AUG 7 2000

SANDIA REPORT

SAND2000-1764/2

Unlimited Release

Printed July 2000

Cassini Spacecraft Uncertainty Analysis Data and Methodology Review and Update

Volume 2: A Technical Description of the Sampling Methods Employed in the Cassini Uncertainty Analysis

Gregory D. Wyss

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of
Energy under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

RECEIVED
AUG 22 2000
OSTI

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States
Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: <http://www.doe.gov/bridge>

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: <http://www.ntis.gov/ordering.htm>

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

SAND2000-1764/2
Unlimited Release
Printed July 2000

Cassini Spacecraft Uncertainty Analysis Data and Methodology Review and Update

Volume 2: A Technical Description of the Sampling Methods Employed in the Cassini Uncertainty Analysis

Gregory D. Wyss
Risk Assessment and Systems Modeling Department
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-0747

RECEIVED
AUG 22 2000
OSTI

Abstract

During 1996 and 1997, Sandia National Laboratories worked in partnership with Lockheed Martin Missiles and Space to assist in the risk analysis and risk uncertainty assessment for the Cassini space probe. After the October 1997 launch, the Department of Energy requested additional documentation, in addition to the review and updating of several aspects of Sandia's work on the original Cassini analysis. This document provides a detailed mathematical description for the unique sampling method that was used in the Cassini risk assessment. This method incorporated aspects of stratified sampling, importance sampling, and Latin hypercube sampling in order to enable the Cassini project team to perform a consequence uncertainty analysis that did not need to homogenize data points and yet captured the tails of the resultant risk distribution using a minimum number of random observations. The report also summarizes the comparison of results computed using this method and those computed using pure Monte Carlo sampling. The results computed using this method were equivalent to those produced by pure Monte Carlo sampling. Furthermore, this method improved resolution of the tail of the risk distribution by more than an order of magnitude while using 78% fewer samples than the pure Monte Carlo analysis.

Acknowledgment

The author wishes to thank several people who have had an important part in the development and execution of the Cassini spacecraft risk uncertainty study. Nelson Deane, Chuong Ha, Frank Kampas, Steve Loughin, Jim Braun, and Larry DeFillipo, all of whom are current or former employees of Lockheed Martin, have made important contributions to this work. They have been an absolute pleasure to work with during all phases of the Cassini project.

The author also wishes to thank Allen Camp, Fred Harper, Vince Dandini, and Dennis Anderson of Sandia National Laboratories for their many consultations and technical contributions to this work. This report owes much to their care and diligence.

The author also wishes to thank Don Owings and Lyle Rutger of the U.S. Department of Energy for their management of this program. Finally, the author would be remiss if he did not also acknowledge the work of Henry Firstenberg and Bart Bartrum of Tetra Tech NUS, whose comments on the Cassini Final Safety Analysis Report provided the impetus for many of the clarifications presented in this report.

Contents

Preface.....	5
Nomenclature.....	6
1 Introduction.....	7
2 Overview of the Cassini Analysis.....	9
2.1 Problem Description	9
2.2 Risk Results to be Computed.....	10
2.3 Comparison With Nuclear Reactor Risk Assessment.....	10
2.4 Variability and Uncertainty.....	11
2.5 Computational Tools.....	12
2.5.1 The LASEP-T Software.....	13
2.5.2 The SPARRC Software.....	14
2.6 Uncertainty Assessment Process.....	14
2.6.1 Direct Substitution Method.....	16
2.6.2 Replica Event Tree Method	17
2.6.3 Mathematical Deconvolution.....	18
2.6.4 Uncertainty Analysis Implementation Issues.....	21
2.7 Summary	21
3 Mathematical Background.....	23
3.1 Monte Carlo Sampling.....	23
3.2 Latin Hypercube Sampling Theory.....	24
3.2.1 Sampling	24
3.2.2 Pairing.....	25
3.2.3 Application to the Cassini Risk Assessment.....	27
3.3 Importance Sampling.....	27
3.4 Stratified Sampling	29
3.5 Summary	30
4 Methods Used for the Cassini Study.....	31
4.1 The Cassini Total Risk Estimate.....	31
4.2 The Cassini Exceedance Frequency Curves	39
4.3 Summary	41
5 Grouping of Release Scenarios.....	43
5.1 Reasons for Grouping	43
5.2 Grouping Rationale.....	43
5.3 Characteristics of Release Scenario Clusters.....	44

6	Grouping of Weather Data.....	45
6.1	Reasons for Grouping	45
6.2	Grouping Rationale.....	45
6.3	Characteristics of Weather Clusters.....	46
7	Implementation of Sampling Strategy	47
7.1	Overview.....	47
7.2	Variability Runs.....	48
7.3	Uncertainty Runs	48
7.4	Comparison With Pure Monte Carlo Sampling.....	49
7.5	Summary	49
8	Summary and Conclusions	51
9	References.....	53

Preface

On October 15, 1997, the Cassini spacecraft was launched from the Kennedy Space Center in Florida to begin its almost 7-year journey to Saturn. Since the spacecraft carried radioactive materials, the final launch decision rested with the Executive Office of the President of the United States. In support of that decision, Lockheed Martin Missiles and Space prepared a final safety analysis report (FSAR) that described the Cassini mission, the spacecraft, and an assessment of the associated public risks. This risk assessment study examined not only the magnitudes of the risks themselves, but also the uncertainties associated with those risk estimates. Both the risk assessment itself and the uncertainty assessment made use of a unique random sampling scheme that incorporates aspects of stratified sampling, importance sampling, and Latin hypercube sampling for selecting the release scenarios, weather data, and uncertain or variable quantity values that should be forwarded for consequence computation.

During the course of a later review of the FSAR, it became apparent that the FSAR and its supporting documentation did not contain sufficient detail regarding this unique sampling method to enable all reviewers to have complete confidence in its correctness. For this reason, the U.S. Department of Energy requested that additional supporting documentation be developed that explained the mathematical basis for this sampling method as well as its use in the Cassini risk analysis. This volume provides that documentation.

This document is part of a larger set of documents that was requested by the Department of Energy from Sandia National Laboratories under a Cassini follow-on project. Specifically, task 2 under that statement of work indicates that several methodologies and data sources are to be reviewed and updated, with the results documented in a report. This document is Volume 2 of the two-volume report that was generated under that task.

Nomenclature

CCDF	complementary cumulative distribution function
CDF	cumulative distribution function
CIR	coincident impact release
DCF	dose conversion factor
DOE	U.S. Department of Energy
EPA	U.S. Environmental Protection Agency
FSAR	final safety analysis report
GPHS	general purpose heat source
ICRP	International Commission on Radiation Protection
INSRP	Interagency Nuclear Safety Review Panel
JPL	Jet Propulsion Laboratory
LASEP-T	Software developed by Lockheed Martin to investigate the progression of potential launch accidents for Titan launch vehicles.
LHS	Sandia National Laboratories' software for performing Monte Carlo and Latin hypercube sampling of statistical distributions
NASA	National Aeronautics and Space Administration
PDF	probability density function
PRA	probabilistic risk assessment
RTG	radioisotope thermoelectric generator
SETAC	Sandia National Laboratories' software for evaluation of large event tree models
SPARRC	A suite of software for the computation of radiological consequences developed by Lockheed Martin and used in the Cassini risk analysis

A Technical Description of the Sampling Methods Employed in the Cassini Uncertainty Analysis

1 Introduction

The Cassini spacecraft is a deep space probe whose mission is to explore the planet Saturn and its moons. Following launch in late 1997, the spacecraft will perform fly-by gravity assist maneuvers with the planets Venus (twice), Earth, and Jupiter on its way to a rendezvous with Saturn in 2004. Because of the extremely low intensity of solar energy at such great distances from the sun, the Cassini spacecraft derives its electrical power from three radioisotope thermoelectric generator (RTG) units that are powered by heat derived from the decay of plutonium-238. As with all space missions involving radioactive materials, the Cassini launch required approval from the Executive Office of the President of the United States.

The space program has a long history of performing risk computations for space missions using probabilistic risk assessment (PRA) techniques. This is especially true for those missions where RTGs are used. However, past assessments have typically yielded quantitative point estimates of risk and qualitative or heuristic assessments of risk uncertainty. The Cassini risk analysis team has determined that advances in the field of risk uncertainty assessment over the past decade warrant a quantitative approach to uncertainty for this mission. The Lockheed Martin Co. and the U.S. Department of Energy (DOE) contracted with Sandia National Laboratories for assistance in applying state-of-the-art quantitative uncertainty assessment methods to the Cassini risk analysis. This report describes the methods developed through this collaboration. It provides an overview of the entire Cassini risk uncertainty assessment process and a detailed explanation of one phase of that assessment: the sampling process. The Cassini risk uncertainty assessment made use of a unique combination of stratified sampling, importance sampling, and Latin hypercube sampling. A description of the derivation and justification of this unique sampling scheme is the purpose of this report.

This document focuses on how the methods developed for the Cassini risk analysis were applied to several accident cases that can be characterized as on-pad/early launch accidents without coincident impact release (CIR). Suborbital/orbital reentry cases were evaluated using a very similar method, although the effects of weather were handled in a different and simpler manner because the consequence models for such releases are less geography-specific, owing to the world-wide nature of such computations. On-pad/early launch accidents with CIR were modeled

using a slightly different formalism that was developed by Lockheed Martin independent of Sandia collaboration. Therefore, the methods used to model CIR cases are beyond the scope of this paper.

The report is organized as follows: Section 2 provides an overview of the Cassini risk analysis, including the expected results and analytical tools used. Section 3 provides mathematical background for the sampling methods that were employed in this uncertainty analysis. Section 4 then describes the actual uncertainty analysis method as implemented for the Cassini study. Sections 5 and 6 describe the need for and methods used in grouping source terms and weather data, respectively. Section 7 describes how the grouping and sampling methods work together to support the uncertainty analysis process. Finally, Sections 8 and 9 provide a summary and references.

2 Overview of the Cassini Analysis

This section describes the constraints under which the Cassini risk assessment operated. These include constraints imposed by the mission itself, the types of risk results desired, the computational tools available to analyze various physical phenomena, and the expectations of review groups.

2.1 Problem Description

The Cassini spacecraft was launched from the Kennedy Space Center at Cape Canaveral Air Station, Florida, on October 15, 1998, aboard a Titan IV launch vehicle with a Centaur upper stage. The spacecraft carried approximately 32.7 kg of plutonium-238 dioxide encapsulated in 54 general purpose heat source (GPHS) modules. While the Cape Canaveral site is relatively remote, there is a human population within several miles of the launch site. The objective of the risk analysis was to quantify the risk to humans in the event of either an accident prior to earth escape or an inadvertent reentry during the earth gravity assist maneuver planned for almost 3 years into the flight. Determining the exposure of humans to plutonium that might be released during such an accident requires one to model

- The characteristics of the accident itself and the thermal and mechanical environments to which the RTGs would be exposed,
- The performance of the RTGs, GPHS modules, and the plutonium fuel itself in these accident environments,
- The dispersal of the plutonium fuel, in the event that its containment is breached by an accident environment, from release to its eventual deposition in a location where it might come into contact with humans, and
- The biological effect that would be induced by this human exposure.

While Lockheed Martin has developed excellent tools to model the various phenomena required for a comprehensive risk assessment, these tools were not designed to be used in a full-scope integrated uncertainty analysis. Thus, the focus of the Sandia task to develop an uncertainty assessment methodology was to design a methodology that would enable the determination of quantitative uncertainty information from analyses performed using the existing computational modeling tools. The existing tools are, in general, computationally intensive, so limiting the number of times each tool must be executed was an important constraint on the development process. The final methodology, then, provided a framework for determining which specific computations would be required of the analysis tools to obtain useful uncertainty information about the then-proposed Cassini mission.

2.2 Risk Results to be Computed

The principal figures of merit, as reported in the Cassini GPHS-RTG Final Safety Analysis Report,¹ are the risk of health effects and the risk of land contamination that exceeds levels established by the U.S. Environmental Protection Agency (EPA). Each of these results was derived by aggregating results over a large number of individual scenarios. For this study, risk was defined as the aggregation of the likelihood of a particular scenario multiplied by the consequences that would be produced by that scenario. Thus, in order to support the desired risk results, one must determine the likelihood, land contamination, and health effects of each scenario analyzed. This information is also expressed as an exceedance frequency graph in which the abscissa is consequence and the ordinate is the likelihood that the value of the consequence will exceed the abscissa value.

There are obviously a multitude of possible accident scenarios that must be considered in the evaluation of risk. It is not generally feasible to perform a detailed analysis of every possible scenario because of limited computational resources. In practice, an analyst generally performs detailed assessments for a finite number of representative accident scenarios. The analyst then performs a qualitative examination for each of the remaining unanalyzed scenarios, and assigns its likelihood of occurrence to the analyzed scenario to which it is most similar. This process of using a single scenario as a surrogate for many actual physical scenarios is common. It produces representative risk results only to the extent that there is good similarity between each physical scenario and some analyzed scenario. If too few scenarios are analyzed, or the analyzed scenarios do not represent the full space of possible scenarios, then the validity of the results suffers. The uncertainty analysis methodology must therefore prescribe scenarios to be analyzed so that the full potential accident space is explored. Since the most severe accident scenarios are often exceedingly rare, selecting only (or even predominantly) the most likely accident scenarios, as might occur in a pure Monte Carlo assessment, would fail this criterion. This is why the Cassini project looked to more advanced uncertainty assessment methods.

2.3 Comparison With Nuclear Reactor Risk Assessment

There are several parallels between the risk analysis method used for launch accidents and that used for the evaluation of risks at commercial nuclear power plants. The first step is an assessment of the probability that the launch vehicle itself will experience a failure that can either jeopardize the mission or threaten the space probe. This is similar to the "Level I" core damage sequence analysis in a nuclear power PRA in that the assessment focuses on the frequency with which the system is placed into a condition where there is a *potential* for the release of radioactive materials. The second step is an assessment of the conditional probability that the launch vehicle failure will *actually* lead to a radiological release and, if a release does occur, the characteristics of the release. This is similar to the "Level II" accident progression and source term analysis in a nuclear power PRA, except that there are dramatic differences between the important release characteristics for a land-based power plant and those for a system that is in flight. The third step is a consequence analysis for the radiological release, and is similar to a "Level III" PRA for a terrestrial nuclear system.

While there are conceptual similarities between the Cassini methods and those used for terrestrial systems, the actual implementations of the methods are quite different, and these differences pose critical problems for an uncertainty analysis. The differences are derived from two major requirements that have been imposed on the uncertainty analysis method: (1) the method must support the separation of natural or random variability from state of knowledge (or "epistemic") uncertainty; and (2) the accident progression, source term, and consequence modeling must be accomplished using the existing LASEP-T and SPARRC software. Therefore, the description of the uncertainty analysis methods that we have developed for the Cassini mission must be preceded by discussions of these two constraints.

2.4 Variability and Uncertainty

One constraint that was placed upon the development of the uncertainty analysis methodology was that it must support the separation of random variability from epistemic uncertainty. The terms "random variability" and "epistemic uncertainty" are closely related in the minds of many, and the distinction between them is best drawn by example. Consider the launch of a spacecraft. Given a particular class of vehicle failure, we know that an accident can progress along any of a number of possible pathways owing to variations in unobserved, unobservable, uncontrolled, or uncontrollable parameters. The path can also be influenced by the inherent stochastic nature of certain physical phenomena. We refer to these potential pathways as "random variability." A traditional risk analysis that seeks to determine all of the possible outcomes from a particular initiating event is seeking to model random variability. This helps decision makers to understand the set of potential situations that could occur given the launch of the spacecraft.

However, there are certain observable, predictable, and controllable parameters and phenomena that may be important in predicting risk for which we simply do not have adequate knowledge. We may have limited experimental data, inadequate model information, or any of a host of other issues that lead us to be uncertain about our predictions of risk. In this study we refer to these limitations as "epistemic uncertainty." It is important that the major sources of epistemic uncertainty be identified and considered in a risk analysis.

To first order, we can think of each variability analysis as representing the universe of possible outcomes *given* a particular "view of the world," and the uncertainty analysis as forcing the analyst to consider alternative world views. If we are to differentiate between epistemic uncertainty and random variability in this fashion, we must then classify individual features of the risk model (*e.g.*, input parameters, submodel results) in this regard. It has been argued by many that there is no such thing as a feature that is either "purely" uncertain or "purely" variable because, in reality, all features contain some aspects of both random variability and epistemic uncertainty. While this is most likely true, the art of trying to determine just how much of the variation of a feature is due to random variability versus epistemic uncertainty is still the subject of intense academic and practical debate. Since such a separation is beyond the practice of current state-of-the-art risk assessment studies, we were satisfied for the Cassini study to differentiate between random variability and epistemic uncertainty on a feature-by-feature basis.

The process for assessing the random variability or epistemic uncertainty of a particular model feature begins with the analysts examining the available data for that feature and developing a statistical distribution to represent the range of values over which it might occur. This feature

must now be classified as either random variability or epistemic uncertainty. If, in the opinion of the analysis team, the distribution for this feature is dominated by random variability, the entire variation in this feature will be classified as random variability (even though some portion of it may be due to epistemic uncertainty). If, however, the analysis team believes that the variation in the feature is predominantly epistemic uncertainty, the entire variation in this feature will be classified as epistemic uncertainty in spite of the fact that some portion of it may be due to random variability. We believe that this approach allows us to grasp the impact of both epistemic uncertainty and random variability on the risk results without having to go significantly beyond the currently accepted state of the art.

An additional aspect of epistemic uncertainty analysis is known as "model uncertainty." Model uncertainty derives from the fact that one may have several models to describe a particular phenomenon or physical principle. While each model has its own pedigree, validation, and perceived region of validity, models are (generally) known to be imperfect representations of reality. In addition, there may be several models that claim to be valid in a particular region and yet produce different results. Obviously, if one had a *perfect* model, then one could dismiss all imperfect models and use only the "perfect" results. This, however, is not the case, so the analyst is left not knowing which model (or combination of models) best represents the unique conditions at hand. Because this situation derives from a lack of knowledge about the models *and* the system being modeled, it is classified as uncertainty — more specifically, model uncertainty. While model uncertainty is both real and important, its evaluation requires actually having more than one model in which one has some confidence. That was not the situation in the Cassini risk assessment, so model uncertainty was not considered in this study.

2.5 Computational Tools

A second constraint that was placed upon the development of the uncertainty analysis methodology was that it must support the use of the existing LASEP-T and SPARRC software for the risk uncertainty assessment. Since the nature of these codes played an important part in the selection of uncertainty analysis methodologies, it is appropriate at this point to briefly describe each of the codes as well as the basic method that is used to perform a point estimate risk computation for the Cassini mission.

A simple stepwise description of a launch-phase risk analysis could be stated as follows: the National Aeronautics and Space Administration (NASA) Lewis Research Laboratory provides a set of accident scenarios and environments in the flight vehicle data book.² Each scenario consists of a technical description of how the launch vehicle fails and a statistical distribution to represent the conditional probability of this failure scenario, given an attempted launch event. The safety analysis team then applies the LASEP-T software to determine the conditional probability that each scenario will result in a release of radioactive material to the environment, and the characteristics of the release should it occur (altitude, mass, particle size distribution, etc.). The SPARRC consequence software is then applied to determine the consequences (especially health effects) that would be expected to occur in the affected population as a result of such a radiological release. The flight vehicle data book is taken as a given input and is not subject to revision or reanalysis during the risk evaluation.

2.5.1 The LASEP-T Software

Lockheed Martin developed the LASEP-T computer software for the Cassini program. It is a descendent of the LASEP software that has been used to analyze vehicle breakup and RTG safety in previous space missions.³ LASEP-T performs computations that are equivalent to both the accident progression and source term computations for a traditional terrestrial nuclear power plant risk assessment. The program accepts as input a description of the accident conditions from the flight vehicle data book and the mission flight profile. This includes such information as the range of times during the flight profile during which a particular accident scenario is possible, the component(s) in the launch vehicle that fail, and a statistical description of the characteristics of the environment that the payload might experience during such an accident.

LASEP-T uses this input data to perform repeated Monte Carlo-based simulations of potential accident environments and fragment fields to determine whether the RTGs in the payload are threatened or breached during the accident and, if so, the mass of fuel that is released, its location, and its particle size distribution. It also tracks any airborne RTG material to its impact point to determine if a fuel release occurs on the ground and, if so, its mass and particle size distribution. The Fireball code⁴ transforms the LASEP-T release results to incorporate any changes to the release caused by the fireball environment that may be predicted to occur during the release scenario.

Each LASEP-T Monte Carlo trial represents an accident scenario that may lead to various combinations of air and ground-based fuel releases and fireball conditions. LASEP-T classifies each scenario that leads to a release according to a set of categories that can be thought of as the end states of a small "accident progression event tree." Once the release category has been determined, LASEP-T records the characteristics of each type of release separately (air, air/fireball, ground, ground/fireball) for mass, particle size distribution, and other important characteristics. These results can then be examined either on a trial-by-trial basis or as aggregate distributions as necessary for the consequence analysis. The conditional probability of each end state is tracked based on the number of Monte Carlo samples that "hit" that end state divided by the total number of samples run during the simulation.

The computational method in LASEP-T can be thought of as equivalent to solving an event tree many times under the constraint that each probability in the tree is either 0 or 1. Monte Carlo sampling is used to determine which branches will have 0s versus 1s in each trial. This method works well for obtaining point estimate probability results (given enough Monte Carlo trials) *if* the number of trials is large.* It is also adequate for examining random variability (without considering epistemic uncertainty) because the sampling of distributions for variable model features is a natural part of the code's normal function. However, performing a Monte Carlo or

* There is a second reason that LASEP-T requires a very large number of trials: the code performs a pure Monte Carlo simulation to characterize a very large space of possible accident scenarios. Those scenarios that produce the greatest release of radionuclides are of great interest to the risk assessment, and yet they are the most rare. Thus, in order to ensure that these low-probability, high-consequence events are not neglected by the risk assessment, the analyst must execute a very large number of LASEP-T trials.

Latin hypercube uncertainty analysis on such a model is computationally intractable because it would require us to place a Monte Carlo driver on a Monte Carlo code.

2.5.2 The SPARRC Software

SPARRC is a system of codes that is used to investigate Space Accident Radiological Release and Consequences. The SPARRC system consists of three separate codes to evaluate the consequences of different classes of release scenarios. They are (1) SATRAP (Site-Specific Analysis of Transport and Dispersion of Radioactive Particles) for launch site accident scenarios, (2) GEOTRAP (Global Transport and Dispersion of Radioactive Particulates) for reentry accident scenarios with particles, and (3) HIAD (High Altitude Aerosol Dispersion) for reentry accident scenarios with fine particle aerosols. The three transport codes share common dose calculation modules (referred to as PARDOS), output formats, and other features so that their results have a consistent meaning and can be quantitatively combined. SATRAP is the primary tool for Phase I (launch abort) accidents within a 200-km square surrounding the Kennedy Space Center launch facility.

SATRAP implements a Lagrangian-trajectory, Gaussian puff model with the capability of handling multiparticle-sized source terms. The transport and diffusion of material in the puff are governed by meteorological data (supplied by the user) that can vary in space and time. These data include wind components at grid points, stability, height of mixing layer, and a roughness length for the surface below. Each source cloud can have its own independent set of characteristics, such as particle size, cloud dimensions, and initial coordinates. Each cloud is independently tracked in time steps through a four-dimensional (x,y,z,t) wind field.

When the source cloud reaches a level where interaction with the earth's surface occurs, SATRAP calculates air and ground concentrations at grid points on the surface. SATRAP then computes the doses and health effects to exposed populations based on user-supplied population density, land use, and food production and consumption patterns. Dose conversion factors (DCF) for the different dose pathways are computed based on methods established by the International Commission on Radiation Protection (ICRP) under ICRP-30.⁵

The SPARRC code system is deterministic in nature. For the Cassini analysis, source term input (mass of fuel released in each particle size category, altitude, etc.) is selected based on the results produced by LASEP-T. An individual SPARRC run also requires as input the specific weather conditions and other non-source term information that is to be used for this particular computation of consequences. Note that some of these features are subject to random variation and hence have distributions that must be sampled. Since SPARRC does not perform statistical sampling on its own, its input distributions are sampled by the LHS Latin hypercube sampling computer software.^{6,7} The LHS results are used to set up input for the many individual SPARRC analyses necessary to account for random variability.

2.6 Uncertainty Assessment Process

In a perfect world, one would consider random accident variability by constructing statistical distributions for the model features that are considered variable (both in LASEP-T and in

SPARRC), and sample these input variables in a Monte Carlo fashion over many LASEP-T trials (within a single code run) and many SPARRC runs (using the LHS code) to account for random variability. Distributions would also be generated for model features that are considered uncertain, and these input parameters would be sampled for an epistemic uncertainty evaluation. The variability analysis process would then be repeated for each observation in the epistemic uncertainty analysis in order to produce a complete picture of random variability for each Monte Carlo observation in the epistemic uncertainty analysis. This process, while intellectually satisfying, would likely take years of computational time, even on the fastest computers available today because it requires the separate application of a Monte Carlo uncertainty process to an underlying Monte Carlo modeling code.

The objective of the Cassini risk assessment was to come as close to the ideal solution as is feasible within the fiscal and time constraints of the project and within the practical constraints of the current state of the art. We approached the problem in two phases: a detailed variability analysis, followed by a combined variability-uncertainty analysis. We began by performing a detailed variability analysis that approximates the computational methodology described above. Since the variability analysis already involves Monte Carlo sampling, it is evident that we cannot successfully "surround" the variability analysis with a Monte Carlo shell to perform an ideal uncertainty analysis. For this reason, we developed three different approaches for using the results of LASEP-T and SPARRC computations in an overall variability and uncertainty analysis methodology: (1) a direct substitution method, where the amount of epistemic uncertainty data that can be propagated through the analysis is severely limited, (2) a "replica event tree" approach in which an event tree is constructed to mimic the results of the Monte Carlo simulation code, and (3) a statistical deconvolution process in which epistemic uncertainty is deliberately intermixed with random variability during the analysis, and the two are mathematically separated after the fact. In practice, only the third method (deconvolution) was used for the Cassini analysis. The first two methods were developed largely as fallback positions ("assured success paths") because the deconvolution method was viewed as largely experimental since it had not previously been applied to risk assessment problems of this scope.

While the methods each have their own distinct probabilistic interpretations, they can all be based on and derived from a single set of LASEP-T and SPARRC runs (provided that those runs are properly designed). This was a critical component of our uncertainty analysis methodology because it enabled analysts to proceed with the confidence that the literally weeks of computational effort that they were putting into the analyses required for the deconvolution analysis would not be wasted if that method were to prove unsuccessful. Anything less would mean opening the project to potentially unacceptable financial burdens and schedule delays.

The following sections present an overview of each method and explain the relative advantages and disadvantages of each. These sections are not intended to be rigorous mathematical treatises, but rather to provide familiarity with the basic characteristics of each method so that the reader may better understand the reasons for selecting particular sampling schemes to implement these uncertainty analysis methods.

2.6.1 Direct Substitution Method

The direct substitution method, which can be thought of as a "sample-by-sample method," involves performing an uncertainty analysis without making any direct use of the detailed variability analysis results. The objective of this method is to mimic the methodology that has been developed and proved for commercial nuclear power reactor analysis. This method provided an assured success path because it had already been demonstrated and would produce predictable results even if neither of the two more advanced uncertainty analysis methods (described below) were to prove satisfactory. While this method was not used for the Cassini uncertainty analysis, it is briefly described here for the sake of completeness.

The direct substitution uncertainty analysis method parallels the traditional nuclear power method as follows: recall that each LASEP-T run computes the conditional probability for the realization of a series of potential release categories, which can be thought of as the end states of a small event tree model. These end states can be compared to the end states of an "accident progression event tree" (formerly called a "containment event tree") from a terrestrial nuclear power plant risk assessment study (e.g., NUREG-1150^{8,9}) because, for both methods, each end state represents one description of what can happen to the system to allow radionuclides to be released to the environment. In terrestrial nuclear power plant risk assessment studies, each group of similar accident progression end states is associated with a source term and a set of consequences (early fatalities, latent cancers, etc., computed to account for the random variability in weather at the time of the accident). Thus, the complete set of accident progression end state definitions can be viewed as an approximation of the random variability inherent in the accident progression (although the analysts who constructed those models have not usually made any real differentiation between random variability and epistemic uncertainty in the sense that we are attempting here).

In a terrestrial reactor Monte Carlo uncertainty analysis, the sampling of the event tree provides many complete sets of accident progression end states which, as a group, can be viewed as a representation of the epistemic uncertainty (with the random variability contained in the list of end states for each Monte Carlo observation). Each end state or group of similar end states is associated with a different source term for each observation (Monte Carlo trial) in the uncertainty analysis. This analogy, while not completely valid for the Cassini study, provides a convenient way to look at the direct substitution uncertainty methodology.

The direct substitution uncertainty analysis method is intended to mimic the terrestrial nuclear power reactor method described previously to the extent that it is possible to do so using the computational tools that are available. In order to make it computationally feasible to perform a stratified Monte Carlo uncertainty analysis for the Cassini mission, we are required to reduce the level of detail in each variability analysis. Whereas in the detailed variability analysis we viewed every LASEP-T *trial* that resulted in release as a separate contributor to random variability, we now simplify our consideration of random variability and treat each LASEP-T *end state* as being similar to a terrestrial reactor accident progression end state (the characteristics of the individual trials that lead to that end state are viewed as representing epistemic uncertainty even though they are based on a sampling of both random variability and uncertainty). In the detailed variability analysis, our Monte Carlo engines were permitted to sample only those model features

that were categorized as being dominated by random variability. In the epistemic uncertainty study, all model features for which distributions have been developed are sampled concurrently.

We can use the characteristics of the individual trials that lead to an end state as the basis for distributions that represent, in this formulation, the epistemic uncertainty in the characteristics that describe that end state. The major difference between this method and the terrestrial reactor method is that a terrestrial reactor event tree will typically produce dozens or even hundreds of end state groups per observation while only a few such end states would be produced for the Cassini assessment.

The epistemic uncertainty methodology for the SPARRC side of the analysis is very similar to that used in the detailed variability study, although a simplification of the variability analysis results for the sake of the uncertainty analysis is also required here. In the variability analysis, the results of each SPARRC run were viewed as separate contributors to random variability. Under this uncertainty method, we draw a number of samples using the LHS code, sampling both uncertain and variable model features concurrently. Each sample contains the information necessary to perform a *series* of SPARRC runs based on variations in weather and source terms. The results generated by all of the SPARRC runs for a single observation are now viewed as representing random variability, while the set of all observations taken as a group represent epistemic uncertainty. One risk exceedance frequency curve is generated for each observation to represent random variability. When this process is performed for all observations, it produces a family of curves that, when taken together, represent the epistemic risk uncertainty. These individual curves can then be summarized to produce a mean curve and various appropriate uncertainty quantile curves. The individual end state results can also be aggregated over all scenarios to present a similar family of curves that represent our epistemic uncertainty in the overall risk for the mission.

We understood that this approach would necessarily intermingle random variability with the epistemic uncertainty results both within LASEP-T and SPARRC, but this is consistent with practice in current state-of-the-art terrestrial reactor risk studies. We accepted this fact and the other stated drawbacks as known limitations of this uncertainty method, and viewed them as a price that must be paid in order to establish an assured success path for this analysis.

2.6.2 Replica Event Tree Method

While the direct substitution method did provide a guaranteed success path, it suffered from some serious drawbacks. The drastic simplification of the treatment of random variability that it requires has been mentioned previously. Another problem, however, derives from the fundamental structure of the LASEP-T code. Recall that LASEP-T determines the conditional probability for each end state based on the typical Monte Carlo formulation of a number of hits divided by the total number of trials. While the simplified treatment of random variability allowed us to gain some insights about epistemic uncertainty in the consequence arena, it could do nothing to illuminate the uncertainties in this key conditional probability.

It would be very helpful if one could find a method that would allow us to consider uncertainties in the frequency domain (such as those described above) while at the same time staying as close as possible to the guaranteed success path described in the previous section. This would provide

an evolutionary (instead of a revolutionary) approach to this risk assessment problem. For this reason, we developed a method that uses all of the same computations described in the previous section for the epistemic consequence uncertainty analysis. The only changes are related to the computation of probabilities.

In the direct substitution method, LASEP-T categorized releases according to the end states of a small event tree. The probability of each end state represents the conditional probability that a particular accident scenario (as defined by the flight vehicle data book) will result in a release of radiological material with characteristics that meet the definition of that end state. LASEP-T of necessity generates only a point estimate of each probability. Thus, in that method, the only epistemic uncertainty in the frequency domain comes from the distribution for each accident scenario's initiating event frequency (as found in the flight vehicle data book).

A second method can be used to gain some insights into the uncertainties in the frequency domain. In this method one would construct a replica of the small event tree used by LASEP-T to categorize its results. Point estimate values for the branch probabilities in this event tree model would be mathematically derived from the LASEP-T results. Sensitivity analyses performed using the LASEP-T code would then be used to provide information for experts to estimate statistical distributions for each branch probability in the event tree (obviously, the expert distributions must be consistent with the point estimate data computed by LASEP-T). Given the model and the branch probability distributions, the event tree would then be solved and subjected to a Monte Carlo uncertainty assessment using software such as Sandia's SETAC/EVNTRE code suite.¹⁰ The result would be estimates for the uncertainty of each of the end state conditional probabilities. These uncertain conditional probabilities would then be convolved with the uncertain accident likelihoods to find the epistemic uncertainty in the overall likelihood of fuel release.

The estimation of branch fraction probability distributions is admittedly an inexact science, but it is conceptually similar to the estimation of release fraction conditional probabilities that was performed for the Interagency Nuclear Safety Review Panel (INSRP) Ulysses uncertainty analysis.¹¹ While this method propagates an estimate of the uncertainty in the frequency domain, it does not provide any better opportunity to understand the dominant contributors to epistemic uncertainty than does the direct substitution method since all of the potential contributors to epistemic uncertainty must be factored into the conditional probability distributions constructed by the experts. Thus, this method offers limited additional insights at the expense of requiring additional model construction and probability distribution development. For these reasons, this method was viewed as having only limited utility for the Cassini risk analysis. Because the deconvolution method was ultimately successful for the actual Cassini study, the replica event tree method was never used.

2.6.3 Mathematical Deconvolution

The third method for characterizing uncertainty represents a fundamental change from previously proven methods. This method, as proposed and developed by Lockheed Martin Astro Space,^{1,12,13} is based on a concept from linear systems theory that is commonly applied in electrical signal analysis: the use of Laplace or Fourier transforms to deconvolve an output distribution function into its original components. In theory, the deconvolution uncertainty

analysis method will allow for a more complete separation between random variability and epistemic uncertainty without forcing the drastic simplifications that were required under the direct substitution method. While this method was viewed as largely experimental (because it had not previously been applied to risk assessment problems of this scope), it ultimately proved successful and was used to provide the final uncertainty analysis results for the Cassini FSAR.

The fundamental idea for the deconvolution uncertainty analysis method is as follows: if we can construct one distribution that represents the system risk considering only random variability, and a second distribution that represents the system risk with epistemic uncertainty and random variability fully intermingled, then by application of Laplace or Fourier transforms we should be able to reconstruct a distribution that represents only the effect of epistemic uncertainty on the variability distribution to obtain the combined uncertainty-variability distribution. The following discussion is intended to provide an overview of deconvolution methods. The reader is referred to a detailed mathematical text for a more rigorous treatment of the mathematical deconvolution.

Let us *assume* that there exists a distribution function which, when convolved with the detailed variability distribution, produces as its result the combined uncertainty-variability distribution. If this distribution function could be found, it would represent the "pure" effect of epistemic uncertainty on the variability distribution. Mathematically this could be written as

$$R = U * V$$

where R is the risk distribution from the combined uncertainty-variability analysis and is found by the convolution of an as-yet undetermined pure epistemic uncertainty distribution U with the detailed variability distribution V . Recall that for both the Laplace and Fourier transforms, the convolution operation becomes a multiplication operation. Thus, if we hope to find the unknown distribution U through the deconvolution of R and V (both of which are known), we must divide their Laplace or Fourier transforms as

$$\bar{U} = \frac{\bar{R}}{\bar{V}}$$

where the bar denotes that the relation holds in the transform domain. The remaining task, then, is to invert the Laplace or Fourier transform for $\bar{U}$ back into the real domain U .

Once the inversion is accomplished, we can plot the overall system risk as a family of curves that is similar to that described for the direct substitution method. The result would be a family of parallel curves that would be generated by a numerical convolution of U and V . In this case, each curve has the basic shape specified by the detailed variability distribution (V), but is transformed based on the action of the uncertainty distribution (U) to obtain the complete range of possible variability distribution curves.

The principle of mathematical deconvolution is, however, not without its limitations. Two limitations are most frequently cited. First, since deconvolution is based on *linear* systems theory, its rigorous application requires that the systems to which it is applied have linear transfer functions. This is most certainly not true in the Cassini risk uncertainty assessment. It should be noted, however, that the method is routinely applied to nonlinear systems in the area of electric

circuit theory with reasonable success. Note also that the objective of the deconvolution analysis is simply to provide an estimate for the degree of confidence that one can place in the risk analysis results (the "spread" of uncertainty about a best-estimate risk profile), and that the process of deconvolution has no effect whatsoever on the best-estimate risk profile itself. Therefore, while a skeptic of deconvolution may hold the uncertainty analysis results in question, his confidence in the best-estimate risk profile should be unaffected because those results are based on long-established practice in the risk and reliability analysis field.

The second limitation of the deconvolution method is that the "level of confidence" curves that it generates are necessarily parallel to the best-estimate risk profile. In other words, it is incapable of reproducing a situation where different views of the world (as expressed through epistemic uncertainty) converge with, diverge from, or even cross the original best estimate risk profile. While this limitation is real, it is believed to be of little consequence for the Cassini study. This limitation would be more important in a case where one is comparing risk results for several different models of the same phenomenon, or contrasting the effects of multiple hazard curves on a system.* In these cases, nonparallelism of the risk curves may in fact be important, and the use of a deconvolution method inappropriate. However, the Cassini study varies neither the hazard curve nor the actual details of the phenomenological modeling (save for the variation of parameter values implemented in a single model). Thus, it is believed that this limitation of the deconvolution method is of only minor importance for the Cassini analysis.

The deconvolution uncertainty analysis method clearly pushes beyond the current state of the art in PRA methodology, and there are aspects of the method that remain to be fully demonstrated. For example, mathematical rigor would require that all of the transfer functions in our study (LASEP-T, SPARRC) be linear in order for the underlying mathematical theory to be applicable. This is clearly not the case for our system. However, initial applications of the method to nonlinear transfer functions have shown promise. Successful application of this method would produce a distribution U that represents the separate effect of epistemic uncertainty on the system, and in doing so, would overcome some of the limitations described for the direct substitution method. This method can make use of the same LASEP-T and SPARRC computations described for the direct substitution method.† The only additional computations would be the generation, manipulation, and inversion of the transform variables. These functions

* A hazard curve describes the likelihood that a threat of a particular magnitude will manifest itself against a system. For example, a seismic hazard curve describes the frequency of occurrence for earthquakes of a particular ground acceleration level. A hazard curve graph uses ground acceleration as the abscissa and frequency as the ordinate. Clearly, seismic hazard curves based on two different sources of data may converge, diverge, or cross. This would likely cause the resultant risk profiles to behave in a similar manner, and the deconvolution method is incapable of resolving this behavior.

† Recall that under the direct substitution method, our computations would consist of a detailed variability analysis in which only model features classified as being dominated by variability would be sampled, followed by a detailed uncertainty-variability analysis in which all model features for which distributions are generated would be sampled concurrently. Under the direct substitution method, the detailed variability computations serve as information only and are not used in the uncertainty-variability analysis. Under the deconvolution method, however, these same computations form the basis for the distribution V and thus must be performed in a manner that is consistent with the uncertainty-variability analysis.

are not computationally intensive (by comparison with LASEP-T and SPARRC) and are currently available in commercial software packages.

2.6.4 Uncertainty Analysis Implementation Issues

While the methods described in this section provide the theoretical underpinnings for the Cassini uncertainty analysis, there are a number of practical issues related to computational feasibility that must be resolved before these methods can be implemented. Specifically, it is not feasible to compute consequences for every known set of weather data and every individual LASEP-T trial that lead to a radiological release. The number of possible release/weather combinations is very large — too large even for a fast-running code such as SPARRC.

The objective is to design a scheme to select input for a limited number of representative SPARRC runs. Several criteria were developed by the Cassini project to help ensure the integrity and defensibility of the final risk uncertainty results:

1. The overall set of runs selected must be a fair representation of the spectrum of results that would be expected were we able to run all cases.
2. In order to ensure that the tails of the resulting risk distribution are not inadvertently neglected, the selection of runs must ensure that high-consequence situations are deliberately sampled (and properly accounted for) in spite of their low probability of occurrence.
3. To enhance the defensibility of the final product, the method should rely on the use of actual measured or computed data wherever possible (it should avoid homogenizing information from "similar" data points into a single "new" surrogate data point unless that point is already a member of the original data set).
4. The scheme must be compatible with the constraints imposed by all three of the uncertainty analysis methods described previously.

The first impulse might be to use random sampling techniques to draw a "representative" subset of data points to be used for the SPARRC analyses. However, if this were to be done, one would be likely to completely miss those rare large release scenarios and unusual weather scenarios that might lead to elevated consequences. Thus, these constraints — especially constraint 2 — cannot be met by a purely random sampling scheme such as a traditional Monte Carlo analysis. Instead, one must select the input for SPARRC runs using more sophisticated techniques. Section 3 lays the mathematical foundation for the advanced sampling techniques that were used in the Cassini risk assessment. The remainder of this paper then deals with the specific issues involved in implementing an advanced sampling scheme for the Cassini risk assessment.

2.7 Summary

This section has described the goals of the Cassini risk assessment project, as well as some of the practical and project-related constraints on the risk and uncertainty analyses. The objective of the analysis was to compute the risk of health effects and the risk of land contamination above EPA-established levels. This required the examination of a large number of scenarios in order to

determine their likelihood and their potential to produce consequences. The results of these scenarios were then aggregated to produce the final risk metrics. The uncertainty analysis was constrained by the following factors:

- More accident scenarios exist than the computational resources required to analyze them (hence, in some instances, surrogate scenarios must be used).
- The project team determined that the uncertainty analysis methodology must support the separation of random variability and epistemic uncertainty.
- The uncertainty analysis methodology must make use of the existing analysis tools (LASEP-T and SPARRC).
- The uncertainty analysis methodology must provide an assured success path while making possible the use of advanced techniques such as mathematical deconvolution.
- The methodology must ensure that high-consequence situations are deliberately sampled in spite of their low probability of occurrence in order to ensure that the tails of the risk distribution are not inadvertently neglected.

The sampling techniques developed for the Cassini uncertainty analysis fulfill these constraints while enabling a representative uncertainty analysis to be performed within the project's computational and schedule constraints.

3 Mathematical Background

There are two basic parts of the Cassini risk assessment: the accident scenario and source term analysis (as performed by LASEP-T) and the consequence analysis (as performed by SPARRC). The LASEP-T software is designed around a Monte Carlo sampling engine that is capable of performing the sampling required to implement all of the uncertainty analysis methods described in Section 2. This is done using two runs of the code: the first sampling only the variability distributions ("variables") and the second sampling both the variability distributions and the epistemic uncertainty distributions ("uncertain parameters"). For this reason, the LASEP-T analysis is not described further in this paper.

SPARRC, however, is a deterministic code that does not contain its own sampling apparatus. Instead, an external sampling engine must be used to select which specific cases are to be run. In an ideal world, one would take each release scenario* from LASEP-T and evaluate its consequences for several representative sets of variables and uncertain parameters under every possible set of weather conditions. However, while SPARRC is a fast-running code, it does require a few minutes of computation time per evaluation. For this reason, the project was forced to limit the number of SPARRC evaluations to be between 2000 and 3000 total. When we realize that there are two LASEP-T runs, each of which can produce a few thousand release scenarios, which would each have to be evaluated for some 150 sets of weather conditions and several representative sets of variables and uncertain parameters, it becomes clear that the ideal situation is not computationally tractable. Instead, we must select our SPARRC runs carefully in order to ensure that the needs of the uncertainty analysis methodology can be met. This selection was done for the Cassini risk assessment using concepts derived from Monte Carlo sampling, Latin hypercube sampling, importance sampling, and stratified sampling. This section provides the mathematical background for each of these methods that is needed to understand the method that was actually used in the Cassini analysis.

3.1 Monte Carlo Sampling

Monte Carlo uncertainty analysis techniques were developed to address the need for uncertainty assessment for a particular class of problems. Consider a variable Y that is a function of other variables $X_1, X_2, \dots, X_k$. This function may be very complicated, such as a computer model (e.g., SPARRC). A question to be investigated is "How does Y vary when the X s vary according to some assumed joint probability distribution?" Related questions are "What is the expected value of Y ?" and "What is the 99th percentile of Y ?"

A conventional approach to these questions is to apply Monte Carlo sampling. By sampling repeatedly from the assumed joint probability density function of the X s and evaluating Y for

* A "release scenario" is a single "trial" from LASEP-T that results in the release of any quantity of plutonium fuel at any point in the accident simulation. LASEP-T performs many thousands of trials in a single run, and it is possible for hundreds or thousands of these to result in some small release of plutonium and hence to become release scenarios.

each sample, the distribution of Y , along with its mean and other characteristics, can be estimated. In Monte Carlo sampling, the samples are drawn purely at random, and large numbers of samples must be drawn in order to ensure that the assumed joint probability density function is adequately represented by the sample population (e.g., that all regions of the density function have in fact been sampled, and that the density of samples approaches that demanded by the density function itself). The large number of samples also helps ensure that unintended correlations between the various X_i are eliminated.

Typical Monte Carlo sampling software (such as Sandia's LHS code⁷) provide as program output, for n Monte Carlo repetitions, a set of n vectors of input variables (each such vector is k -dimensional). Each input vector can then be evaluated by the function or program to generate n values of the result Y (Y may be a scalar or a vector whose dimensionality is determined by the function or program of interest). This approach yields reasonable estimates for the distribution of Y if the value of n is large. The value of n must be especially large if one wishes to characterize the tails of the distribution for Y . Since the LASEP-T results are likely to contain a small number of high-consequence release scenarios, and since these scenarios are critical to the final risk results, we are, in essence, required to perform a detailed characterization of the tail of the distribution provided by LASEP-T. Thus, if pure Monte Carlo were to be employed to select the runs to be performed by SPARRC, one would either be required to perform a prohibitively large number of SPARRC runs or risk failing to sample from this very important region. Therefore, pure Monte Carlo sampling is not an appropriate method for selecting the runs to be performed by SPARRC.

3.2 Latin Hypercube Sampling Theory

3.2.1 Sampling

An alternative approach, which can yield more precise estimates, is to use a constrained Monte Carlo sampling scheme. One such scheme, developed by McKay, Conover, and Beckman,¹⁴ is Latin hypercube sampling. Latin hypercube sampling selects n different values from each of k variables $X_1, \dots, X_k$ in the following manner. The range of each variable is divided into n nonoverlapping intervals on the basis of equal probability. One value from each interval is selected at random with respect to the probability density in the interval. The n values thus obtained for X_1 are paired in a random manner (equally likely combinations) with the n values of X_2 . These n pairs are combined in a random manner with the n values of X_3 to form n triplets, and so on, until n k -tuplets are formed. These n k -tuplets are the same as the n k -dimensional input vectors described for pure Monte Carlo sampling. This is the Latin hypercube sample. It is convenient to think of this sample (or any random sample of size n) as forming an $(n \times k)$ matrix of input where the i^{th} row contains specific values of each of the k input variables to be used on the i^{th} run of the computer model.

Latin hypercube sampling improves on the efficiency of traditional Monte Carlo sampling by ensuring that the n sampled values are selected so that they cover as much of the specified probability density as possible within the constraints that the resulting samples remain equally likely and still contain the element of randomness. To illustrate this point, consider drawing two samples from a known probability density function. If the samples are drawn purely at random,

then there is a 50% likelihood that both samples will be drawn from the same side of the median. If a third sample is drawn at random, there is still a 25% likelihood that *all three* samples will lie on the same side of the median. Clearly this is not desirable if one wishes to characterize the entire density function in terms of the sampled values. To overcome this artifact with pure random sampling, one must draw a very large number of samples — large enough that the probability of any important region remaining unsampled becomes small. Now consider drawing two samples from that same density function using Latin hypercube sampling. The Latin hypercube method first divides the density function into two equally likely regions (above the median and below the median), and then randomly draws exactly one sample from each of these regions of equal probability. In this way, one is guaranteed that the resulting samples are equally likely and yet still drawn at random while covering the entire domain of the density function in the fewest number of samples.

The Latin hypercube sampling technique has been applied to many different computer models since 1975.^{15, 16, 17, 18, 19} A more complete description of the method can be found in the user's manual for Sandia's LHS software.⁷

3.2.2 Pairing

It should be noted that even though two variables are sampled independently and paired randomly, the sample correlation coefficient of the n pairs of variables in either a random sample or a Latin hypercube sample will, in general, not equal zero because of sampling fluctuations. In order to obtain a sample in which the sample correlations more nearly match the assumed, or intended, correlations, Iman and Conover²⁰ proposed a method for restricting the way in which the variables are paired. While a full description of the restricted pairing methodology is beyond the scope of this paper, a heuristic description and justification is provided. The reader should consult the references^{6, 7, 20} for a more complete description of the method and its justification.

Recall that in the process of sampling a single distribution (generating n observations of values for that distribution), there is no significance to the order in which those observations are obtained. What matters for the purposes of the uncertainty analysis is that the entire *collection* of observations faithfully preserves the properties of the distribution from which it was sampled. For this reason, we can view the random pairing process (the basic Latin hypercube method) as follows. Imagine writing the value of each observation for distribution X_1 on a separate slip of paper and placing those slips into a hat. There would now be n separate slips of paper in that hat. Now imagine following the same process for distribution X_2 , and placing the slips in a second hat. The same process could be followed for distributions X_3 through X_k , so that at the end of the process one would have k hats, each containing n slips of paper with distribution values written on them. To accomplish the random pairing process described previously, one would simply draw one slip of paper at random from each of the k hats, and the group of values written on those k slips of paper would form the first observation of the output data set. In the language of the previous section, these values would form the k -dimensional input vector for the first computer run, or the first row in the $(n \times k)$ matrix of sampling results. The second set of values drawn randomly from the k hats would form the second observation, k -dimensional input vector, or matrix row, and so forth until the last, or n^{th} , set of values was drawn to form the last such observation. Since exactly n values were generated for each distribution, all of the hats are now

empty. Thus, all of the generated distribution values were used, so each column in the $(n \times k)$ matrix faithfully represents the distribution from which its values were drawn.

One could, however, visualize a different pairing process in which the slips of paper from the k hats were poured out to make k separate piles. The slips of paper from those piles could then be arranged in columns to form the $(n \times k)$ matrix described previously. Since there is no significance to the *order* in which the individual slips of paper are arranged in these columns, one might imagine a person examining the entire $(n \times k)$ matrix and *deliberately* moving slips of paper around within columns to achieve some goal for the overall matrix. If one were clever, one could order each column so that the correlation between its values and those of every other column in the matrix was as small as possible. When this process is completed, the $(n \times k)$ matrix still contains the same values in the same columns, but each column has been intentionally reordered. Each column in the matrix faithfully represents the distribution from which its values were drawn because, once again, all of the values generated for the distribution were used. Also, each row still represents the k -dimensional input vector for a computer run because it contains one value from each of the distributions X_1 through X_k . Thus the $(n \times k)$ matrix is in every sense equivalent to that generated by the random pairing process described earlier. In fact, there is some small probability that the random pairing algorithm would generate this matrix if exactly the right sequence of random draws were to occur during the process. This intentional selection of the pairing of the selected observations produces results that are directly analogous to the restricted pairing procedure developed by Iman and Conover. Since this process explicitly removes unintended correlations from the sample set, there is no longer any need to remove such correlations using the principle that large number of observations cause random pairing to asymptotically approach statistical independence. Thus, by intentionally pairing the data, one can reduce the number of observations required to appropriately represent a set of input distributions. For many problems, reasonable results can be obtained when the number of observations is at least $4/3$ the number of random variables being sampled.⁶

If one uses Latin hypercube sampling with restricted pairing, as implemented in Sandia's LHS software,⁷ it is important to understand the software's degree of success in achieving zero correlation between the sampled random variables. The LHS software computes a measure of this success known as the variance inflation factor (VIF). The VIF is defined as the largest element on the diagonal of the inverse of the achieved correlation matrix. If the VIF gets significantly larger than 1, some undesirably large pairwise correlations may be present. Marquardt and Snee²¹ deal with some very large VIFs ($> 2 \times 10^6$) and provide a readable explanation on reasonable sizes of VIFs. Marquardt²² indicates that there can be serious colinearity (i.e., large pairwise correlations present) for $VIF > 10$. Thus, there is certainly no problem as long as the VIF is close to 1.⁶

It is logical to ask then whether this restricted pairing technique can be used to induce specified *nonzero* correlations between random variables since it is already being used to cause these same correlations to tend toward zero (or, more properly, become as small as possible). The answer is yes, with the limitation that such induced correlations are based on the nonparametric technique known as rank correlation. Such a measure is used since it remains meaningful in the presence of non-normal distributions on the input variables.

3.2.3 Application to the Cassini Risk Assessment

There are four classes of information that must be selected in order to provide input to the SPARRC consequence analysis software: (1) values from random variability distributions, (2) values from epistemic uncertainty distributions, (3) the release scenario and source term to be modeled, and (4) the weather conditions under which this release is presumed to occur. Latin hypercube sampling is clearly applicable to the first two classes of information because the distributions for these variables and parameters are generally continuous* and well behaved (e.g., their density functions do not have extraordinarily long tails,† and there are no regions of the distribution that exhibit both a low probability density and a high importance for the analysis results).‡ However, Latin hypercube sampling is not appropriate for the third class of information because the release scenario and source term data contain a class of very low-probability, high-consequence (hence, high importance) events that would likely not make it into the sampled data. Latin hypercube sampling is also less appropriate for the fourth class of information for the same reason — the weather information is dominated by a large number of “normal” days and contains only a few days that would be expected to produce the most severe consequences. Thus, while Latin hypercube sampling forms an important part of the Cassini uncertainty assessment, other techniques must be used for the release scenario and weather data.

3.3 Importance Sampling

Consider a function f , and suppose one wished to integrate f over some region in x from a to b . The basic theorem of Monte Carlo integration²³ states, in essence, that the value of this integral I is equal to the average value of f on the interval $[a, b]$, times the “volume” of that interval $(b-a)$, or

* Or, where the distributions are discontinuous or discrete, it is possible to perform an *a priori* ordering of the domain so that the regions of equal probability generated within LHS represent regions where the entities that they contain are in some sense similar to one another.

† A “long tail” might, for example, be found in a density function where the difference between the 99th percentile value and the 99.9th percentile differed by more than an order of magnitude. Latin hypercube sampling can perform poorly for distributions with long tails. Latin hypercube sampling fails for distributions with long tails because a typical Latin hypercube sample consists of only a few tens or hundreds of observations, and only one or two of those observations are taken from the tail of the distribution. This makes the mean, variance, and other statistics of the sample much more highly sensitive to the exact values selected for those one or two samples.

‡ Latin hypercube sampling also fails when a region of a distribution that is highly important to the analysis results exhibits a relatively low probability density. This occurs because the Latin hypercube method divides the domain for each random variable into regions of *equal probability*, and draws one sample from each such region. Obviously, if a region of importance has a very low probability, it is likely to be fully contained within one of these equal probability regions. In fact, it may make up just a very small portion of one such region. Thus, it is likely that the important value for this random variable will not be sampled *at all* because the one sample drawn from this region is most likely to be drawn from the larger, less important portion of the region. Pure Monte Carlo sampling also suffers from this same problem unless an extraordinarily large number of samples are drawn. This problem can be overcome using importance sampling or more general stratified sampling techniques.

$$I = \int_a^b f \cdot dx \approx (b-a) \cdot \left(\langle f \rangle \pm \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{n}} \right) = (b-a) \cdot \langle f \rangle \pm (b-a) \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{n}}$$

where the angle brackets denote the arithmetic mean over the n sample points. The “plus or minus” term represents a one standard deviation error estimate for the integral. Thus, given n points that have been chosen at random (uniformly) from the x domain within the region $[a, b]$, denoted x_i , the value of I can be estimated as

$$I \approx (b-a) \cdot \frac{1}{n} \sum_{i=1}^n f(x_i)$$

This equation would estimate the integral I using pure Monte Carlo sampling. Now suppose that the original function f can be written as the product of two functions such that $f = g \cdot h$. One could then rewrite the original integral I as

$$I = \int_a^b f \cdot 1 \cdot dx = \int_a^b \left(\frac{f}{g} \right) \cdot g \cdot dx = \int_a^b h \cdot g \cdot dx$$

Suppose furthermore that g represents a probability density function such that

$$\int_a^b g(x) \cdot dx = 1$$

If one then chooses points x_i randomly based upon the nonuniform density function $g(x) \cdot dx$, instead of the uniform density function $1 \cdot dx$ used in the original Monte Carlo integration function, then one can estimate the value of I by sampling the function h at these nonuniform points x_i . The estimate for I is then²³

$$I = \int_a^b f \cdot dx = \int_a^b \frac{f}{g} \cdot g \cdot dx \approx (b-a) \cdot \left[\left\langle \frac{f}{g} \right\rangle \pm \sqrt{\frac{\langle f^2/g^2 \rangle - \langle f/g \rangle^2}{n}} \right] = (b-a) \cdot \left[\langle h \rangle \pm \sqrt{\frac{\langle h^2 \rangle - \langle h \rangle^2}{n}} \right]$$

or, again based on the nonuniform x_i ,

$$I \approx (b-a) \cdot \frac{1}{n} \sum_{i=1}^n h(x_i) = (b-a) \cdot \frac{1}{n} \sum_{i=1}^n \frac{f(x_i)}{g(x_i)}$$

This technique is known as *importance sampling*, and can be thought of as a “variance reduction” technique in that it provides a method to reduce the variance of the sample mean that must be calculated during a Monte Carlo integration. It can be shown that the optimal choice for the function g (the choice that achieves the greatest variance reduction and thus enables one to perform the integration using the fewest samples) is to make g proportional to the absolute value of f .²³ However, since it is often very computationally expensive to determine values for f (and

we want to reserve that computational expense for the actual integral computation), a good choice for g can be derived from any reasonable approximation for f .

To further illustrate the method of importance sampling, consider the integration of a function f whose value is very small over most of the domain of integration, but very large in some small portion of the domain. Since the function g is optimally selected to be roughly proportional to the absolute value of f , and since g represents the density function from which the sample points x_i will be drawn, it is clear that the best estimate for I is obtained when one samples heavily from the areas where f is large and lightly from the areas where f is small. This variation in sampling frequency is accounted for by the fact that the final value of I is now derived from the function

$h = \frac{f}{g}$ instead of simply from f as in the pure Monte Carlo integration.

Concepts from importance sampling were used in the selection of source terms for the Cassini uncertainty analysis. The specifics of that method are described at the end of this section.

3.4 Stratified Sampling

A second variance reduction technique for Monte Carlo integration is stratified sampling. Latin hypercube sampling is one example of stratified sampling. More general stratified sampling techniques²⁴ do not make use of a single observation per stratum, as was the case with Latin hypercube sampling. This section provides a sufficient overview of stratified sampling to allow the reader to understand how this technique was used in the selection of weather data for the Cassini uncertainty analysis. For a more detailed description, the reader is referred to the literature.

In stratified sampling, a population of size N is first divided into L subpopulations (called "strata") of sizes $N_1, N_2, \dots, N_L$, respectively. These subpopulations are nonoverlapping, and together they comprise the whole population. The objective of the stratification process is to improve the precision in the estimates of the characteristics of the whole population by dividing a heterogeneous population into subpopulations, each of which is internally more homogeneous. A more homogeneous population within a stratum implies less variation within that population, so precise estimates of the statistical characteristics of that stratum can be obtained from a small sample in that stratum. These estimates can then be combined into a precise estimate for the whole population — usually with a significantly smaller sample size than would be required were one to perform random sampling on the entire population N . Each subpopulation is sampled independently using an appropriate sampling technique such as simple random sampling or Latin hypercube sampling.

In stratified sampling, the size of the subpopulations can vary by stratum, as can the fraction of the subpopulation that is sampled (or the number of samples selected per stratum). The objective is to select enough samples from each stratum that its statistical properties can be accurately determined. Obviously, a highly homogeneous stratum will require less rigorous sampling than a heterogeneous stratum. More specifically, the optimal allocation of sampling points among subpopulations is to have the number of sampled points in each subpopulation j proportional to σ_j , where σ_j is defined as the square root of the variance within that subpopulation.²³ If a stratum

is small and heterogeneous, this may require the sampling of every member of that subpopulation. This will not, however, increase the *overall* sampling requirements if other strata are largely homogeneous and, hence, can be sampled more lightly than would have been required had the members of the heterogeneous stratum been force-fit into other strata.

In the Cassini risk assessment study, analysts divided the measured weather data into strata. The objective was to produce several strata that were relatively homogeneous with respect to wind characteristics, expected consequences for a series of trial source terms, and so forth. Two final strata were reserved for weather data that produced either very high or very low consequences, respectively. The heterogeneous "bad days" category was kept small enough that every member of this population could be sampled during the analysis. Other strata were sampled using Latin hypercube techniques. More details on weather grouping and sampling are provided in later sections of this paper.

3.5 Summary

The mathematical methods described in this section each played a role in the final sampling scheme that was implemented for the Cassini risk uncertainty analysis. Latin hypercube sampling was used to sample the variability distributions and uncertain parameters, importance sampling concepts were used to select release scenarios, and stratified sampling was used to select weather data. The next section describes how these techniques were brought together mathematically in the Cassini analysis.

4 Methods Used for the Cassini Study

The objective of the Cassini risk uncertainty analysis was to provide two sets of information for decision makers: total risk estimates and exceedance frequency curves. We first describe the construction of the total risk estimate, and as part of that discussion, lay out and justify the sampling method used for the Cassini analysis. We then show how the samples drawn for the total risk estimate were used to construct the exceedance frequency curves that were both plotted individually and used in the deconvolution assessment method.

4.1 The Cassini Total Risk Estimate

Let us now turn to the computation of risk for the Cassini mission. For this project, the risk R was to be computed for each release scenario as the product of its frequency f and its consequences C , where consequences are a function of vectors that describe variability distributions v , uncertain parameters p , the release characteristics m , and weather conditions W . Thus, each of the s release scenarios generated by LASEP-T could lead to n consequence computations as all possible combinations of variables, parameters, and weather data are considered. The risk was to be summed over all possible release scenarios and potential consequence computations to obtain the total mission risk as follows:

$$R = \sum_{i=1}^s R_i = \sum_{i=1}^s \sum_{j=1}^n f_{i,j} \cdot C(v_j, p_j, m_i, W_j)$$

Here $f_{i,j}$ is the overall frequency associated with each consequence calculation. This frequency can be broken down into two parts: the release scenario frequency f_i and the conditional probability $P_{C,i,j}$ that release scenario i will result in the j^{th} consequence computation being performed. For the Cassini analysis, f_i can be further broken down into two parts: the accident frequency (as defined by the launch vehicle databook²) f_{DB} and the conditional probability $P_{R,i}$ of this release scenario given the occurrence of an accident. This conditional probability is "computed" by LASEP-T because, since LASEP-T is a pure Monte Carlo simulation tool, every trial is viewed as equally likely* (there is no mathematical basis for assuming otherwise). Thus, $P_{R,i} = P_R$, where P_R is computed as 1 divided by the total number of trials t (since this represents one trial out of the t trials that were generated during this LASEP-T analysis). In other words, for each databook accident, $P_{R,i}$ is the same for all release scenarios generated by LASEP-T.[†] This means that the equation for R can be rewritten as

* Assuming, of course, that LASEP-T generates enough Monte Carlo trials to be an adequate statistical representation of the problem at hand. That, however, is assumed not to be an issue in this discussion.

† For the sake of completeness, the v_i and p_i are also equally likely because they are selected through Latin hypercube sampling. Each individual set of weather conditions W_i is also equally likely, although these will also be sampled, as described later in this paper.

$$\begin{aligned}
R &= \sum_{i=1}^s R_i = \sum_{i=1}^s \sum_{j=1}^n f_{DB} \cdot P_{R,i} \cdot P_{C,i,j} \cdot C(v_j, p_j, m_i, W_j) \\
&= f_{DB} \cdot \sum_{i=1}^s P_{R,i} \cdot \sum_{j=1}^n P_{C,i,j} \cdot C(v_j, p_j, m_i, W_j) = f_{DB} \cdot P_R \cdot \sum_{i=1}^s \sum_{j=1}^n P_{C,i,j} \cdot C(v_j, p_j, m_i, W_j)
\end{aligned}$$

Note that since $P_{C,i,j}$ is a series of conditional probabilities, it must sum to unity when summed over all i and j . If we perform the same consequence computations for every release scenario, then $P_{C,i,j}$ becomes independent of i . In this case, the risk computation for the i^{th} release scenario becomes simply

$$R_i = f_{DB} \cdot P_R \cdot \sum_{j=1}^n P_{C,j} \cdot C(v_j, p_j, m_i, W_j)$$

Let us now further assume that the series of n consequence computations that are to be performed on the i^{th} release scenario is built around two independent schemes: one that selects the values to be used for variables and parameters, and a second that selects the weather conditions that are to be used for each consequence computation. If one were to, as assumed thus far, perform every possible combination of consequence computations, then the n consequence computations would come about by combining each of the v sets of variable and parameter values to be evaluated with each of the u sets of possible weather conditions, so that in fact $n = u \cdot v$. Since these sets of input parameters are generated independently of one another, the conditional probability $P_{C,j}$ could be further broken down into $P_{V,k}$, the conditional probability that the k^{th} set of variables and parameters was used for this computation, and $P_{W,l}$, the conditional probability that the l^{th} set of weather conditions was used. Then the risk computation for the i^{th} release scenario becomes

$$R_i = f_{DB} \cdot P_R \cdot \sum_{k=1}^v \sum_{l=1}^u P_{V,k} \cdot P_{W,l} \cdot C(v_k, p_k, m_i, W_l)$$

Note that all sets of weather data are assumed to be equally likely since each represents a set of observed conditions from one day within the launch window during previous years. Therefore, all values of $P_{W,l}$ are also equal. In addition, if the sets of variables and parameters are developed in such a manner as to be of equal probability (using methods such as Monte Carlo or Latin hypercube sampling), then all values of $P_{V,k}$ are equal, so the risk computation for the i^{th} release scenario becomes

$$R_i = f_{DB} \cdot P_R \cdot P_V \cdot P_W \cdot \sum_{k=1}^v \sum_{l=1}^u C(v_k, p_k, m_i, W_l)$$

Recall that P_R is simply the inverse of the number of release scenarios, P_V is the inverse of the number of sets of variables and parameters to be evaluated, and P_W is the inverse of the number of known weather data sets. If these values are used in the above equation, it becomes

$$R_i = f_{DB} \cdot \frac{1}{t} \cdot \frac{1}{v} \cdot \frac{1}{u} \cdot \sum_{k=1}^v \sum_{l=1}^u C(v_k, p_k, m_i, W_l)$$

Thus far, we have not reduced the number of consequence computations that must be performed. We have simply applied some constraints on *how* those consequence computations would be performed — namely, that all release scenarios are equally likely, as are all sets of weather data and all sets of variable and parameter values. Let us now consider applying a stratified sampling technique to the weather data so that, instead of performing the consequence computation for every set of weather data, we perform it only for some sampled subset of that data. In principle, we can state that the sum of consequences over all sets of weather data and over all sets of variable and parameter values is identical to the number of weather data sets u , times the number of variable and parameter value sets v , times the mean value of the consequences (if one were to actually perform the computation for every combination of these data sets and include it in the computation of the mean). Thus,

$$\sum_{l=1}^u \sum_{k=1}^v C(v_k, p_k, m_i, W_l) = u \cdot v \cdot \langle\langle C_i \rangle\rangle$$

where $\langle\langle C_i \rangle\rangle$ is the exact mean value of consequences over all of the weather data sets for the i^{th} release scenario and the k^{th} set of variable and parameter values.* Let us now replace this exact mean with an estimator $\langle C_i \rangle$ that is derived from n_s values. The above equation then becomes

$$\sum_{l=1}^u \sum_{k=1}^v C(v_k, p_k, m_i, W_l) \approx u \cdot v \cdot \left(\langle C_i \rangle \pm \sqrt{\frac{\langle C_i^2 \rangle - \langle C_i \rangle^2}{n_s}} \right) \approx u \cdot v \cdot \langle C_i \rangle$$

Given that the further tracking of the error term would only add complexity to the discussion without providing most readers with additional insight, it will be omitted from the remainder of this discussion. The next step is to replace this simple estimator with one that is derived from sampling. We will use a stratified sampling scheme for the weather data, and Latin hypercube sampling for the variable and parameter data. Furthermore, we will draw the same number of weather samples from each stratum as there are sets of variable and parameter data (v).

Let the u members of the weather data population now be separated into L subpopulations, with N_p members in the p^{th} subpopulation. Let each subpopulation be sampled v times. Thus, in the k^{th} sample, one will randomly draw one weather data set from each of the L subpopulations. These L sets of weather data will then be paired with the k^{th} set of variable and parameter values to produce L consequence evaluations for the k^{th} sample. This process is repeated v times — each repetition uses a different set of v_k and p_k , plus a different set of L weather data sets. The estimator for the mean for this sampling process is

* Here we have reversed the order in which the sums are taken. This can be done without a problem because there are no quantities within the double sum that depend only upon one index or the other.

$$\langle C_i \rangle = \frac{\sum_{p=1}^L N_p \cdot \frac{\sum_{k=1}^v C_{i,k,p}}{v}}{u}$$

where $C_{i,k,p}$ represents a single consequence computation for the i^{th} release scenario using the k^{th} set of variable and parameter values as well as the k^{th} set of weather data from the p^{th} stratum. This estimator itself consists of two parts: the innermost sum is a simple sample mean for those samples that are within a single stratum. The outer sum then constructs the mean for the stratified sample based upon the appropriate weighting of means from the individual strata.²⁴ Note that the sum of all N_p must be u . This estimator can be rewritten as

$$\langle C_i \rangle = \frac{1}{u \cdot v} \sum_{p=1}^L N_p \cdot \sum_{k=1}^v C_{i,k,p}$$

so that

$$\begin{aligned} \sum_{l=1}^u \sum_{k=1}^v C(v_k, p_k, m_i, W_l) &\approx u \cdot v \cdot \langle C_i \rangle \\ &= u \cdot v \cdot \frac{1}{u \cdot v} \sum_{p=1}^L N_p \cdot \sum_{k=1}^v C_{i,k,p} \\ &= \sum_{p=1}^L N_p \cdot \sum_{k=1}^v C_{i,k,p} \end{aligned}$$

and, thus,

$$\begin{aligned} R_i &\approx f_{DB} \cdot \frac{1}{t} \cdot \frac{1}{v} \cdot \frac{1}{u} \cdot \sum_{p=1}^L N_p \cdot \sum_{k=1}^v C_{i,k,p} \\ &\approx \frac{f_{DB}}{t} \cdot \frac{\sum_{p=1}^L N_p \cdot \frac{\sum_{k=1}^v C_{i,k,p}}{v}}{u} \end{aligned}$$

This is the risk associated with a single release scenario. The first form of the above equation is more suitable for software development, while the second provides a better physical interpretation. The first term represents the frequency of the release scenario as the frequency of the databook accident divided by the number of LASEP-T trials (the conditional probability of *this* release scenario is simply the inverse of the number of trials since all trials are equally likely). The second term (the first sum) represents the stratification of the weather data, where we must construct a weighted average over all strata (N_p/u is the weight of the p^{th} stratum). The final term represents a simple sample mean over all of the consequence computations performed for weather within the p^{th} stratum. Each such consequence computation uses different weather data and different variable and parameter values.

If one were actually to be able to perform this assessment for every release scenario, then one could simply sum all of the R_i to obtain the total risk for this databook accident. However, there are simply too many scenarios for this to be computationally possible. Therefore, let us return to the original formulation for R and consider what happens when one performs a form of importance sampling on the release scenarios. Originally,

$$R = \sum_{i=1}^s R_i$$

Let us again replace the sum over all scenarios with the number of release scenarios s times the exact average risk, and then create an estimator for that average.

$$R = \sum_{i=1}^s R_i = s \cdot \langle R \rangle \approx s \cdot \left(\langle R \rangle \pm \sqrt{\frac{\langle R^2 \rangle - \langle R \rangle^2}{n_s}} \right) \approx s \cdot \langle R \rangle$$

where n_s is the number of samples examined when computing the mean risk $\langle R \rangle$, and the "plus or minus" term represents a one standard deviation error estimate that, for the sake of clarity, will be neglected for the rest of this discussion. We seek an efficient method to compute the mean risk. Since the individual values R_i within this average can vary by several orders of magnitude, we need a variance reduction scheme to make the computation tractable. We will use a crude version of importance sampling in which the function g from the previous discussion is represented by a simple three-step function.

Recall from the previous discussion that the most efficient sampling scheme is obtained when the function g is proportional to the magnitude of the function being summed (in this case, R_i). We do not know R_i (otherwise it would not be a problem to compute the mean risk). However, we do have a surrogate by which we can estimate the relative magnitude of each R_i : the mass of plutonium fuel released during each release scenario m_i . Generally, where m_i is large, there is likely to be larger consequences than where m_i is small. For this reason, we sorted the release scenarios in order of decreasing m_i , and sampled those scenarios with the greatest values the most rigorously. We did this by selecting the cutoff points between the three portions of the step function so that there were very few points M_1 in the first section, more points M_2 in the second section, and the vast majority of points M_3 in the third section (all of which had very small releases and, thus, would be expected to produce extremely small consequences and extremely small values for R_i). Note that $M_1 + M_2 + M_3 = s$, which is the total number of release scenarios generated by LASEP-T. Thus, the risk equation could be rewritten as

$$R = \sum_{h=1}^{M_1} R_{h,1} + \sum_{h=1}^{M_2} R_{h,2} + \sum_{h=1}^{M_3} R_{h,3} \approx M_1 \cdot \langle R_1 \rangle + M_2 \cdot \langle R_2 \rangle + M_3 \cdot \langle R_3 \rangle$$

where $\langle R_g \rangle$ is the sample mean risk of the release scenarios within the g^{th} section of the importance step function. The importance of each section is incorporated in the size of that

section. Let us now estimate each of these means by simple random sampling.* Let us take v samples from each of these sections, and use those v values for $R_{h,g}$ to estimate the respective sample means.† The risk expression, using these sample means, can then be written as

$$\begin{aligned} R &\approx \frac{M_1 \cdot \sum_{h=1}^v R_{h,1}}{v} + \frac{M_2 \cdot \sum_{h=1}^v R_{h,2}}{v} + \frac{M_3 \cdot \sum_{h=1}^v R_{h,3}}{v} \\ &\approx \frac{1}{v} \sum_{h=1}^v (M_1 \cdot R_{h,1} + M_2 \cdot R_{h,2} + M_3 \cdot R_{h,3}) \\ &\approx \sum_{z=1}^3 \frac{M_z}{v} \cdot \sum_{h=1}^v R_{h,z} \end{aligned}$$

Recall that

$$R_i \approx f_{DB} \cdot \frac{1}{t} \cdot \frac{1}{v} \cdot \frac{1}{u} \cdot \sum_{p=1}^L N_p \cdot \sum_{k=1}^v C_{i,k,p}$$

Transforming this to the notation required by the importance sampling formulation, we get

$$R_{h,z} \approx f_{DB} \cdot \frac{1}{t} \cdot \frac{1}{v} \cdot \frac{1}{u} \cdot \sum_{p=1}^L N_p \cdot \sum_{k=1}^v C_{h,z,k,p}$$

Thus,

$$\begin{aligned} R &\approx \sum_{z=1}^3 \frac{M_z}{v} \cdot \sum_{h=1}^v R_{h,z} \\ &\approx \sum_{z=1}^3 \frac{M_z}{v} \cdot \sum_{h=1}^v f_{DB} \cdot \frac{1}{t} \cdot \frac{1}{v} \cdot \frac{1}{u} \cdot \sum_{p=1}^L N_p \cdot \sum_{k=1}^v C_{h,z,k,p} \end{aligned}$$

Since the indices h and k are over the same range, let us assign them to be equal. Recall that k was used to identify which set of variable and parameter values, as well as which set of weather data would be used (from the p^{th} weather stratum) in the consequence computation. Since h was a replacement for i , it denotes which release scenario is to be modeled (from the z^{th} section of the importance step function) in the consequence computation. The practical implication of this step is as follows: If h and k are not assigned to be equal, then we take each sampled release scenario h and evaluate it over *all values* of k — in other words, each sampled release scenario is evaluated for all sampled weather data and variable and parameter value sets. When h and k are

* In reality, Latin hypercube sampling was used, but that is not important for this discussion.

† The decision to use v samples — the same number of samples used to generate the variable and parameter value sets — was made in order to allow all such sampling to be performed within a single execution of the LHS Latin hypercube sampling software.

assigned to be equal, we evaluate that same release scenario h only for the $k = h^{\text{th}}$ set of weather data and variable and parameter values. Since we are only constructing a statistical estimator here, the restriction is not a problem *per se*, so long as enough samples are taken to adequately characterize the overall population. In other words, this restriction causes the size of the sample to be reduced by a factor of v , and this increases the error estimator accordingly, but that is the only statistical effect on this calculation.

When the indices h and k are assigned to be equal, the formula for risk changes slightly. After rearranging the above equation slightly, it becomes

$$R \approx f_{DB} \cdot \frac{1}{t} \cdot \frac{1}{u} \cdot \frac{1}{v} \cdot \frac{1}{v} \cdot \sum_{z=1}^3 M_z \cdot \sum_{p=1}^L N_p \cdot \sum_{k=1}^v \sum_{k=1}^v C_{k,p,z}$$

Note that we are now summing over the same index twice. The first (rightmost) summation aggregates the various consequence computations into a single value. The second summation simply takes v instances of this identical value and adds them up. This is equivalent to multiplying that original single value by the constant v , so the equation above can be reduced to its final form as follows:

$$R \approx f_{DB} \cdot \frac{1}{t} \cdot \frac{1}{u} \cdot \frac{1}{v} \cdot \frac{1}{v} \cdot \sum_{z=1}^3 M_z \cdot \sum_{p=1}^L N_p \cdot v \cdot \sum_{k=1}^v C_{k,p,z}$$

or, after canceling one factor of v , we obtain the final integrated risk equation

$$R \approx \frac{f_{DB}}{t \cdot v \cdot u} \cdot \sum_{z=1}^3 M_z \cdot \sum_{p=1}^L N_p \cdot \sum_{k=1}^v C_{k,p,z}$$

Let us now compare this formula with the probability of consequence (POC) formula used in the Cassini analysis.¹

$$POC = POF \times POR \times PFC \times PWC \times POB$$

where POC is the unconditional probability associated with a particular consequence computation, POF is the probability of failure of the launch vehicle, POR is the conditional probability that a fuel release occurs given that an accident has occurred, PFC is the conditional probability that this particular fuel release falls within the range for this cluster given that an accident causing a fuel release has occurred, PWC is the conditional probability that the fuel release falls on a weather day that is within this weather cluster, and POB is the probability of this LHS observation. If we rewrite these factors in terms of the notation in this paper, we obtain

$$POF = f_{DB}$$

$$POR = \frac{s}{t}$$

$$PFC = \frac{M_z}{s}$$

$$PWC = \frac{N_p}{u}$$

$$POB = \frac{1}{v}$$

so that

$$\begin{aligned} POC &= f_{DB} \cdot \frac{s}{t} \cdot \frac{M_z}{s} \cdot \frac{N_p}{u} \cdot \frac{1}{v} \\ &= \frac{f_{DB}}{t \cdot u \cdot v} \cdot M_z \cdot N_p \end{aligned}$$

This can be seen to be equivalent to the coefficient for the k^{th} observation in the z^{th} portion of the importance step function and the p^{th} weather stratum from the final integrated risk equation. Thus, the sampling method used for the Cassini risk assessment is a faithful implementation of the sampling methods described in this section.

Let us now seek a physical interpretation of the final integrated risk equation. In order to make this interpretation more clear, let us rewrite that equation as

$$R \approx f_{DB} \cdot \sum_{z=1}^3 \frac{M_z}{t} \cdot \sum_{p=1}^L \frac{N_p}{u} \cdot \frac{\sum_{k=1}^v C_{k,p,z}}{v}$$

Now let us think of the three portions of the importance step function and the L weather strata as forming a $(3 \times L)$ grid, and every possible combination of release scenarios and weather conditions must fall into exactly one of these grid spaces. Under this interpretation, f_{DB} represents the frequency with which an accident occurs, $\frac{M_z}{t}$ represents the conditional probability that this accident causes a release of nuclear material that falls within the parameters of the z^{th} portion of the importance step function, and $\frac{N_p}{u}$ represents the conditional probability that such a release occurs on a day with weather that is described by a member of the p^{th} weather stratum. The final term represents an estimator for the mean value of the consequences that would occur over the entire population of release scenarios occurring within the (z,p) cell of this grid. That estimator considers v samples from this population, and each such sample consists of a unique set of variable and parameter values, a unique set of weather data (from the p^{th} stratum), and a unique release scenario (from the z^{th} portion of the importance step function). The v in the denominator of this term is necessary in order for it to be an estimator for the mean of the stated population. From this description, it is clear that each $C_{k,p,z}$ represents one evaluation by the

consequence modeling software, and that using this estimator for total risk requires $e = 3 \cdot L \cdot v$ such evaluations. Thus, using 10 weather strata and 30 samples (v) requires 900 executions of the consequence software to obtain an estimate of the total risk. This number increases rapidly with increasing v .

Note that each k in the above equation represents a random draw from *each* of the populations. Thus, it represents a random selection for each variable and each parameter, as well as a random weather sample and a random release scenario from each cell. In practice, this sampling was accomplished in the LHS software⁷ as follows: Each variable (and for the variability/uncertainty combination runs, each parameter) was represented as a distribution input to LHS. Furthermore, a separate discrete distribution was developed for each weather data stratum so that all data points within that stratum were equally likely. Similar discrete distributions were developed for each portion of the importance step function. This enables one to generate a point within each cell of the $(3 \times L)$ grid for each observation by combining the randomly generated release scenario and the randomly generated weather sample that fall within that cell. An artifact of this sampling method is that all three release scenarios for a particular LHS observation will be evaluated for the same L weather data points, and that all L weather data points will similarly be evaluated using the same three release scenarios (both the weather data points and the release scenarios will change for the next observation). This was done in order to reduce the number of random variables that were to be sampled by LHS to $(3 + L)$ as opposed to the $(2 \times 3 \times L)$ that would have otherwise been required. This enabled LHS to represent and pair the data using fewer observations. It is believed that this sampling method had a negligible impact on the statistical validity of the analysis results.

While it will not be demonstrated here, there is one more important aspect of the total risk formula derived above: the samples that it generates can be used not only to generate an estimate for the total risk, an exceedance frequency curve, and an estimate of the uncertainty in that exceedance frequency curve (by way of the deconvolution method), but they can also be used to support the characterization of risk uncertainty using the direct substitution and replica event tree methods described earlier in this paper. This is important because it ensured that even if the then-unproven deconvolution method failed, the project could achieve some measure of risk uncertainty characterization using the same sampled data (and the implied computationally expensive set of consequence computations) to characterize the risk uncertainty to a greater degree than had been done for any previous RTG-powered space mission.

4.2 The Cassini Exceedance Frequency Curves

We now turn our attention to the development of exceedance frequency curves from the set of consequence computations that were performed as a result of the sampling method described in the previous section. Let us begin by rewriting the total risk function as

$$R \approx \sum_{z=1}^3 \sum_{p=1}^L \sum_{k=1}^v \frac{f_{DB}}{t \cdot v \cdot u} \cdot M_z \cdot N_p \cdot C_{k,p,z}$$

To construct an exceedance frequency curve, one must have a series of ordered pairs, each of which consists of a consequence value and the frequency associated with that consequence

value.* Recall that the computation of total risk requires one to multiply frequency times consequences and sum over all scenarios. We can separate the total risk equation into two parts to obtain a consequence estimate part ($C_{k,p,z}$) and a frequency estimate part (all of the other terms within the triple sum). These two parts can then be viewed as an ordered pair, and the set of ordered pairs can then be plotted as an exceedance frequency. This ordered pair, which will be denoted as R^P , can then be written as

$$R_{k,p,z}^P = \left(\frac{f_{DB} \cdot M_z \cdot N_p}{t \cdot v \cdot u}, C_{k,p,z} \right)$$

Again, this ordered pair represents the frequency and consequence for the p^{th} weather stratum, the z^{th} portion of the importance step function, and the k^{th} LHS sample. The sampling analysis produces $e = 3 \times L \times v$ such ordered pairs. Note that the frequency portion of this ordered pair is identical to the formula for POC described in the previous section and implemented in the Cassini risk assessment.¹

The actual construction of the exceedance frequency curve from this set of ordered pairs is now a simple task. We sort the set of e ordered pairs in order of *decreasing consequence*. For the sake of convenience, an individual ordered pair will now be referred to as R_i^P , where i ranges from 1 to e . Also, we will insert into the ordered pair place holders so that we remember that each such ordered pair is explicitly related to particular values of k , p and z . The augmented pair can now be written as

$$R_i^P = \left(k_i, p_i, z_i, f_i = \frac{f_{DB} \cdot M_z \cdot N_p}{t \cdot v \cdot u}, C_i = C_{k,p,z} \right)$$

The q^{th} point in the exceedance frequency curve represents a point that consists of the consequence C_q from the q^{th} augmented pair R_q^P as the abscissa, and the sum of all f_i where $1 \leq i \leq q$ as the ordinate. This point can be written as

$$E_i = \left(C_i, \sum_{i=1}^q f_i \right)$$

Since the C_i have been sorted into decreasing order, increasing i results in points that have continuously decreasing abscissas and continuously increasing ordinates, so that the e^{th} point plots the lowest consequence against the total frequency.

* Another way to view the total risk estimate computation is that it was originally based on these ordered pairs, and that one computes total risk by multiplying the frequency times the consequence within each ordered pair, and then sums over all such ordered pairs. All release scenarios must be represented in either computation in order for the risk estimate or exceedance frequency curve to be valid, and the method by which the statistical estimators were constructed for the total risk computation ensures that this does in fact occur. Therefore, it is natural to separate the total risk estimate formula into its constituent parts to recover the individual "ordered pairs" so that they can be plotted as an exceedance frequency curve.

The construction of these exceedance frequency curves is critical to the Cassini risk assessment because the deconvolution algorithm makes use of these curves in order to obtain error bounds for the risk exceedance frequency curves.

4.3 Summary

This section has shown how the mathematical concepts of Latin hypercube sampling, importance sampling, and stratified sampling have been used to generate total risk estimates and exceedance frequency curves for the Cassini risk assessment. This discussion has *assumed* that the release scenarios and weather data can be grouped in such a manner as to make the importance sampling and stratified sampling processes appropriate. More detailed descriptions of the methods used to perform these groupings are contained in the following sections.

This page intentionally left blank.

5 Grouping of Release Scenarios

The importance sampling method described in the previous section for release scenarios is based upon a three-part step importance function. This importance function allows one to perform uniform sampling within each of the three regions prescribed by the step function, thus enabling the sampling to be performed using standard Monte Carlo or Latin hypercube techniques. This section provides a brief description of the method used to “bin” the release scenarios into the three prescribed regions. A more detailed description of this process can be found in the FSAR.¹

5.1 Reasons for Grouping

The release scenarios were generated by running the LASEP-T Monte Carlo simulation software to simulate a large number of possible accident scenarios for a given accident case. A typical case was run for 10,000 simulation trials, although more than 175,000 simulation trials were performed for one case.¹ Each simulation trial may or may not result in a radiological release, and a single LASEP-T run might identify between 1500 and 11,000 scenarios that lead to release. For each release scenario, LASEP-T computes the mass of the expected radiological release as well as its characteristics (location, particle size distribution, etc.). The mass of fuel released can vary by several orders of magnitude, and the vast majority of the release frequency corresponds to low-mass release scenarios. Clearly a pure Monte Carlo sampling of this distribution is likely to neglect the few larger releases that were found by LASEP-T. This is the reason for the importance sampling scheme and the grouping of release scenarios that was performed for the Cassini risk analysis.

5.2 Grouping Rationale

As previously mentioned, the Cassini risk analysis used three “bins” of unequal size based upon the mass of radioactive fuel released. Recall that a good importance function should, to the degree possible, be proportional to the value of the final outcome that is sought. In our case, we are seeking the consequences associated with each release scenario. Of the many release characteristics computed by LASEP-T, the mass of fuel released is arguably the single characteristic that comes closest to fulfilling this need. For this reason, each release scenario was scored based upon its total mass release in order to determine its likely importance to the final risk computation, and those release scenarios with the lowest scores were associated with the first section of the importance function, and so forth. The number of release scenarios in each section of the importance function (the size of each release bin) was set using an *ad hoc* procedure that will be described below. Note, however, that the sole factor used to score the scenarios was the mass of fuel released. The Cassini risk analysis team did not attempt to find a function that might more accurately relate scenario definition to consequences because of project schedule constraints and because it was believed that the total mass of fuel released is already a sufficient predictor of consequences for the purposes of this analysis (the next strongest contributor to consequences would likely be weather, which was sampled separately because it is viewed as independent of the release scenario for the purposes of this analysis).

A major goal of the source term binning process was to provide a mechanism by which the low-probability, high mass release scenarios would be intentionally sampled. This was done in order to ensure that these scenarios are not neglected from the overall risk results because of insufficient sampling. Binning is carried out by constructing a source term CCDF and dividing its probability range into three parts. To do this, the analysis team took the ratio of the largest complementary cumulative probability, $p_{\max}$, to the smallest such probability $p_{\min}$. This ratio is denoted as r . The dividing points between the three bins are then selected as $p_{1-2} = p_{\min} \times r^{1/3}$ and $p_{2-3} = p_{\min} \times r^{2/3}$. This has the effect of dividing the log-log CCDF into three equal vertical ranges.* In linear space, this often results in the selection of a few tens of release scenarios for the high-mass release cluster, a few hundred in the moderate-mass release cluster, and the remainder — possibly a few thousand release scenarios — in the low-mass release cluster. In this way, we assure that the low probability but highly important larger releases are indeed thoroughly sampled, while the numerous low-release bins (which are of minimal importance to aggregated risk) are sampled only to the extent necessary.

5.3 Characteristics of Release Scenario Clusters

This clustering method produces results that are characterized by a few tens of release scenarios in the risk-important high-mass release cluster. This means that, for an analysis using 30 observations,[†] some of the accident cases will be expected to compute consequences for every release scenario in the high-mass cluster. For a 50-observation analysis, most accident cases will consider every member of the high-mass cluster, and some will likely be considered multiple times (i.e., for different sets of weather conditions).[‡] This helps ensure that the risk results computed for the Cassini mission accurately represent the low-probability, high-consequence release scenarios that are so important in the aggregation of total mission risk while not neglecting the less important small releases that dominate the high-frequency portion of the total risk CCDF.

* In the special case where $r > 10^6$, $p_{\min}$ is replaced in these formulae by $(10^{-6} \times p_{\max})$ in order to prevent the largest mass bin from consisting only of very low-probability releases.¹

[†] The variability analysis used 30 observations, while 50 observations were used in the uncertainty analysis.

[‡] Each selected release scenario will, because of the nature of the sampling technique described in the previous section, be evaluated for ten weather scenarios — one from each of the ten weather clusters. A release scenario that is considered “multiple times” is actually evaluated for some multiple of ten weather scenarios. Thus, one need not worry about the pathological condition where a particularly large release only gets paired with a particularly benign weather scenario. The sampling method used in the Cassini analysis ensures that this does not happen.

6 Grouping of Weather Data

An important part of the sampling method described previously was the stratification of the u individual weather scenarios into L subgroups. While the details of that stratification are beyond the scope of this paper, it is important for the reader to have a basic understanding of how this stratification was performed. This will engender a better understanding of the practical sampling performed in the actual analysis. This section provides a brief description of the methods and rationale used to perform that stratification. A more detailed description of this process can be found in the FSAR.¹

6.1 Reasons for Grouping

Meteorological data was collected at a variety of locations near the Kennedy Space Center in Florida for October and November over a period of several years in support of the Cassini launch. October and November were chosen since the primary launch window for the Cassini spacecraft spanned those months. The result was a set of 150 weather scenarios that both fell within the launch windows and satisfied launch meteorological requirements. Thus, the value of u from Section 3 is 150. Since it is not possible to perform every release computation for each of these 150 unique weather conditions, the analysis team decided to divide the 150 scenarios into approximately 10 groups (strata). The objectives of the stratification were (1) to group weather scenarios that have similar meteorological characteristics for more efficient sampling (i.e., reduction of sampling variance), and (2) to produce a small, possibly heterogeneous stratum that contained only those weather scenarios that are likely to produce the highest consequences in order to ensure that they are not inadvertently neglected during the sample selection process (this would produce an unrealistically low estimate for the final mission risk).

6.2 Grouping Rationale

A two-phase approach was used to stratify the weather data. The first phase separated the 150 weather scenarios into nine categories based upon observed wind characteristics. The analysis team first divided these scenarios based upon whether the time-averaged wind direction was offshore at the time of launch and, if so, whether there were portions of the axial wind profile directed onshore either within the launch window or within the 8 hours that follow the launch window. This rationale allowed the offshore-wind scenarios to be grouped into three categories. The remaining days with on-shore time-averaged winds were grouped into five additional categories based upon the time-averaged wind direction (with particular attention being paid to weather scenarios that might direct a release toward a populated area). A ninth category was established for scenarios with offshore average winds but onshore surface winds.

The second phase of the stratification process involved the use of actual consequence computations for each of the 150 weather scenarios using four postulated radiological source terms. These computations, called the "sensitivity runs" in the FSAR, were used to examine the homogeneity of the consequences produced within each of the nine weather categories. While some of the categories produced relatively homogeneous consequence results, the results in others varied widely. In order to reduce the variation within the weather strata, the weather

categories were modified based on the sensitivity runs to obtain the final weather clusters that were sampled in the Cassini risk analysis. Specifically, the modifications consisted of (1) combining two adjacent categories that each had few members and produced similar consequences, and (2) defining two new nondirectional clusters to reduce the amount of variation seen within the remaining clusters. One of these new clusters was a "minimum consequence" cluster and the other a "maximum consequence" cluster. The maximum consequence cluster was populated by grouping the eight highest consequence weather scenarios regardless of their observed meteorological characteristics or initial category. The minimum consequence category was similarly populated with the eight lowest consequence scenarios over the entire set of 150 candidates. This "removal of outliers" from the initial eight clusters improved their homogeneity, and their segregation into a new small stratum ensured that these unique cases would be appropriately represented in the final risk results. Note that these outlier strata were small enough that each scenario would be expected to be selected at least three times in a 30-observation Monte Carlo sample.* Had the outliers not been treated separately, the sampling process might have missed their contribution to risk altogether because many of them were taken from a category that originally contained more than 60 scenarios.

6.3 Characteristics of Weather Clusters

The ten final weather clusters consisted of eight clusters that were based primarily upon wind direction, plus the maximum- and minimum-consequence clusters. Half of the clusters contained fewer than 10 weather scenarios, while the remaining clusters contained 11, 12, 13, 25, and 56 scenarios, respectively. Thus, in a Monte Carlo sample with 30 observations, one would expect that each scenario would be evaluated at least twice in 8 of the 10 clusters, and one of the remaining two clusters would likely be completely sampled once. In a 50-observation analysis, even the largest cluster would enjoy almost complete sampling under this stratification. Since the outliers are placed into separate strata, one would expect that the sampling of this stratified data would represent the actual characteristics of the weather data far more accurately than an equal number of samples drawn randomly from the entire weather population.

* The variability analysis used 30 observations, while 50 observations were used in the uncertainty analysis.

7 Implementation of Sampling Strategy

To this point we have examined the mathematical foundations for the sampling method employed in the Cassini risk analysis as well as the practical methods used to group the release scenarios and weather data for sampling. We have also briefly discussed the mathematical deconvolution method that was used to examine the random variability and epistemic uncertainty of the risks associated with the mission. In this section we examine how the sampling scheme described in Section 4 was used with the release scenario clusters developed in Section 5 and the weather clusters developed in Section 6 to achieve the deconvolution analysis described in Section 2.6.3.

7.1 Overview

A detailed description of the deconvolution uncertainty analysis method is beyond the scope of this paper.^{1,12,13} Recall from Section 2.6.3, however, that, as the basis for the deconvolution analysis, we assumed that there exists a distribution function U which, when convolved with the detailed variability distribution V , produces as its result the combined uncertainty-variability risk distribution R . Mathematically this was written as

$$R = U * V.$$

Recall that the convolution operation becomes a multiplication operation under Laplace and Fourier transforms. In the Cassini analysis, we can construct a distribution for V by sampling all aspects of the risk assessment that represent random variability while holding those aspects that are associated with epistemic uncertainty constant at their nominal values. We can also construct a distribution for R by sampling all aspects of the risk assessment that represent either random variability or epistemic uncertainty *simultaneously*. We seek the unknown distribution U through the deconvolution of R and V (both of which are known). This is done by dividing their Laplace or Fourier transforms as

$$\bar{U} = \frac{\bar{R}}{\bar{V}}$$

where the bar denotes that the relation holds in the transform domain. The remaining task, then, is to invert the Laplace or Fourier transform for $\bar{U}$ back into the real domain U . The distribution U now represents only the effect of epistemic uncertainty on the variability distribution. This distribution can be used to obtain a family of curves that represent the combined uncertainty-variability distribution, which is the objective of our analysis. Thus, the Cassini risk assessment study applied the formulae developed in Section 4 in two steps: once during the variability-only computations (to obtain V), and once during the fully integrated variability-uncertainty computations (to obtain R).

Note that we are interested in obtaining the representations for *functions* V and R (here assumed to be exceedance frequency curves, as computed by the method of Section 4.2), and not in the actual individual sampled values that make them up *per se*. For this reason, it is acceptable to

perform the sampling for these two analyses using different methods or different numbers of computations as long as the integrity of the final functions V and R is maintained. For the Cassini analysis, the same sampling strategy was used during both computations, but 30 observations were used to determine V while 50 observations were used to determine R . The additional observations were used because more distributions were being sampled in this analysis.

7.2 Variability Runs

Recall that the total risk computation prescribed in Section 4.1 was performed using a formula that can be rewritten as:

$$R \approx \sum_{k=1}^v \sum_{p=1}^L \sum_{z=1}^3 \frac{f_{DB}}{t \cdot v \cdot u} \cdot M_z \cdot N_p \cdot C_{k,p,z}$$

Recall that the exceedance frequency curves were constructed from a variation of this same formula. Note also that the first summation is over all Monte Carlo observations. Thus, each Monte Carlo observation actually prescribes $3 \cdot L$ actual consequence computations because of the inner sums. Recall that L , the number of weather clusters, is 10 for the Cassini study, so each Monte Carlo observation results in 30 actual consequence computations. These come about because, for each observation, we select one member at random from each of the 10 weather clusters, and one member at random from each of the 3 release clusters, and perform consequence computations for all possible combinations thereof. Thus, 13 random variables are required to fully prescribe these 30 consequence computations. This is true for both the variability and the uncertainty runs.

The variability analysis then used Sandia's LHS software⁷ to sample these 13 random variables plus 1 additional input to the SPARRC model (the time of the launch accident, to give a total of 14 random variables). Those random variables that were determined to primarily represent epistemic uncertainty rather than random variability were held constant at their nominal values during the variability runs and were not sampled. The analysis used 30 observations to model the random risk variability for on-pad/early launch accidents, and each observation prescribed 30 consequence computations, so a total of 900 consequence computations were performed to characterize the V in this analysis.

It should be noted that V was determined separately for several scenarios in each phase of on-pad/early launch operations. Thus, several thousand consequence computations were performed during the on-pad/early launch risk variability characterization for the Cassini mission.

7.3 Uncertainty Runs

The uncertainty analysis runs were performed using the same sampling strategy as the variability runs. The only difference is that here all random variables were sampled simultaneously regardless of whether they were determined to represent random variability or epistemic uncertainty. This introduced 22 additional random variables into the problem (principally inputs to the SPARRC model), for a total of 36 random variables to be sampled. The analysis used 50

observations to model the risk uncertainty for on-pad/early launch accidents, and each observation prescribed 30 consequence computations, so a total of 1,500 consequence computations were performed to characterize the R in this analysis. In addition, just as V was determined separately for several scenarios in each phase of on-pad/early launch operations, separate R computations were also performed for the same sets of scenarios. Thus, several thousand consequence computations were performed during the on-pad/early launch risk uncertainty characterization for the Cassini mission.

7.4 Comparison With Pure Monte Carlo Sampling

After the Cassini spacecraft was launched, the Lockheed Martin Co., at the request of the Department of Energy, undertook a study to determine the impact of using fuel release and weather-day clusters and Latin hypercube sampling on the Cassini analysis results. The study examined case 1.1z with no CIR using 4000 Monte Carlos samples, and compared it with the results obtained during the original study, which had used 900 consequence computations that were obtained using the sampling methods described in this document. In the Monte Carlo study, fuel release and weather-day clustering were removed, and all consequence computations were specified using pure Monte Carlo sampling.²⁵

The results of that study are noteworthy for two reasons. First, the results generated by the pure Monte Carlo computation were in every case both qualitatively and quantitatively equivalent to those generated for the original Cassini study. Deviations between the CCDFs generated by the two studies are easily explained by random sampling deviations. However, the CCDF results from the Monte Carlo computation end sooner than those generated for the original study. That is, the CCDF generated for the original study shows accident sequence frequencies (and their associated consequences) that are more than an order of magnitude more rare than those generated by pure Monte Carlo sampling. It is apparent, then, that the sampling methods used for the original Cassini study produce results that are not only comparable to pure Monte Carlo results (in the region where pure Monte Carlo predicts *any* results), but are also better than the pure Monte Carlo results because they capture more of the tail of the CCDF that is neglected by the pure Monte Carlo computation. Furthermore, the original Cassini study produced these improved results using 78 percent fewer samples than the pure Monte Carlo computation. Given the extremely large computational burden posed by the original Cassini study, it would have been virtually impossible to carry out the computations that would have been required for this pure Monte Carlo analysis over all cases and phases of the original Cassini analysis.

7.5 Summary

The sampling methodology that was developed for the Cassini risk uncertainty study was specifically designed to ensure that the tails of the risk distribution were not neglected. The program specifically wanted to avoid any appearance that fuel releases and weather scenarios that might have a low probability and yet cause a very large consequence were in any way neglected in the consequence computations. This is mathematically equivalent to ensuring that the tails of the CCDF for the overall risk are adequately captured in the sampling strategy. The results of the Monte Carlo comparison study indicate that this objective was accomplished very well.

A second objective for the sampling methodology was to ensure that the risk computations could be accomplished using a minimum number of random samples (and, hence, a minimum number of computationally expensive consequence calculations). The fact that the original Cassini risk study produced better results using 78 percent fewer samples than a pure Monte Carlo computation indicates that this objective for the sampling strategy was also achieved.

8 Summary and Conclusions

This paper has described the state-of-the-art quantitative uncertainty assessment methods that were developed for the Cassini risk analysis in a collaborative research effort between Lockheed Martin Co., Sandia National Laboratories, and the U.S. Department of Energy. These were applied to several accident cases that can be characterized as on-pad/early launch accidents without coincident impact release.* The methods make use of importance sampling methods, stratified sampling methods, and Latin hypercube sampling to characterize the risks associated with the launch of the Cassini spacecraft. Sampling methods were required because it was not computationally feasible to compute consequences for every postulated radiological release under every set of observed weather conditions.

The methods described here were developed under the explicit assumption that all postulated release scenarios are equally likely for a given accident case — a constraint that is consistent with the characteristics of the LASEP-T software that developed those scenarios. Further development will be required in order to formally extend these methods to consider situations where all release scenarios are not equally likely. However, the methods have proven robust and accurate in a number of sensitivity computations that were performed subsequent to the Cassini launch.²⁵ The methods take great care to ensure that the statistical sampling processes employed do not neglect the low-probability high-consequence cases that are often major contributors to the overall aggregated risk. The utility and accuracy of these methods will make them useful risk assessment tools for many years to come.

* Suborbital/orbital reentry cases were evaluated using a very similar method, although the effects of weather were handled in a different and simpler manner because the consequence models for such releases are less geography-specific owing to the world-wide nature of their computations. On-pad/early launch accidents with CIR were modeled using a slightly different formalism that was developed by Lockheed Martin independent of Sandia collaboration. Therefore, the methods used to model CIR cases are beyond the scope of this paper.

This page intentionally left blank.

9 References

- ¹ Lockheed Martin Missiles and Space, "GPHS-RTGs in Support of the Cassini Mission: Final Safety Analysis Report (FSAR)," Lockheed Martin CDRL C.3, Lockheed Martin Missiles and Space, Valley Forge Operations, Valley Forge, PA, June 1997.
- ² Lockheed Martin Astro Space, "Cassini Titan IV/Centaur RTG Safety Databook," Rev. B, NAS3-00031, Prepared for NASA Lewis Research Laboratory, Cleveland, OH, March 1997.
- ³ NUS Corporation, "Final Safety Analysis Report for the Ulysses Mission," ULS-FSAR-006, Prepared for the U.S. Department of Energy Office of Special Applications, Washington, DC, 1990.
- ⁴ Dobranich, D., Powers, D.A., and Harper, F.T., "The Fireball Integrated Code Package," SAND97-1585, Sandia National Laboratories, Albuquerque, New Mexico, July 1997.
- ⁵ International Commission on Radiation Protection, "Limits for Intakes of Radionuclides by Workers," ICRP Publication 30, Pergamon Press, Elmsford, New York, 1979.
- ⁶ Iman, R.L., and Shortencarier, M.J., "A FORTRAN 77 Program and User's Guide for the Generation of Latin Hypercube and Random Samples for Use With Computer Models," NUREG/CR-3624, Prepared by Sandia National Laboratories for the U.S. Nuclear Regulatory Commission, Washington, DC, 1984.
- ⁷ Wyss, G.D., and Jorgensen, K.H., "A User's Guide to LHS: Sandia's Latin Hypercube Sampling Software," SAND98-0210, Sandia National Laboratories, Albuquerque, NM, February 1998.
- ⁸ U.S. Nuclear Regulatory Commission (USNRC), "Reactor Safety Study — An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants," WASH-1400 (NUREG-75/014), Washington, DC, 1975.
- ⁹ U.S. Nuclear Regulatory Commission (USNRC), "Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants," NUREG-1150, Washington, DC, 1990.
- ¹⁰ Griesmeyer, J.M., and Smith, L.N., "A Reference Manual for the Event Progression Analysis Code (EVNTRE)," NUREG/CR-5174, Prepared by Sandia National Laboratories for the U.S. Nuclear Regulatory Commission, Washington, DC, 1989.
- ¹¹ Interagency Nuclear Safety Review Panel, "Uncertainty Analysis Report for Ulysses," INSRP 90-07, Washington, DC, 1990.
- ¹² Kampas, F.J., and Loughin, S., "Deconvolution of the Variability and Uncertainty in the Cassini Safety Analysis," in *Proceedings of the 15th Symposium on Space Nuclear Power and Propulsion*, University of New Mexico Institute for Space and Nuclear Power Studies, Albuquerque, NM, 1998.

- ¹³ Kampas, F.J., "Cassini Memo No. 568: Nominal Parameters in Uncertainty Analysis," Lockheed Martin Astronautics, Valley Forge, PA, March 17, 1998.
- ¹⁴ McKay, M.D., Conover, W.J., and Beckman, R.J., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, 221, 239-245, 1979.
- ¹⁵ Steck, G.P., Iman, R.L., and Dahlgren, D.A., "Probabilistic Analysis of LOCA, Annual Report for FY1976," Technical Report SAND76-0535 (NUREG76-6513), Sandia National Laboratories, Albuquerque, NM, 1976.
- ¹⁶ Iman, R.L., Helton, J.C., and Campbell, J.E., "An Approach to Sensitivity Analysis of Computer Models, Part 1. Introduction, Input Variable Selection and Preliminary Variable Assessment," *Journal of Quality Technology*, 13(3), 174-183, 1981.
- ¹⁷ Iman, R.L., Helton, J.C., and Campbell, J.E., "An Approach to Sensitivity Analysis of Computer Models, Part 2. Ranking of Input Variables, Response Surface Validation, Distribution Effect and Technique Synopsis," *Journal of Quality Technology*, 13(4), 232-240, 1981.
- ¹⁸ Iman, R.L., and Conover, W.J., "Sensitivity Analysis Techniques: Self-Teaching Curriculum," Nuclear Regulatory Commission Report, NUREG/CR-2350, Technical Report SAND81-1978, Sandia National Laboratories, Albuquerque, NM, 1982.
- ¹⁹ Iman, R.L., and Helton, J.C., "A Comparison of Uncertainty and Sensitivity Analysis Techniques for Computer Models." Technical Report SAND84-1461, Sandia National Laboratories, Albuquerque, NM, 1985.
- ²⁰ Iman, R.L., and Conover, W.J., "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables," *Communications in Statistics*, B11(3), 311-334, 1982.
- ²¹ Marquardt, D.W. and Snee, R.D., "Ridge Regression in Practice," *American Statistician*, 29(1), 3-20, 1975.
- ²² Marquardt, D.W., "Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation," *Technometrics*, 12(3), 591-612, 1970.
- ²³ Press, W.H., et al., *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, 1992.
- ²⁴ Cochran, W.G., *Sampling Techniques*, 2nd ed., Wiley, New York, 1963.
- ²⁵ Deane, Nelson A., "Cassini Memo No. 591: Monte Carlo Sampled Case 1.1 with No CIR Analysis Without Fuel and Weather Day Clusters," Lockheed Martin Astronautics, San Jose, CA, August 14, 1998.

Distribution

External Distribution

5	Nelson Deane
Lockheed Martin Missiles & Space
2770 De la Cruz Blvd
Santa Cruz, CA 95050	1	Bart Bartrum
Tetra Tech NUS
Suite 400
910 Clopper Rd.
Gaithersburg, MD 20878
1	Chuong Ha
Lockheed Martin Missiles & Space
2770 De la Cruz Blvd
Santa Cruz, CA 95050	1	Earl Wahlquist
U.S. Department of Energy
Germantown Building, NE-50
19901 Germantown Rd.
Germantown, MD 20874
1	Larry DeFillipo
Lockheed Martin Missiles & Space
Rm. U8620, Bldg. 100
P.O. Box 8555
Philadelphia, PA 19101	1	Rich Furlong
U.S. Department of Energy
Germantown Building, NE-50
19901 Germantown Rd.
Germantown, MD 20874
1	Frank Kampas
Lockheed Martin Missiles & Space
Rm. U8620, Bldg. 100
P.O. Box 8555
Philadelphia, PA 19101	1	Art Mehner
U.S. Department of Energy
Germantown Building, NE-50
19901 Germantown Rd.
Germantown, MD 20874
1	Jack Chan
Lockheed Martin Missiles & Space
Rm. U8620, Bldg. 100
P.O. Box 8555
Philadelphia, PA 19101	1	Don Owings
U.S. Department of Energy
Germantown Building, NE-50
19901 Germantown Rd.
Germantown, MD 20874
1	Hank Firstenberg
Tetra Tech NUS
Suite 400
910 Clopper Rd.
Gaithersburg, MD 20878	1	Lyle Rutger
U.S. Department of Energy
Germantown Building, NE-50
19901 Germantown Rd.
Germantown, MD 20874

1	Jack Wheeler
U.S. Department of Energy
Germantown Building, NE-50
19901 Germantown Rd.
Germantown, MD 20874	1	Yale Chang, 1E-260
The Johns Hopkins University
Applied Physics Laboratory
11100 Johns Hopkins Rd.
Laurel, MD 20723-6099
1	F. Eric Haskin
ERI Consulting
901 Brazos Place SE
Albuquerque, NM 87123	1	Orbital Sciences Corp.
Attn: Cassini Project Office
2031 Century Blvd.
Germantown, MD 20874-1110
1	Michael V. Frank
Safety Factor Associates
4401 Manchester Ave., Suite 106
Encinitas, CA 92024	1	Dr. Louis Goossens
Delft University of Technology
TBM/VK
P.O. Box 5050
NL-2600 GB Delft
The Netherlands
1	Capt. Ed Jakes, USAF
HQ AFSC/SEWE
9700 Avenue G, Bldg. 2499, Suite 120
Kirtland AFB, NM 87117-5670	1	Bernt Kraan
Delft University of Technology
Mekelweg 4
Room 5.050
2627 CA Delft
The Netherlands
1	Gareth Parry
U.S. Nuclear Regulatory Commission
NRR/DSSA
11555 Rockville Pike, MS O-10A1
Rockville, MD 20852-2738		
1	Nathan Siu
Probabilistic Risk Analysis Branch
MS 10E50
US Nuclear Regulatory Commission
Washington, DC 20555-0001		

Sandia Distribution

1	MS 0191	3010	J.J. Gregory
1	MS 0405	12333	Todd Jones
1	MS 0405	12333	Kevin Maloney
1	MS 0428	12300	David D. Carlson
1	MS 0428	12301	William H. McCulloch
1	MS 0449	6234	D.R. Gallup
1	MS 0736	6400	Thomas Blejwas
1	MS 0739	6415	John Guth
1	MS 0744	6420	D.L. Berry

1	MS 0746	6411	Dennis Anderson
1	MS 0746	6411	Robert Cranwell
1	MS 0746	6411	Laura Swiler
1	MS 0747	6410	A. L. Camp
1	MS 0747	6410	Roger Cox
5	MS 0747	6410	Vince Dandini
1	MS 0747	6410	Felicia Duran
1	MS 0747	6410	Jeff LaChance
1	MS 0747	6410	Teresa Sype
12	MS 0747	6410	Gregory Wyss
1	MS 0748	6413	David Robinson
1	MS 0748	6413	Bob Waters
5	MS 0748	6413	Tim Wheeler
1	MS 0748	6413	Donnie Whitehead
1	MS 0767	5317	Frederick T. Harper
1	MS 0767	5317	Bruce Boughton
1	MS 0828	9133	Martin Pilch
1	MS 0828	9133	William L. Oberkampf
1	MS 0829	12323	Floyd Spencer
1	MS 0830	12335	Kathleen V. Diegert
1	MS 0977	6534	Sharon M. DeLand
1	MS 1136	6424	Paul Pickard
1	MS 0918	8940-2	Central Technical Files
2	MS 0899	9616	Technical Library
1	MS 0612	9612	Review and Approval Desk for DOE/OSTI