for participation is provided through the HEDR budget). The VAMP program is concerned with models and transfer data relevant to transfer in terrestrial, aquatic, and urban environments. It is not concerned with models for atmospheric transport, but does consider the interaction of aerosols with terrestrial and aquatic surfaces. The principal objectives of VAMP are

- facilitate the validation of assessment models by acquiring suitable sets of environmental measurement data from the results of national research and monitoring programs, especially those established following the Chernobyl accident in 1986

- guide environmental research and monitoring efforts to acquire data for the validation of models used to assess the radiologically most significant exposure pathways

- produce reports reviewing the current status of environmental assessment modeling, including principal remaining areas of uncertainty in models used for radiation dose assessment.

The HEDR Task 02 Leader is a member of the multiple pathways working group of VAMP. The activities of this working group are described in the Progress Reports of the VAMP program (IAEA 1990, 1991, 1992). In brief, this group has established a set of "blind tests," in which the participants take basic input information and prepare assessments of environmental behavior of specific environmental contamination incidents. To date, the "CB" and "S" scenarios (for Central Bohemia and Sweden), based on Chernobyl data, have been investigated. The HEDR Phase I model was the best predictor of 3-year cumulative dose among the 18 international models participating. However, there were several instances of compensating over- and under-prediction with the HEDR Phase I model. The final documentation of the CB scenario is expected to be forthcoming shortly.

4.1 The Vamp CB Scenario

The "CB" scenario was based on measured deposition in the area of Prague, in what was then Czechoslovakia. The HEDR Phase I model was the model most appropriate for simulating the regional nature of the contamination, and it was one of the very few models that included stochastic calculations of uncertainty. It was also one of the best predictors, although detailed analysis indicated several areas for which improvement could be desired.

The VAMP procedure is to allow the participating modelers to improve their models and then demonstrate the improvement with the original data set. Because of the delays in finishing the new HEDR models, the HEDR Project has not been able to formally complete the CB demonstration. However, the CB data set, which includes air, soil, vegetation, animal product, and human concentrations of cesium-137, remains one of the best data sets for comparison purposes. It is planned to repeat the CB scenario when the DESCARTES/CIDER combination is completed.

4.2 The Vamp H Scenario

As a participant in the multiple pathways group, the HEDR Project has volunteered a Hanford-based data set and scenario description of the use of the VAMP participants. The "H" (for Hanford) scenario is described in Section 3.7 above--the 1963 PUREX release of radioiodine. A scenario description has been drafted. This will be discussed at the July 1993 meeting of the VAMP members, and then finalized for use by the multiple pathways working group.

Use of the Hanford data by up to 20 outside modeling groups will provide additional information to the HEDR Project about the nature of model uncertainty, parameter variability, and the range of results that other users would predict for the Hanford environment. International participation in the Hanford dose calculations will provide additional exposure and credibility to the HEDR Project.

5.0 Individual Validation Report Generic Contents

Each of the individual validation exercises described in Section 3.0 will result in an internal validation report. The validation report will serve as the project record of the validation activity and as an input to the Model Reliability Report scheduled for Fiscal Year 1994. The generic contents of each of these validation reports is described here. In addition, an overall project matrix will be prepared listing all of the HEDR models, the DQOs applicable to each, and the particular validation exercises that demonstrate the fulfillment of the DQOs.

5.1 Test Procedures

Prior to implementing the tests outlined in Section 3.0, a detailed test plan will be prepared. This plan will enumerate the steps required to prepare the input data and perform the evaluation. The methods for comparison to the acceptance criteria will be described. The test procedures will form the first section of the test report.

5.2 Configuration Information

This section of the validation report will define the version number and date of the computer code and input data used for the test. The computer platform and operating system will be described. Any special hardware, software, or operating procedures will be defined. If minor modifications to the code must be made to provide for output of intermediate information, descriptions of the change, and regression tests to ensure that no errors have been introduced, will be prepared. (In the computer science sense, a "regression test" is performed to show that changes have not introduced errors into the code, i.e., to show that the code has not "regressed.")

5.3 Independent Data Set

The data with which the computed results are being compared will be identified. Source of data, location of special data files, and other information will be recorded. In most instances, it is expected that the full data set will be provided to ensure its insertion into the HEDR Project records.

5.4 Application Log

A complete list of the runs performed for the test will be provided. File names, input variables, and other information sufficient to allow replication of the calculation will be provided.

5.5 Test Results

The details of the comparison of the test runs to the independent data set will be provided. This may include scatter plots, regression analyses, and other statistical interpretations of the paired data. If hand calculations are used, records will be retained in this section. Evidence of independent technical review of calculations will be retained. A conclusion documenting the achievement of the DQOs (Section 1.3), or lack thereof, will be prepared.

6.0 Application of Validation Results

Most of the code validation efforts described in this document are scheduled to occur after the initial production calculations and possibly after the publication of calculated results. It is possible that the code validation efforts will identify errors in codes or that they will illustrate areas in which additional work would provide great benefits to the project.

Staff will attempt to incorporate lessons learned from the validation tests into the HEDR computational toolbox. However, the current workplan does not include an explicit cycle to allow model and code refinements after the validation steps prior to the end of the current Battelle contract with the Centers for Disease Control. In light of this, the staff intend to provide the Technical Steering Panel with recommendations for additional work, to be carried out by Battelle or the successor contractor(s), following analysis of the results of the validation tests.

7.0 Projected Schedules and Costs

An overall schedule and cost estimate, following the HEDR work plan (Shipler 1993) and expanding on correspondence with the TSP and CDC[a] is provided.

7.1 Schedule

The schedule for overall implementation of the model validation plan is as follows.

1. Battelle establishes data analysis methods for conducting validation studies (June 30, 1993).
   - Define numerical characteristics of computed values.
   - Define characteristics of validation data set(s).
   - Define technique(s) for comparing the computed values and data sets.
   - Include detailed techniques in this plan.

2. TSP/CDC selects the acceptable level of predictive accuracy desired for calculational model. (July 15, 1993). These criteria establish the cost/benefit tradeoff on number and completeness of studies and how the results will be evaluated.

3. TSP/CDC/BNW establishes final set of validation studies to be performed (sufficient comparisons to be performed over the range of conditions to which the models will be applied). (July 15, 1993)

4. Battelle conducts validation studies. (November 30, 1993)

5. Battelle/TSP/CDC conduct peer review of study results. (January 31, 1994)
   - Battelle conducts internal reviews.
   - TSP/CDC conducts reviews.

6. Battelle documents validation results in the model reliability report. (February 28, 1994)

7. TSP/CDC evaluates results of validation studies to determine if accuracy levels have been achieved.

7.2 Cost Estimate

A preliminary time-phased spending plan is provided by HEDR Task in Table 7.1. Current funding is scheduled in Tasks 02, 04, 05, 07, and 08 to cover these expenditures.

(a) Letter, D. B. Shipler to J. E. Till and M. R. Donnelly, March 5, 1993.

Table 7.1. Time-Phased Spending for HEDR Validation Activities Through January 1994 (Hours)

Staff	June	July	Aug.	Sept	Oct.	Nov.	Dec.	Jan 94	TOTAL
Task 2	40	60	60	40	40	40		40	$27k
Task 8	40	20	40	40	40	40		40	$26k
Task 7			40	60	40	60		40	$24k
Task 5			40	40					$8k
Task 4			40	80	40				$16k
Editor							40	40	$8k
TOTAL	$6k	$6k	$21k	$26k	$16k	$14k	$4k	$16k	$109k

VAMP	FY93	$5K	FY94	$10K	TOTAL	$15K
Participation	Meeting Jul 93		Meeting Mar. 94			

8.0 References

Beck, D. M., R. F. Darwin, A. R. Erickson, and R. L. Eckert. 1992. Milk Cow Feed Intake and Milk Production and Distribution Estimates for Phase I. PNL-7227 HEDR, Pacific Northwest Laboratory, Richland, Washington.

Denham, D. H., R. L. Dirkes, R. W. Hanf, T. M. Poston, M. E. Thiede, and R. K. Woodruff. 1993. Phase I Summaries of Radionuclide Concentration Data for Vegetation, River Water, Drinking Water, and Fish. PNWD-2145 HEDR, Battelle, Pacific Northwest Laboratories, Richland, Washington.

General Electric Company. 1948. P-Department Daily Report. HAN-45954-Pts 1 through 6, November 1944 through December 1947, General Electric Company, Hanford Works, Richland, Washington.

Heeb, C. M. 1993. Iodine-131 Releases from the Hanford Site, 1944 Through 1947. PNWD-2033 HEDR Vols. 1-2, Battelle, Pacific Northwest Laboratories, Richland, Washington.

Ikenberry, T. A. 1991. Evaluation of Thyroid Radioactivity Measurement Data from Hanford Workers, 1944-1946. PNL-7254 HEDR, Pacific Northwest Laboratory, Richland, Washington.

International Atomic Energy Agency (IAEA). 1989. Evaluating the Reliability of Predictions Made Using Environmental Transfer Models. Safety Series No. 100, Vienna, Austria.

International Atomic Energy Agency (IAEA). 1990. VAMP Progress Report Number 2. Vienna, Austria.

International Atomic Energy Agency (IAEA). 1991. VAMP Progress Report Number 3. Vienna, Austria.

International Atomic Energy Agency (IAEA). 1992. VAMP Progress Report Number 4. Vienna, Austria.

Jenne, D. E., and J. W. Healy. 1950. Dissolving of Twenty Day Metal at Hanford. HW-17381-DEL, General Electric Company, Hanford Works, Richland, Washington.

Parker, H. M. 1950. Health Instrument Divisions Report for Month of December, 1949. HW-15550-E-DEL, General Electric Company, Hanford Works, Richland, Washington.

Robkin, M. A. 1992. "Experimental Release of I^{131}: The Green Run," Health Physics 62(6):487-495.

Shipler, D. B. 1993. Integrated Task Plans for the Hanford Environmental Dose Reconstruction Project, FY 1992 Through May 1994. PNWD-2020 HEDR, Battelle, Pacific Northwest Laboratories, Richland, Washington.

Singlevich, W. 1950. H.I. Environs Report for the Month of December 1949. HW-15593, General Electric Company, Hanford Works, Richland, Washington.

Snyder, S. F., W. T. Farris, B. A. Napier, T. A. Ikenberry, and R. O. Gilbert. 1992. Parameters Used in the Environmental Pathways (DESCARTES) and Radiological Dose (CIDER) Modules of the Hanford Environmental Dose Reconstruction Integrated Codes (HEDRIC) for the Air Pathway. PNWD-2023 HEDR, Battelle, Pacific Northwest Laboratories, Richland, Washington.

Soldat, J. K. 1965. "Environmental Evaluation of an Acute Release of I^{131} to the Atmosphere." Health Physics 11(10):1009-1015.

Swanberg, F., Jr. 1962. "Quantitative Measurements of Some Gamma-Ray-Emitting Radionuclides in Nuclear Industrial Workers by Whole-Body Counting Techniques." Health Physics 8(1):67-71.

Walters, W. H., R. L. Dirkes, and B. A. Napier. 1992. Literature and Data Review for the Surface-Water Pathway: Columbia River and Adjacent Coastal Areas. PNWD-2034 HEDR, Battelle, Pacific Northwest Laboratories, Richland, Washington.

Distribution

No. of Copies

OFFSITE

Technical Steering Panel

D. S. Barth
University of Nevada
4505 Maryland Parkway
Las Vegas, NV 89154

W. A. Bishop
2503 Wedgewood Court S.E.
Olympia, WA 98501

M. L. Blazek
Oregon Department of Energy
625 Marion Street N.E.
Salem, OR 97310

G. G. Caldwell
Tulsa City-County Health Dept.
4616 East 15th Street
Tulsa, OK 74112

S. N. Davis
Dept. Hydrology & Water Res.
Building 11
University of Arizona
Tucson, AZ 85721

N. J. Germond
224 Iron Mountain Blvd.
Lake Oswego, OR 97034

P. C. Klingeman
Civil Engineering Dept.
Apperson Hall 202
Oregon State University
Corvallis, OR 97331-2302

No. of Copies

K. J. Kopecky
Hutchinson Cancer Res. Center
1124 Columbia Street
Seattle, WA 98104

P. D. McGavran
Dept. of Health and Welfare
450 W. State Street, 4th Floor
Boise, ID 83720-5450

R. L. Morrill
Dept. of Geography, DP-10
University of Washington
Seattle, WA 98195

A. H. Murphy
3115 NW McKinley Drive
Corvallis, OR 97330

D. W. Price
Agricultural Economics
Hulbert Hall Room 211
Washington State University
Pullman, WA 99164-6210

M. A. Robkin
Radiological Sciences, SB-75
University of Washington
Seattle, WA 98195

G. S. Roessler
Route 1, Box 139H
Elysian, MN 56028

B. Shleien
2421 Homestead Drive
Silver Springs, MD 20902

No. of Copies

A. P. Slickpoo, Sr.
809 Nez Perce Lane
P.O. Box 311
Kamiah, ID 83536

J. E. Till
Route 2 Box 122
Neeses, SC 29107

D. E. Walker, Jr.
P.O. Box 4147
Boulder, CO 80306

OTHER

2 DOE Office of Scientific and Technical Information
Technical Information Center
P.O. Box 62
Oak Ridge, TN 37830

B. G. Brooks, EH-421
Room J-112
Department of Energy
Germantown, MD 20545

18 K. CharLee
Office of Nuclear Waste Mgmt.
Department of Ecology
719 Sleater Kinney Road, S.E., Suite 200
Olympia, WA 98504

L. E. Denham
Centers for Disease Control and Prevention
4770 Buford Highway N.E., F-35
Atlanta, GA 30341-3724

No. of Copies

M. R. Donnelly
Public Health Service
Centers for Disease Control and Prevention
2201 Sixth Avenue, RX-22
Seattle, WA 98121

H. A. Haerer
Golder Associates, Inc.
4104 148th N.E.
Redmond, WA 98052

F. O. Hoffman
SENES, Oak Ridge Inc.
Center for Risk Analysis
677 Emory Valley Road
Suite C
Oak Ridge, TN 37830

J. P. Thomas
HEAL
1720 N. Ash
Spokane, WA 99205

ONSITE

6 **DOE Richland Operations Office**

R. F. Brich, TSD	A5-55
Public Reading Room (5)	A1-65

50 **Battelle, Pacific Northwest Laboratories**

D. M. Anderson	K8-15
G. L. Black (5)	K1-25
S. D. Cannon	K1-25

No. of Copies			No. of Copies		
	D. H. Denham	K3-56		B. A. Napier (5)	K3-54
	D. E. Deonigi	K6-54		J. V. Ramsdell	K6-03
	P. W. Eslinger	K6-96		J. A. Shaw	K1-25
	W. T. Farris	K3-54		D. B. Shipler	K1-25
	M. D. Freshley	K6-77		J. C. Simpson	K7-34
	R. O. Gilbert	K7-34		S. A. Sage	K6-03
	W. A. Glass (14)	K4-13		C. D. Taylor	B1-40
	R. H. Gray	K1-33		M. E. Thiede	K6-13
	S. P. Gydesen	P8-55		W. H. Walters	K6-09
	G. L. Harvey	K1-77		Publishing Coordination	K1-06
	C. M. Heeb	K6-42		Records Center (2)	K3-70
	T. A. Ikenberry	K3-54		Technical Library (2)	P8-55

END

DATE FILMED

8/19/93