Paper SY12

SAS® Application for Human Health Risk Assessment for Hazardous Waste Combustion Facilities

Samuel E. Stinnette, Science Applications International Corporation (SAIC), Oak Ridge, Tennessee

Dennis J. Beal, Sharon K. Robers, Charles T. Hadden, Barney W. Cornaby, SAIC, Oak Ridge, Tennessee

ABSTRACT

The U.S. Environmental Protection Agency (EPA) has published guidance on performing human health and ecological risk assessments for hazardous waste combustion facilities. This guidance includes mathematical models for determining emissions from such facilities (e.g., based on the hazardous wastes being burned), as well as resulting air, soil, water, plant, and animal food chain concentrations. EPA favors a commercially available combustion risk assessment software package to perform these calculations, ensuring consistency and accuracy of results. Although flexible in nature, this software package does not allow for special site-specific scenarios such as special human receptors and risks from radiological contaminants. Science Applications International Corporation (SAIC) in Oak Ridge, TN, has developed SAS code and SAS data sets to allow for the mathematical modeling of soil, water, plant, and animal concentrations, as well as the estimation of human health risks from exposures to approximately 470 chemicals and radionuclides. The code combines database management, input from and output to Microsoft® Excel files, complex mathematical modeling, and use of SAS macros to perform the necessary calculations and produce report-ready tables. The modeling and calculation process results in a series of SAS data sets with thousands of observations and hundreds of variables to better assess realistic exposure scenarios.

INTRODUCTION

This paper provides an overview of the human health risk assessment process for hazardous waste combustion facilities. Through the use of SAS programming, the daunting challenge of quantifying risks for human exposures to emissions from a hazardous waste combustion facility is met. Mathematical models from EPA are programmed, allowing for accurate computation of thousands of environmental exposure concentrations and risks. A "system" utilizing SAS/ACCESS Interface to PC Files, SAS macros, and relatively new breakthroughs in outputting SAS data into Microsoft Excel tables (with formatting) is created, producing tables for a report that answers the question: do emissions from the hazardous waste combustion facility produce unacceptable risks relative to thresholds for potential human and ecological receptors in the vicinity of the facility? SAS version 8.2 under Windows 2000 is used for this project.

OVERVIEW OF THE RISK PROCESS FOR WASTE COMBUSTION FACILITIES

The EPA provides specific guidance on conducting human health risk assessments for hazardous waste combustion facilities: *Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities* (EPA 1998a) and *Region 6 Risk Management Addendum – Draft Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities* (EPA 1998b). EPA guidance is also provided for conducting ecological risk assessments for hazardous waste combustion facilities: *Guidelines for Ecological Risk Assessment* (EPA 1998c) and *Screening Level Ecological Risk Assessment Protocol for Hazardous Waste Combustion Facilities* (EPA 1999). These guidance documents provide detailed instructions on the risk assessment process, from determining which chemicals should be evaluated at a hazardous waste combustion facility, to modeling emissions that will exit stacks of the facility and be deposited in nearby locations, modeling environmental concentrations for other media (e.g., soil, surface water, sediment, plants, and animals), and assessing the health risks to human and ecological receptors exposed through various pathways to these media. Although there are separate guidance documents for human health and ecological risk assessment, there is an overlap (the modeling of some of the environmental concentrations). This paper will focus on the common parts, then focus on the human health risk assessment process after the determination of environmental concentrations.

The initial risk assessment is an assessment of expected emissions and expected exposures, based on engineering data and modeling. The waste stream being burned is evaluated and models applied to predict emissions from multiple stacks and flues within stacks. Site-specific meteorological data are assessed and the deposition of emissions from the stacks and flues are modeled, using EPA-approved software. The resulting emissions data and

other site-specific and contaminant-specific parameters are used to model contaminant concentrations in media other than air: soil, sediment, surface water, plants, and animals. The resulting concentrations in the various media are then used to evaluate risks to potential human and ecological receptors in the vicinity of the hazardous waste combustion facility.

The goal of the risk assessment process is to determine whether emissions from the hazardous waste combustion facility will have detrimental effects to potential human receptors. EPA seeks demonstration of low risk (i.e., below established thresholds) via a formal risk assessment, prior to issuing final permits to operate the facility.

For potential human receptors, EPA has established two types of "risks" that must be assessed:

1. Cancer risk from exposure to contamination is expressed as Incremental Lifetime Cancer Risk (ILCR), or the increased chance of cancer above the normal background rate of cancer. Thus, the ILCR is a probability, so that the larger the ILCR, the greater the chance of cancer.

2. Non-carcinogenic health hazards are characterized using a hazard quotient (HQ) and hazard index (HI) approach. The HQ is the ratio of the calculated average daily dose (ADD) to the reference or "safe" dose established by research (and approved by the EPA), for a specific pathway such as ingestion, dermal contact, or inhalation. The HI is the sum of HQs, across all pathways and all chemicals.

The EPA combustion guidance (EPA 1998a, 1998b, 1998c, 1999) provides thresholds for ILCRs, and HIs. A demonstration that these thresholds have been met must be made prior to issuance of a permit to operate the facility. The thresholds are as follows:

- The ILCR should be less than 10^{-5} (1 chance in 100,000) for all receptors.
- The HI must be less than 0.25 for all receptors.

A report is produced that summarizes the health risks to potential receptors, with the goal of demonstrating that the ILCR and HI criteria have been met for all potential receptors in the vicinity of the hazardous waste combustion facility. Applicable combinations of contaminants (chemicals and radionuclides), potential receptors, media, and locations are evaluated in the risk assessment. Thousands of calculations are performed and many tables of results are produced, so that the decision can be made regarding the ILCR and HI levels for the facility. Accuracy of results is extremely important.

For this case study, more than 470 contaminants are evaluated.

Five separate locations are evaluated:

1. On-site Ground Maximum.
2. Off-site Maximum.
3. Mountain Maximum (Native Americans only).
4. Maximum at the local river.
5. Maximum location for acute exposures.

Three type of human receptors are evaluated:

1. Industrial Workers (including the workers and nursing infants of workers);

2. Non-Native American Residents (including resident adults and children, nursing infants of residents, resident subsistence farmer adults and children, nursing infants of resident subsistence farmers, and resident subsistence fisher adults and children); and

3. Native American Residents (including Native American subsistence resident adults and children, as well as nursing infants of Native American subsistence residents).

Seven types of media are evaluated in the human health risk assessment:

1. Chronic air emissions (chronic air concentrations, evaluated for direct inhalation and external exposure to radionuclides in air);

2. Soil concentrations (evaluated for incidental ingestion, inhalation of resuspended soil, and external exposure to radionuclides in soil);

3. Plant concentrations (evaluated for ingestion of homegrown produce and wild produce);

4. Water concentrations (evaluated for ingestion of drinking water, dermal contact with steam produced in a Native American sweat lodge, and inhalation of vapors produced in a Native American sweat lodge);

5. Animal tissue concentrations (evaluated for ingestion of fish, beef, milk, pork, chicken, chicken eggs, wild fowl, wild fowl eggs, and wild game);

6. Specific combinations of the above named media and pathways to evaluate an infant's ingestion of breast milk (media and pathways are specific to the media and pathways for which the infant's mother is exposed); and

7. Acute emissions (acute air concentrations, evaluated for direct inhalation).

Figure 1 displays the combinations of human receptors, media, locations, and timeframes (current or future), as well as the plausibility (plausible or worst-case) of each exposure scenario.

Although a specific software (i.e., other than SAS) has been developed and marketed to perform these calculations and produce tables of results, the SAIC project team evaluated the software and made a decision to use SAS for the following portions of the process:

1. Modeling of contaminant concentrations in soil, plants, animal tissue, and surface water;
2. Calculating human health risks; and
3. Outputting all results into report-ready tables, using Dynamic Data Exchange (DDE).

The basis of this decision follows.

WHY USE SAS?

The commercially available software was evaluated by SAIC for potential use on this project. One general result of this evaluation was that the software, which is designed to evaluate default exposure scenarios specified in the human health combustion guidance (EPA 1998a), could perform calculations for most of the site-specific scenarios required for this project, but many of them would require multiple runs, followed by manually adding the outputs together in another software (e.g., Microsoft Excel). The commercially available software handles a limited number of chemicals only, which causes problems for this project, as this project includes nearly twice as many chemicals as found in the software, as well as radionuclides (no evaluation of these contaminants is provided in the software). Another general conclusion was that many calculations, especially for certain receptors, could not be performed within the software. These specific "problems" noted with the software are further explained below. All of these problems could be solved by using the SAS software. Thus, the decision was an easy one and was made early in the project: SAS would be used to maintain all databases and perform all calculations.

Several site-specific exposure scenarios could be evaluated using the commercially available software, but would require a series of runs, followed by a summation of the results to obtain the total exposure. For example, for the resident farmer (a default scenario evaluated with the software), the evaluation of the ingestion of beef and milk is made based on the assumption that cows eat a combination of forage grown in root-zone soil and grain grown in tilled soil. For this project the Native American gets his milk from cows in a similar manner as the resident farmer, but the Native American obtains his meat from wild game (e.g., deer) that eats only forage grown in the wild (root-zone soil). Thus a special, separate run would have to be made to obtain exposures for milk and wild game, followed by the summation of these exposures (to obtain the total exposure) for the Native American. This process is easily accomplished within one program in SAS.

A similar example involves the ingestion of fish pathway, again for residents versus Native Americans. The resident is evaluated using a single fish ingestion rate in the software. The Native American is assumed to ingest both fish and fish parts (head, fins, tails) – using different ingestion rates. The commercially available software could do these calculations, but only by performing two separate runs, followed by the addition of the results from the two runs to obtain the total ingestion of fish pathway results. Several other scenarios followed this pattern of requiring more than one run of the software, along with adding the results together to obtain the total exposure for the scenario. Again, these situations are easily handled with SAS programming.

Another problem with the commercially available software was that it was created to evaluate chemicals, but not radionuclides. This project included 370 organic chemicals, 54 inorganic chemicals, and 46 radionuclides. Modeling equations, as well as equations to quantify ILCRs, are different for radionuclides versus chemicals. This difference was a major obstacle when considering the commercially available software. Carbon-14 and tritium are two special-

Figure 1. Human Receptor Populations, Locations, Timeframes, and Exposure Pathways

Receptor	Location	Exposure Pathways: Inhalation of emissions	Ingestion of soil	Inhalation of resuspended soil	Ingestion of drinking water	External exposure to radionuclies in air	External exposure to radionuclides in soil	Ingestion of homegrown produce	Ingestion of wild produce	Ingestion of wild game	Ingestion of wildfowl and eggs	Ingestion of milk	Inhalation of vapors in sweat lodge	Dermal absorption in sweat lodge	Ingestion of fish	Ingestion of homegrown beef and milk	Ingestion of homegrown chicken and eggs	Ingestion of homegrown pork	Ingestion of breast milk
Current Plausible Exposure Scenarios																			
On-site Industrial Worker	Receptor works at On-site Ground Maximum	■	■	■	■	■	■												
	and lives at Off-site Maximum	■	■	■	■	■	■	■											
Nursing Infant of Industrial Worker	Mother works at On-site Ground Maximum and lives at Off-site Maximum																		■
Resident Adult and Child	Off-site Maximum	■	■	■	■	■	■	■											
Nursing Infant of Resident	Mother lives at Off-site Maximum																		■
Native American Subsistence Resident Adult and Child	Receptor lives at Off-site Maximum, visits Mountain Maximum, and harvests food on-site.	■	■	■	■	■	■		■	■	■		■	■	■				
Nursing Infant of Native American Subsistence Resident	Mother lives at Off-site Maximum, visits Mountain Maximum, and harvests food on-site.																		■
Current Worst-Case Exposure Scenario																			
Resident Subsistence Farmer Adult and Child	Off-site Maximum	■	■	■	■	■	■	■				■				■	■	■	
Nursing Infant of Resident Subsistence Farmer	Off-site Maximum																		■
Resident Subsistence Fisher Adult and Child	Off-site Maximum	■	■	■	■	■	■	■							■				
Acute Exposure	Acute Maximum	■																	
Future Plausible Exposure Scenarios																			
On-site Industrial Worker	Receptor works at On-site Ground Maximum	■	■	■	■	■	■												
	and lives at Off-site Maximum	■	■	■	■	■	■	■											
Nursing Infant of On-site Industrial Worker	Mother works at On-site Ground Maximum and lives at Off-site Maximum																		■
Future Worst-Case Exposure Scenarios																			
Resident Adult and Child	On-site Ground Maximum	■	■	■	■	■	■	■											
Nursing Infant of Resident	Mother lives at On-site Ground Maximum																		■
Resident Subsistence Farmer Adult and Child	On-site Ground Maximum	■	■	■	■	■	■	■				■				■	■	■	
Nursing Infant of Resident Subsistence Farmer	Mother Lives at On-site Ground Maximum																		■
Resident Subsistence Fisher Adult and Child	On-site Ground Maximum	■	■	■	■	■	■	■							■				
Native American Subsistence Resident Adult and Child	Receptor lives at On-site Ground Maximum, visits Mountain Maximum, and harvests food on-site.	■	■	■	■	■	■		■	■	■		■	■	■				
Nursing Infant of Native American Subsistence Resident	Mother lives at On-site Ground Maximum, visits Mountain Maximum, and harvests food on-site.																		■

■ = Exposure pathway is evaluated in the risk assessment.

case radionuclides that have their own special exposure models (different from the other 44 radionuclides). These special cases could not be handled within the commercially available software, but are easily programmable in SAS.

Also, the commercially available software only provided chemical-specific information for approximately 200 chemicals. The software did allow for adding chemicals to the database, so a lot of research and keystrokes would be required, either within the commercially available software or within the software chosen for this project (since 470 contaminants would be evaluated). SAS can handle all of these requirements.

The nursing infant scenario is evaluated for dioxins only in the commercially available software, whereas for this project, other chemicals (co-planar polychlorinated biphenyls) and radionuclides need to be addressed in this exposure scenario. Further, the commercially available software does not address the new EPA guidance for calculating lifetime risks resulting from infant exposures. SAS can handle all of these cases.

Another problem with the commercially available software was that it did not include equations/calculations to evaluate the Native American sweat lodge exposures, including inhalation and dermal exposure. The commercially available software could have possibly obtained results for the inhalation pathway, but definitely not for the dermal exposure pathway, which was required by regulators for this project. Again, this scenario is easily handled via SAS programming.

Research indicated that for some contaminants, uptake factors used in the modeling of various environmental concentrations result in more contaminant accumulated than is physically possible. A mass-limited uptake factor approach is used to avoid overestimating concentrations (and ultimately ILCRs and HIs). The use of mass-limited uptake factors is not present in the commercially available software, but is programmable in SAS.

In summary, based on the many special cases described above, a decision was made early in the process to create all databases and perform all calculations using the SAS software.

MODEL DESCRIPTION

The SAS model used for this project involved four major tasks:

1. Creating databases.
2. Modeling concentrations for the risk assessment.
3. Calculating risks and hazards for human receptors.
4. Producing output for reports.

A flow chart of this process is shown in **Figure 2**, with each of these four tasks described below. This paper focuses on the modeling of environmental concentrations, the calculation of human health risks and hazards, and producing report-ready output from the risk assessment. The simplified description of the process involves three sets of calculations and a series of programs that present tables in report-ready format. The three sets of calculations are

1. combining emissions data output from the air dispersion models with modeling parameters to estimate environmental concentrations;

2. combining modeled environmental concentrations with risk assessment parameters to determine human health intakes; and

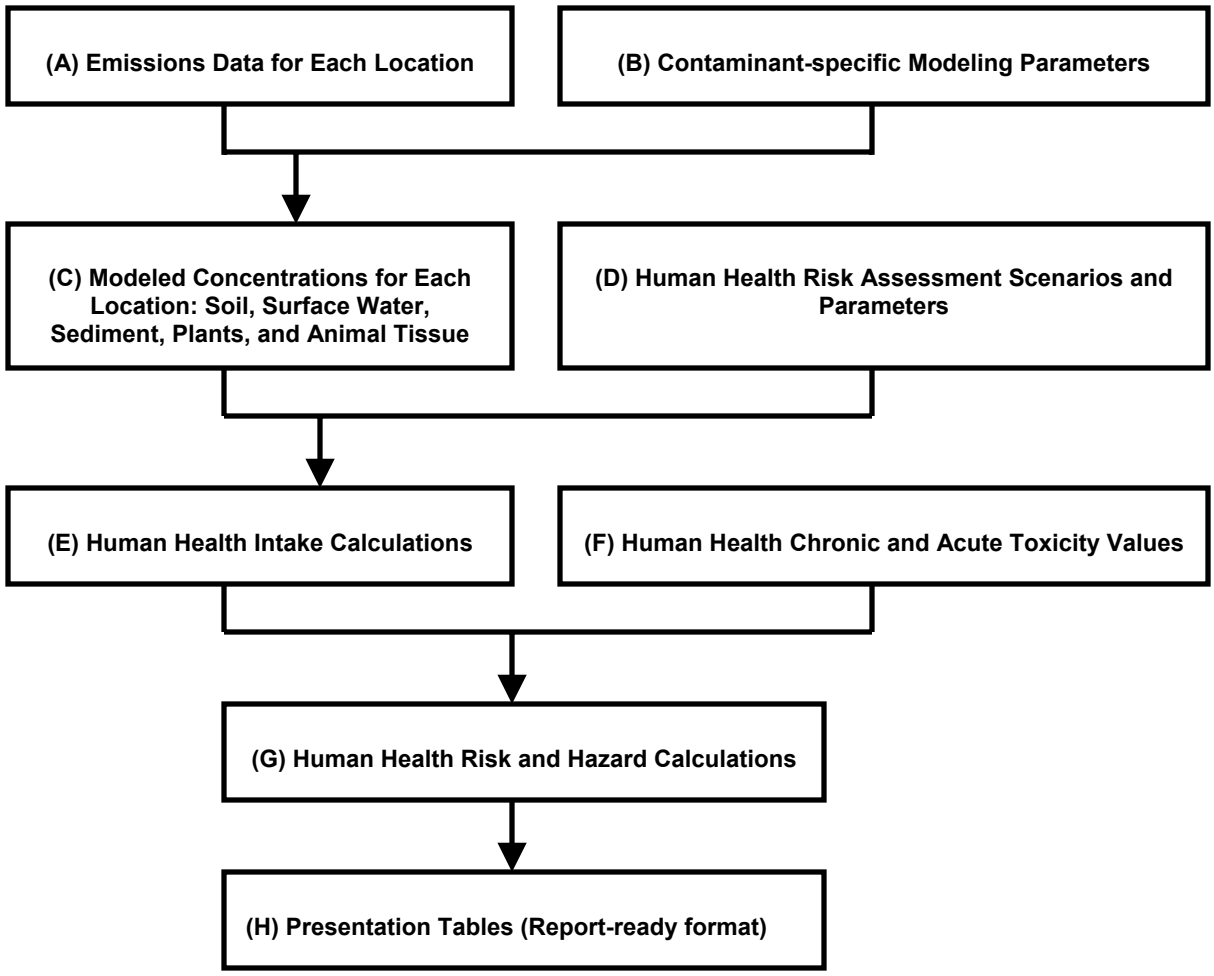
3. combining human health intakes with human health toxicity values to estimate human health risks and hazards.

CREATING DATABASES

The development of the databases is the most time-consuming step of this risk assessment process, requiring lots of research and hand-entry of parameter values. Different types of contaminant-specific parameters are needed to model environmental concentrations and to quantify risks (ILCRs) and hazards (HIs) for the human health risk assessment. Other parameters needed are either site-specific or default values from the EPA, provided in the risk assessment guidance (EPA 1998a, 1998b, 1998c, and 1999). These various parameters are represented within boxes B, D, and F in **Figure 2**.

Initially, separate data sets are created for modeling parameters (e.g., the air-to-plant uptake factor), human health risk assessment scenarios and parameters (e.g., the body weight of each receptor), and human health toxicity values (e.g., the oral cancer slope factor). EPA guidance provides contaminant-specific parameter values for approximately 220 chemicals. However, this project includes the evaluation of 470 contaminants. Extensive research is required to find contaminant-specific parameter values for the remaining 250 contaminants. This process of researching and developing the database, especially for the modeling parameters (box B in **Figure 2**), is very time-consuming and is

Figure 2. Human Health Risk Assessment Flow Chart



on-going. Some modeling parameters are derived from other parameters (e.g., for organic chemicals, transfer factors can be derived from the log of the octanol-water partitioning coefficient). The approach for this project is to store parameter values in Microsoft Excel spreadsheets, followed by reading these data into SAS data sets, using the CIMPORT procedure.

After all data sets of parameter values are stabilized, the data sets are merged into one "master" data set of parameter values. This master data set is used to model environmental concentrations (box C), calculate human health intakes (box E), and calculate human health risks and hazards (box G). This master data set contains 470 observations (one for each contaminant evaluated in the risk assessment) and 360 variables. Thus, there is a potential for $470 \times 360 = 169{,}200$ values in this master data set.

Based on new research and the continuing effort to complete parameter values for all 470 contaminants, updates are periodically made to the master data set. Most often, these updates affect only a few contaminants. These updates are made within SAS programs, usually through IF-THEN statements. If the update affects a large number of contaminants, the new parameter values are read into a temporary SAS data set (e.g., from a Microsoft Excel spreadsheet, using PROC CIMPORT) and subsequently merged with the master data set to effectively overwrite the previous values for the affected parameters. Thus, the process used to make updates to the master data set is a professional judgment call as to which is more efficient (i.e., least time-consuming): using a SAS program with IF-THEN statements to update values or reading in data from an external file followed by merging this data onto the master data set.

MODELING CONCENTRATIONS FOR THE RISK ASSESSMENT

The process of modeling environmental concentrations for the risk assessment (box C) does not require sophisticated techniques in SAS, but is the most challenging step of the SAS system developed for this project. The

calculations of environmental concentrations, described below, are performed within a series of SAS programs. Each of these programs creates output SAS data sets by merging the master data set with another input data set. The modeling equations from the combustion guidance are programmed, using IF-THEN statements to address specific scenarios. The end result is a data set with modeled (i.e., calculated) concentrations for each contaminant, following the EPA combustion guidance.

The process of modeling environmental concentrations for several media (e.g., soil, surface water, plants) could be contained within a single SAS program for each location (e.g., the On-site Ground Maximum). However, the equations are quite different from one medium to another. Furthermore, some equations are complex and some specific scenarios require detailed attention. Therefore, a series of SAS programs is developed for each location and medium combination (e.g., soil concentrations at the Mountain location).

Based on the modeling equations from the EPA combustion guidance, the SAS programs to model concentrations for the risk assessment must be executed in a specific order. For a specific location (e.g., On-site Ground Maximum), the following programs are executed:

1. **Air modeling data** (including emission rates and air concentrations) are read from a Microsoft Excel spreadsheet into a SAS data set via PROC CIMPORT. The air modelers populate the spreadsheet with data in a specified format, allowing the data to be read into a SAS data set smoothly. Two noteworthy specific scenarios requiring detailed attention are: (a) separate emissions and air concentrations are provided for eight separate flues, within three separate stacks coming from the hazardous waste combustion facility; and (b) each contaminant is categorized and modeled as being in one of three phases: particle, particle-bound, or vapor.

2. Two SAS programs are created and executed to estimate **soil concentrations**, one program for current concentrations and the other program for future concentrations, based on the emissions data (box A) and modeling parameters (stored on the master data set; see box B) as inputs. EPA equations for soil modeling (EPA 1998a) are programmed. There are several versions of the soil equations, depending on several factors, including, but not limited to, whether the contaminant is carcinogenic or non-carcinogenic, if the soil-loss constant is zero or non-zero, and the phase (particle, particle-bound, or vapor) of the contaminant. Another unique scenario requiring detailed attention is that special models are required to determine soil concentrations for mercury (these mercury models are described in the EPA combustion guidance). Dealing with the emissions and air concentrations from the multiple flues and stacks into a soil concentration (for each contaminant) at a single location requires complex models and programming. The resulting output is a SAS data set for the location being evaluated. The data set contains soil concentrations at three depth intervals: untilled (1-cm depth), root-zone (15-cm depth), and tilled (20-cm depth). The output data set from executing a soil-modeling program contains approximately 470 observations and 540 variables (i.e., there is a potential for 253,800 values on this data set), depending on the specific location being modeled.

3. The modeled soil concentrations data set serves as the input data set for the determination of **plant concentrations**. Several types of plant concentrations are modeled, including but not limited to, aboveground produce based on air-to-plant transfer, aboveground produce from direct deposition, aboveground produce from root uptake, and belowground produce from root uptake. Some or all of these four types of plant concentrations are modeled for homegrown produce and wild plants eaten by humans, as well as forage, silage, and grain grown as animal feed. The program has several complexities, including the determination of plant concentrations for certain combinations of depths and subsequent receptors. For example, plants are grown in either root-zone soil (e.g., wild plants, consumed by Native Americans) or tilled soil (e.g., homegrown vegetables), but never in untilled soil. Another example of a complexity in the programming is that forage and silage are only estimated for aboveground categories, with forage from root uptake calculated from root-zone depths only and silage from root uptake calculated from tilled soil only; grain is only estimated for the aboveground/root-uptake category, from both tilled and root-zone depths. Again, having emissions from multiple flues and stacks, as well as the phase (particle, particle-bound, or vapor) and the potential for using mass-limited uptake factors come into play for modeling plant concentrations, thus requiring unique models and programming. Carbon-14 and tritium have their own models that are rather different from all other contaminants. All project-specific calculations requiring detailed attention are programmed within the data step. The output data set from executing a plant-modeling program contains approximately 470 observations and 680 variables (i.e., there is a potential for almost 320,000 values on this data set).

4. The modeled plant concentrations data set serves as the input data set for the determination of **animal tissue concentrations** for human consumption. In this program the animal tissue concentrations estimated for each contaminant at this particular location include beef, milk, pork, chicken, chicken eggs, wild fowl, wild fowl eggs, and wild game (e.g., deer). The equations used include a food intake component for each animal being modeled. For example, a beef cow (as opposed to a dairy cow) ingests specific amounts of soil (from 1-cm depth), as well

as forage, silage, and grain (all grown at the 20-cm depth). The previously calculated soil, forage, silage, and grain concentrations, along with soil-to-beef uptake factors, are used to estimate the beef concentration. Similar approaches are used to estimate concentrations for milk, pork, chicken, chicken eggs, wild fowl, wild fowl eggs, and wild game. There are several project-specific scenarios requiring detailed attention, including dealing with the phase of the contaminant (particle, particle-bound, or vapor) and the potential for using mass-limited uptake factors are abundant. The output data set from executing an animal tissue-modeling program contains approximately 470 observations and 730 variables (i.e., there is a potential for more than 343,000 values on this data set).

5. The air modeling data set serves as the input data set for the determination of **surface water, sediment, and fish concentrations**. These concentrations are modeled for the river maximum location only. Surface water modeling is the most complex, with many equations containing parameters that are functions of other parameters that, in turn, are functions of yet other parameters. For example, one of the parameters used to estimate the water concentration is the total contaminant load to the water body. The total contaminant load to the water body is a function of several other parameters, including deposition and runoff, which are calculated from several other parameters. Although the equations themselves are not complex, the order in which the calculations are performed is essential. To obtain the final surface water and sediment concentration for a contaminant, approximately 35 equations are programmed, in a specific order. Water-to-fish and sediment-to-fish uptake factors are used to model the fish concentrations from the water or sediment (depending on the type of contaminant). Again, there are several project-specific scenarios requiring detailed attention, including dealing with multiple flues and stacks and the phase of the contaminant (particle, particle-bound, or vapor). The output data set from executing a program to calculate surface water, sediment, and fish concentrations contains approximately 470 observations and 930 variables (i.e., there is a potential for more than 437,000 values on this data set).

Once all calculations of concentrations have been made, the estimation of risks and hazards to specific human receptors can be performed. The SAS data sets containing modeled concentrations (box C) are used to estimate human health intakes (box E) and eventually risks and hazards (box G).

CALCULATING RISKS AND HAZARDS FOR HUMAN RECEPTORS

Estimating risks and hazards for human receptors is a two-step process. First, intakes are calculated by combining exposure concentrations and exposure parameters; and second, the resulting intakes are combined with toxicity values to determine risks and hazards. For chemicals, intakes are quantified as the lifetime average daily dose (LADD) and average daily dose (ADD) in units of mg/kg/day. The LADD defines a dose level that is distributed (averaged) over an entire lifetime. Unlike the LADD, the ADD is averaged over a specific incremental exposure period rather than an entire lifetime. For radionuclides, intakes are quantified as a total intake in units of picocuries (pCi). Equations and parameter values that are used to quantify intakes are taken from EPA guidance (EPA 1998a). The general intake equation is of the form:

$$Intake = EPC \times Parameters,$$

where

Intake = Chemical or radionuclide intake (mg/kg/day or pCi),
EPC = Exposure point concentration (modeled; media-specific: mg/m^3, mg/kg, or mg/L for chemicals and pCi/m^3, pCi/g, or pCi/L for radionuclides), and
Parameters = Various exposure parameters (receptor- and/or contaminant-specific).

The general risk and hazard equations are

$$Risk = Carcinogenic\ intake \times CSF$$

and

$$HQ = Non\text{-}carcinogenic\ intake \div RfD,$$

where

Risk = Cancer risk (unitless),
Carcinogenic intake = Carcinogenic intake (mg/kg/day for chemicals and pCi for radionuclides),
CSF = Cancer slope factor [$(mg/kg/day)^{-1}$ for chemicals and pCi^{-1} for radionuclides],
HQ = Non-carcinogenic hazard quotient (unitless), for chemicals only,

Non-carcinogenic intake = Non-carcinogenic intake (mg/kg/day), for chemicals only, and
RfD = Reference dose (mg/kg/day), for chemicals only.

On **Figure 2**, this two-step process involves combining boxes C (EPCs) and D (risk scenarios/parameters) to obtain box E (intakes), then combining box E (intakes) with box F (toxicity values) to obtain box G (risks and hazards). This process is accomplished through a series of SAS programs, each program executed for a specific combination of locations and media. For example, a program called calc_risk_air_ground_max_current.sas calculates chronic intakes, risks, and hazards for all receptors exposed via air-related pathways at the On-site Ground Maximum location, using current concentrations.

For this project, 26 different SAS programs are used to quantify intakes, risks, and hazards (as noted above, one of these programs is called calc_risk_air_ground_max_current.sas). Each program performs 4 primary steps: (1) merging the data set containing EPCs and toxicity values with a data set containing the human health risk scenarios and parameters (this data set is described below); (2) making a copy of each set of EPCs for each location/medium/pathway combination; (3) executing a do-loop to calculate intakes, risks, and hazards for all appropriate location/medium/pathway combinations; and (4) summing risks (hazards) across all pathways for each receptor/location to obtain a total risk (hazard) for each contaminant, as well as a total risk (hazard) across all contaminants for each receptor/location. Twenty-one specific SAS macros are written to perform the calculations in steps 3 and 4. Specific macros are called within the 26 different SAS programs; these macros are described below. The resulting total risks and hazards are then compared against the thresholds (the total risk should be less than 10^{-5} and the total hazard should be less than 0.25).

The human health scenarios and parameters are created and stored in a separate SAS data set (box D). This data set contains all unique exposure scenarios evaluated in the human health risk assessment, including all combinations of locations (e.g., On-site Ground Maximum), receptors (e.g., Native American subsistence resident adult), media (e.g., soil), and pathways (e.g., incidental ingestion of soil). Each observation on this data set contains a variable called MACRO, whose value points to a SAS macro that contains the actual equations for intake, risks, and hazards.

When the macro is executed, a temporary SAS data set (with the same name as the SAS macro) is created. The macro is geared toward the calculation of intakes, risks, and hazards for a specific pathway/medium combination (e.g., ingestion of soil) and, therefore, can be used for multiple location/receptor combinations. Each macro combines boxes C and D from **Figure 2** to estimates intakes (box E) and then combines the intakes with toxicity values (box F) to quantify risks and hazards (box G). Note that the toxicity values are contaminant-specific and are already in the same data set as the modeled concentrations (the toxicity values are part of the original "master" data set of contaminant-specific values; see discussion above under DATABASES). A total of 21 distinct SAS macros are used in the human health risk assessment for this project. An example of one of the 21 macros is provided below.

The example macro is called INGSATS (indicating ingestion of soil). This macro estimates carcinogenic intakes (SAS variable GSCCDI), non-carcinogenic intakes (SAS variable GSNCDI), risks (SAS variable GSRISK), and hazards (SAS variable GSHQ). The exposure concentration in soil is assigned to the variable called CS prior to the execution of the INGSATS macro. The macro INGSATS contains the following code:

```
   %macro INGSATS;
   data ingsats; * ingsats -> INGestion/Soil/Adult (single receptor using chronic
RfD)/using toxicity equivalency factors (TEFs) and Summations of risks/hazards;
   set &INDS; * INDS = name of the input data set with soil concentration (CS);
   * Note: Averaging Times (AT_C and AT_N) are in years (not days);

   * (1) Chemical CDI/Risk/HQ calculations:;

   if anatype ne 'Radionuclides' then do;
   * Adjust concentrations for those with toxicity equivalency factors (TEFs):;
   if tef ne . and cs ne . then cs_tef=cs*tef;

   * (a) gsccdi = ingestion of soil/carcinogenic/chronic daily intake:;
   if cs ne . and tef =  . then gsccdi=(cs*cr_soil*fi*ef*ed)/(bw*at_c*365);
   if cs ne . and tef ne . then gsccdi=(cs_tef*cr_soil*fi*ef*ed)/(bw*at_c*365);

   * (b) gsrisk = ingestion of soil/carcinogenic/risk:;
   if gsccdi ne . and sfo ne . and tef =  . then gsrisk=sfo*gsccdi;
   if gsccdi ne . and              tef ne . then do;
```

```
   if tefgroup='PAH'  then sfo_tef=&SFO_PAH ; * SFO_PAH =Macro variable containing
value of oral cancer slope factor (SFO) for benzo(a)pyrene;
   if tefgroup='TCDD' or tefgroup='PCB' then sfo_tef=&SFO_TCDD; * SFO_TCDD=Macro
variable containing value of SFO for 2,3,7,8-TCDD;
   if sfo_tef ne . then gsrisk=gsccdi*sfo_tef;
   end;

   * (c) gsncdi = ingestion of soil/non-carcinogenic/chronic daily intake:;
   if cs ne . and tef =  . then gsncdi=(cs*cr_soil*fi*ef*ed)/(bw*at_n*365);
   if cs ne . and tef ne . then gsncdi=(cs_tef*cr_soil*fi*ef*ed)/(bw*at_n*365);

   * (d) gshq = ingestion of soil/non-carcinogenic/hazard quotient:;
   if gsncdi ne . and rfdoc ne . and tef=. then gshq=gsncdi/rfdoc;
   if gsncdi ne . and tef ne . then do;
   if tefgroup='PAH' then rfdoctef=&RFDOCPAH; * RFDOCPAH=Macro variable containing
value of oral chronic reference dose (RFDOC) for benzo(a)pyrene;
   if tefgroup='TCDD' or tefgroup='PCB' then rfdoctef=&RFDOCTCD; * RFDOCTCD=Macro
variable containing value of RFDOC for 2,3,7,8-TCDD;
   if rfdoctef ne . then gshq=gsncdi/rfdoctef;
   end;
   end;

   * (2) Radionuclide CDI/Risk/HQ calculations:;
   if anatype='Radionuclides' then do;
   * (a) gsccdi = ingestion of soil/carcinogenic/dose (like CDI):;
   if cs ne . then gsccdi=cs*cr_soil*fi*ef*ed*1000; * 1000=conversion factor (g/kg);

   * (b) gsrisk = ingestion of soil/carcinogenic/risk:;
   if gsccdi ne . and sfo_soil ne . then gsrisk=gsccdi*sfo_soil;
   end;
   label cs_tef  ='TEF-applied Soil Conc.'
         sfo_tef ='Oral S.F. used with TEF'
         rfdoctef='Oral Chronic RfD used with TEF'
         gsccdi  ='Ingestion/Soil Carcinogenic CDI'
         gsrisk  ='Ingestion/Soil Carcinogenic Risk'
         gsncdi  ='Ingestion/Soil Non-carcinogenic CDI'
         gshq    ='Ingestion/Soil Non-carcinogenic HQ' ;
   format cs_tef sfo_tef rfdoctef e9. gsccdi gsrisk gsncdi gshq e8.;
   run;

   * Run TOTALPW1 macro to sum across all contaminants and determine % contribution:;

   %TOTALPW1(ingsats,gsrisk,gshq,totgscr,totgsch,totgsrr,Total Chemical Risk for
Ingest./Soil,Total Chemical Hazard for Ingest./Soil,Total Rad Risk for Ingest./Soil,
ingsrisk,Ingestion/Soil Risk,pctingsr,pctingsh,Percent of Total Risk for
Ingest./Soil,Percent of Total Hazard for Ingest./Soil,coc_ings,COC for
Ingestion/Soil);
   run;
   %mend INGSATS;
```

Execution of all 26 SAS programs (which include the calling and execution of all 21 macros) results in risks and hazards calculated for all human receptors exposed to all appropriate media and all appropriate locations (see Figure 1 for the specific combinations of receptors, media, and locations). The output data sets have a wide range for the number of observations and variables. Some of these data sets are as small as 470 observations and 110 variables, while some are as large as 10,900 observations and 810 variables. These output data sets are used to create tables for a report.

PRODUCING OUTPUT FOR REPORTS

Several SAS programs are executed to summarize results and produce report-ready tables. Because the desired appearance of the Microsoft Excel tables is known prior to making the calculations, the process of producing tables is streamlined using SAS programming. Dennis Beal, one of the authors of this paper, is responsible for writing the SAS code that exported the results into the Microsoft Excel tables, including complex formatting, using the DDE tool in SAS (see Mr. Beal's paper entitled *Using Dynamic Data Exchange to Customize Formatted Reports in Microsoft*

Excel, presented in the Data Presentation section of this conference). This programming allows for consistency between tables and ensures accuracy in the presentation of results, avoiding copy-and-paste errors, as well as formatting inconsistencies; it also saves time from formatting the many tables (within Microsoft Excel) required in reporting results.

CONCLUSION

SAS is used to create databases, model concentrations for the various environmental media evaluated in the risk assessment, calculate risks and hazards for human receptors, and produce report-ready risk prediction tables for a hazardous waste combustion facility. There are several benefits from using SAS as opposed to the commercially available software, including:

- Completeness of results (all of the many project-specific scenarios are programmable in SAS; some scenarios are not obtainable using the other software).
- Using SAS macros ensures consistency/accuracy of exposure, risk, and hazard results.
- Report-ready, formatted tables are produced in a timely, efficient manner.

REFERENCES

EPA (U.S. Environmental Protection Agency). 1998a. *Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities*, Peer Review Draft, EPA/530/D-98/001B, U.S. Environmental Protection Agency, Washington, D.C., available at <http://www.epa.gov/epaoswer/hazwaste/combust/risk.htm>.

EPA. 1998b. *Region 6 Risk Management Addendum – Draft Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities*, EPA-R6-98-002, U.S. Environmental Protection Agency, Washington, D.C.

EPA. 1998c. *Guidelines for Ecological Risk Assessment*. EPA/630/R-95/002F. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, D.C.

EPA. 1999. *Screening Level Ecological Risk Assessment Protocol for Hazardous Waste Combustion Facilities,* Peer Review Draft, EPA 530-D-99-001A, Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, D.C. Available at <http://www.epa.gov/epaoswer/hazwaste/combust/ecorisk.htm>.

ACKNOWLEDGEMENTS

The author acknowledges the support and contributions of Samantha R. Pack of Science Applications International Corporation in Oak Ridge, Tennessee.

CONTACT INFORMATION

Samuel E. Stinnette
Science Applications International Corporation
151 Lafayette Drive
Oak Ridge, TN 37830
Work Phone: 865-481-8796
Fax: 865-481-8714
E-mail: samuel.e.stinnette@saic.com