Prediction Techniques for Passive Systems’ Probability of Failure

by

Edoardo Cavalieri d’Oro

Laurea in Ingegneria Nucleare, Politecnico di Milano, Italy (2002)

Submitted to the Department of Nuclear Engineering in Partial Fulfillment of the Requirements for the Degree of

MASTER in NUCLEAR SCIENCE and ENGINEERING

at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2007

©2007 Massachusetts Institute of Technology, all rights reserved
TABLE OF CONTENTS

TABLE OF CONTENTS .............................................................................................................................. 1
TABLE OF FIGURES ................................................................................................................................. 3
LIST OF TABLES ........................................................................................................................................ 5
Abstract .................................................................................................................................................. 7
Acknowledgments ................................................................................................................................... 9
NOMENCLATURE .................................................................................................................................... 10

1. INTRODUCTION .............................................................................................................................. 11
   Organization of the Work ..................................................................................................................... 11
   1.1 The Gas Fast Reactor Plant ............................................................................................................. 12
   1.2 Plant Modeling ................................................................................................................................ 14
      1.2.1 The GFR Core ................................................................................................................................. 15
      Decay Heat Curves ................................................................................................................................. 17
      Core Heat Capacity ................................................................................................................................. 19
      Core Hot Spot and Core Average Specific Power .................................................................................. 20
      1.2.1.1 Calculation of Temperatures in Cladding and Fuel for a Lumped Core .................................. 21
      Overall Heat Transfer Coefficient ........................................................................................................ 22
      Thermal Conductivity in the Fuel .......................................................................................................... 22
      Thermal Conductivity in the Cladding ................................................................................................... 24
      Thermal conductivity across the Fuel-Clad Gap .................................................................................... 24
      Geometrical Lumping Parameters ....................................................................................................... 25
      1.2.2 The Decay Heat Removal System ............................................................................................... 27
      1.2.3 The Selected Time-Frame ........................................................................................................... 28
      1.2.4 The Reference Scenario ............................................................................................................. 30
   1.3 Modeling Techniques ...................................................................................................................... 32
      1.3.1 Systems Dynamics Modeling ......................................................................................................... 32
      System Dynamics to Model Transients in T-H Plants ......................................................................... 32
      Lumped Modeling versus Traditional T-H Codes ................................................................................ 32
      Respect of SD Software Capabilities during the Selected Simulation Time-Frame ........................... 33
      Improvements in the Computational Capability to Explore Transients ............................................. 33
      Other General Reasons ....................................................................................................................... 34
1.3.2 Supervised Learning Modeling

2. THE METHODOLOGY

2.1 About Passive Systems

2.2 Available Resources

2.3 Possible Applications and Main Objectives

2.4 The Methodology

2.5 Tools and Computational Methods

2.5.1 Classification Techniques

Classification Trees

Logistic Regression Classifiers

2.5.2 Latin Hypercube for Sensitivity Analyses

2.5.3 Statistical Indicators

P-value

Wald Statistics

2.5.4 The Software Used in the Simulations

2.5.5 Data Interpretation and Final Remarks

3. RESULTS

EXAMPLE I: Core Failure of the GFR as due to a Negative Heat Pulse

The model

The Accident and the Computational Time Required Running the Scenarios

Uncertainties Quantification

Resulting Output from the Dynamic Model

Creating the Dataset for the Prediction Models

Results from the Classification Models

Analysis of the Scatter Plots

Findings

EXAMPLE II: Leakage from a Check Valve in a Helium Cooled GFR

The Model

The Selected Accident

Failure Modes and Sampling Techniques

Results from the Classification Models

Findings

EXAMPLE III: Explanatory Example about LHS
TABLE OF FIGURES

Figure 1: 1200 MWe direct stacked cycle isometric and top view views .......................................................... 12
Figure 2: Loop diagram of the GFR plant. The diagram highlights the different loops generated by a 5 nodes representation of the plant ................................................................. 14
Figure 3: Results from the optimization on the dataset of the Super Phoenix reactor compared to a LWR. ................................................................. 18
Figure 4: Schematization and fundamental geometrical parameters for the core channel ............................ 21
Figure 5: Estimated thermal conductivity of BeO/(U,TRU)O2 fuel. Source of the figure: M. Pope, 2006 page 59 [1] .................................................................................................................. 23
Figure 6: Reference case of the fuel temperature, as given by the GFRDYN_LUMPED Model ........... 31
Figure 7: Reference case simulations of the cladding and fuel margins. Blue line represents the temperature transient while the red straight lines show the limits imposed by the materials. ............ 31
Figure 8: Sketch illustrates the main variables and parameters modeled by the GFRDYN Model ............. 34
Figure 9: Roadmap for the evaluation of parameters in a PSR contest ...................................................... 43
Figure 10: Visual of the scheme used to create the dataset which is used as input for the supervising models ............................................................................................................................................ 46
Figure 11: Vensim Sketch of the molecule describing the GFR core under the selected transient ........ 55
Figure 12: Temperature of the GFR core during a LOCA as obtained by the core's sub-model. Reference Case (bottom red line) and Forced Pulse Heat Case (upper blue peaked line) ................................................................. 56
Figure 13: Input dataset used to model the HTC [Core], Cp and Initial temperature in the fuel .............. 57
Figure 14: Temperature of the cladding in 200 realizations and deterministic cladding limits. ..........58
Figure 15: Output distributions of the fuel temperature for the average and hottest channels of the GFR. 58
Figure 16: Classification tree for the dependent variable temperature of the cladding as a function of the
three independent variables HTC, Cp, Initial Temp of the Fuel. .........................................................62
Figure 17: Scatter plot of the Specific heat for CO2 versus Initial Temperature of the fuel. ............63
Figure 18: Scatter plot of the specific heat for CO2 versus Heat Transfer Coefficient in the core. ....64
Figure 19: Scatter plot of the initial temperature of the fuel versus the heat transfer coefficient in the core.
......................................................................................................................................................64
Figure 20: Realizations for the Maximum Temperature in the Hot Leg of DHR Loops [F. MacKay et al.,
2007]. ...............................................................................................................................................68
Figure 21: Classification tree resulting from LHS sampling of six parameters in the Helium GFR. ....70
Figure 22: Scatter plot of the heat transfer coefficient versus Maximum Leakage for the Helium GFR ..71
Figure 23: Measured PF deltas on the first 400 trials compared to the whole 10,000 sequence. ..........74
Figure 24: PF of the cladding for 2000 simulations. Data are collected every 20 trials. .................75
Figure 25: PF of the cladding for 10,000 simulations. Data is collected every 500 trials. ...............76
Figure 26: Shape Distribution Effects on linear and non linear models as a function of the number of
samplings, M .......................................................................................................................................80
Figure 27: Load distribution and Capacity for the linear and non linear cases (left and right graphs) for
M=10,000 ...................................................................................................................................................81
Figure 28: Schematic of the GFR showing the plant subjected to multiple failures........................84
Figure 29: Agreement in prediction as a function of the number of trials .......................................87
Figure 30: Scatter-plot of the failures in the core for a number of samplings M = 60 ..................88
Figure 31: Scatter-plots of response and failures as a function of the Number of Realizations M. GFR’s
core Model (Example I) for 400, 2000 and 10,000 realizations. ......................................................89
Figure 32: Scatter-plots of response and failures as a function of the Number of Realizations M. GFR’s
core Model (Example II). Threshold lines are indicated in the graphs .............................................90
# LIST OF TABLES

Table 1: Status of GFR Plant Features as of July 2007 ................................................................. 13
Table 2: Fraction of nominal power for the Super-Phenix reactor. Source: NT-SERI/LCSI/03/4005 ...... 17
Table 3: Confidence intervals on the parameter estimation of the ANS Curve for the Super Phoenix Reactor ............................................................................................................. 18
Table 4: Tube-In-Duct (TID) Type Core (11ringrev2) .................................................................... 19
Table 5: Parameters and geometrical dimensions used for the GFRDYN Model ............................... 26
Table 6: Input Parameter Distribution Table .................................................................................... 57
Table 7: Dataset matrix used as an input for the classifiers ............................................................... 59
Table 8: Summary Statistics from the Logistic Classifier ................................................................. 61
Table 9: Classification Agreement Table ........................................................................................ 61
Table 10: Selected Input Parameters and their probability distributions [adopted from F. MacKay et al., 2007] .................................................................................................................. 66
Table 11: Summary Statistics from the Logistic Classifier .............................................................. 68
Table 12: Classification Agreement Table ....................................................................................... 69
Table 13: Probabilities of failure, PF, obtained from 140 Latin Hypercube samplings ....................... 77
Table 14: Input datasets and Models used for the calculations provided in Example III .................... 78
Table 15: Distribution parameters for two different input datasets: uniform versus non-uniform ........ 79
Table 16: Probability of failures as a function of M for the four different cases ....................... 82
Table 17: Total PF for a two component system. Linear versus non-Linear for two different input sets and seeds ........................................................................................................... 82
Table 18: CT thresholds for HTC Core and Summary p-tests for LC ............................................. 87
Out of Chaos Brilliant Stars are Born...

I-Ching, Hexagram #3
Abstract

This work fits into the wider framework of the on-going debate centered on Passive System reliability. Its aim is to provide insights into the design of a dependable method to evaluate the reliability of Passive Systems.

In order to achieve this, a method is proposed that focuses on the identification of the fundamental parameters that are critical in leading the system to failure. The selection of these parameters was done through the use of Latin Hypercube Sampling (LHS) combined with an analysis centered on the use of two statistical tools, Logistic Regression and the Classification Tree. The results yielded by this study, made it necessary to perform a systematic statistical evaluation of the efficiency of the LHS when used in the context of sensitivity analyses. The study was conducted via the visual and statistical investigation of the scatter-plots derived from the propagation of the uncertainties associated with the fundamental parameters of the plant.

In order to validate the proposed method, two examples involving a Gas Fast Reactor (GFR) plant have been set up. The two examples differ, among other aspects, in the number of realizations, M, used to carry out the analyses.

The first example - used to illustrate the method - is a representation of the core derived from the application of System Dynamics modeling.
The second example is a RELAP5-3D model of a two-loop passive Decay Heat Removal system of the GFR. This case was designed in order to test the method in a more realistic scenario.

Important findings about the applicability of the method as a function of M, are given by way of comparison between the results obtained from the two cases. The results reveal that the numbers of realizations, provided by LHS, are insufficient when used to predict and interpret the propagation of the failures in the plant. The second important conclusion is that the resulting Probability of Failure (PF), for low values of M, does not converge to an accurate estimate. The implications of these findings are investigated through a third study.

The third example is a purely mathematical model specifically designed to test the assumptions made for the first two cases. It provides additional analysis on Examples I and II offering further support for the findings from the two.

The results attained by this work suggest that further studies of this kind should be conducted in this area.

THESIS SUPERVISOR: Michal W. Golay

TITLE: Professor of Nuclear Science and Engineering

THESIS READER: George E. Apostolakis

TITLE: Professor of Nuclear Science and Engineering
Acknowledgments

This thesis has been supported by a cooperative agreement between the U.S. Nuclear Regulatory Commission and MIT.

First, I would like to express my gratitude to my advisor Professor Golay, who gave the opportunity to work on this project.

I also wish to thank Prof. Apostolakis for his help and comments, Prof. Driscoll and Dr. Hezlar for his cooperation in make me understanding the technical aspects of the. Special thanks go to Professor Buongiorno for the precious ideas and comments.

I also would like to thank the colleagues and the staff of the Italian Ministry of Interiors who made the initial part of this adventure possible.

Lastly and most importantly, I could not have done any of this without my family and in particular to my mother and to her courage.
**NOMENCLATURE**

T-H = Thermal Hydraulics
CT = Classification Tree
LR = Logistic Regression
LHS = Latin Hypercube Sampling
GFR = Gas Fast Reactor
SSCP = Systems Structures Components Phenomena
RMPS = Reliability Measures for Passive Systems
PSR = Passive System’s Reliability
PF = Probability of Failure
M = Number of Realizations
Ns = Number of Systems
PS = Passive System
DHR = Deacay Heat Removal System
ECCS = Emergency Core Cooling System
HXC = Heat Exchanger
AHP = Analytical Hierarchy Process
PEPSR = Parameters Evaluation for Passive System’s Reliability

PSA = Probabilistic Safety Assessment
SM = Safety Margins
MC = Monte Carlo
PRA = Probabilistic Risk Assessment
IAEA = International Atomic Energy Agency
SD = System Dynamics
PDF = Probability Density Function
SPX = Super Phenix
LWR = Light Water Reactor
LOCA = Loss of Coolant Accident
GFRDYN_LUMPED = lumped model of the GFR during transients
PCS = Power Conversion System
FAST = Fourier Amplitude Sensitivity Test
FPI = Fast Probability Integration.
1. INTRODUCTION

Organization of the Work

This chapter presents an overview of the main topics discussed in this work. The present chapter is divided into three parts.

The first section describes the features of the plant used in Examples I and II, the former of which was specifically created in order to validate the method proposed later on in the document. The second part describes the core part of an overall model, developed with a special modeling technique named System Dynamics, and provides also details of the reference scenario analyzed with special emphasis to the passive systems utilized to remove the decay heat in the event of an accident. The third section provides general information about the different modeling techniques used to carry out the analyses performed in Chapter 3.

Chapter 2 moves the discussion towards the wider debate centered on the need for an overall method that would cover all the questions connected with passive system's reliability. After a brief overview of the supervised learning methods and their assumed applicability to reliability studies, we discuss an overall methodology to account for some of the typical problems encountered in Passive System’s reliability studies. The core of this chapter focuses mainly on investigating the use of Classification Trees in order to predict the failure probability of the plant more rigorously, and to determine a possible ranking criterion of the parameters of interest.

Chapter 3 provides the set of examples used to validate the suggested method code along with a few observations on the subject. Final conclusions, future work and suggestions on how to broaden the field of applicability to wider and more complex engineering problems are given at the end of this Chapter.

References and appendixes, containing the source codes and detailed simulation results, are presented at the end of this document.
1.1 The Gas Fast Reactor Plant

The nuclear plant that has been selected for our analyses is the Gas Fast Reactor, GFR, of 2400 MWth that is under design at MIT (see Figure 1). While the research presented in this work deals mainly with the natural circulation DHR systems, it is equally important to define and understand the complete design of the Gas-cooled Fast Reactor (GFR) design.

The reactor is planned to use the carbon dioxide as coolant and term vector fluid at a nominal pressure of 20 MPa with an inlet temperature of 480 °C during normal operations. The plant is constituted by the Emergency Core Cooling System, ECCS, the core, the vessel, the Power Conversion Units, PCUs, the Core and the Containment plus the active systems used to remove the heat in the first minutes of a severe accident. A comprehensive list of all major components of the plant is summarized by Table 1.

![Figure 1: 1200 MWe direct stacked cycle isometric and top view views.](image)

The entire system operates as one passive system whose initiation is triggered by the reactor scram after a Loss-of-Coolant Accident (LOCA).
<table>
<thead>
<tr>
<th>Subsystem</th>
<th>Features</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Core</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fuel</td>
<td>UO$_2$ + BeO</td>
<td>LWRTRU fissile ± MA</td>
</tr>
<tr>
<td>Clad</td>
<td>ODS-MA956, or HT-9</td>
<td>SiC a long range possibility</td>
</tr>
<tr>
<td>Configuration</td>
<td>tube-in-duct fuel assemblies, &quot;hexnut&quot; pellets, vented</td>
<td>pin-type core as fallback</td>
</tr>
<tr>
<td>Thermal-Hydraulics</td>
<td>axial peaking factor $\leq 1.3$ radial peaking factor $\leq 1.2$ power density $\sim 85$ W/cc</td>
<td>Vary BeO fraction to flatten power. Lower than GA GCFR of 1970's@ 235 W/cc</td>
</tr>
<tr>
<td>Burnup</td>
<td>$\geq 120$ MWe$/$kg(avg)</td>
<td>In single batch no-reshuffle core, 17-yr</td>
</tr>
<tr>
<td><strong>Safety Systems</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aux. Loops</td>
<td>combined shutdown &amp; emergency, 3 or 4 x 100% capable, active forced convection; but passive natural convection supplemented; water boiler heat sink</td>
<td>Based on RELAP parallel loop calculations. For $P \geq 0.7$ MPa natural convection alone may suffice</td>
</tr>
<tr>
<td>Emergency Power</td>
<td>Fuel cells to supplement diesels</td>
<td>Projected to be more reliable than diesels alone in long</td>
</tr>
<tr>
<td><strong>Plant</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power Conversion System (PCS)</td>
<td>supercritical CO$_2$ Brayton direct 2 x 600 MWe loops=1200 MWe; 650°C core exit/turbine inlet, pressure: 20</td>
<td>AGRs in UK use CO$_2$ coolant at 4 MPa and have T=650°C</td>
</tr>
<tr>
<td>Reactor Vessel</td>
<td>PCI V</td>
<td>Vessel houses loop isolation and check valves plus shutdown cooling heat</td>
</tr>
<tr>
<td>Containment</td>
<td>PWR type, steel liner reinforced concrete 0.7 MPa design capability 70,000 m$^3$ free volume filtered/vented</td>
<td>CO$_2$ can be added to adjust pressure. Internally insulated</td>
</tr>
<tr>
<td>H$_2$ production by steam electrolysis (optional)</td>
<td>Separate water boiler loops (4) @ 10% of reactor power Recuperation of H$_2$ &amp; O$_2$ heat allows cell operation at 850°C</td>
<td>Water boiler loops can also serve for self-powered decay heat removal</td>
</tr>
</tbody>
</table>
1.2 Plant Modeling

The following Sections provide the list of equations used in the modeling phase of the GFR core. This section is extracted from a more complicate model which is currently in phase of construction at MIT. The model has been named GFRDYN_LUMPED model for its capability to simulate dynamically the GFR plant (containment, core and the DHR passive system) during the reference accident (in this case a Loss of Coolant Accident, LOCA). The model has been built by modeling the minimum number of nodes necessary to consider the interaction between the sub-systems which compose it. A visualization of the model in form of a sketch, which emphasizes the loops followed by the coolant in the accident sequence, is given in Figure 2.

In this work, only a portion of the overall model presented in the figure has been selected; the core sub-module has been preferred to portrait the transient which occurs during a LOCA and therefore “disconnected” from the remaining part of the model. The following Sections introduce first introduce the equations which has been used to model the GFR core in accident scenario, then focuses on the uncertainties around the parameters introduced in the modeling phase, and lastly describes in the detail the reasons behind the choice of the reference scenario and the results obtained from it.

Figure 2: Loop diagram of the GFR plant. The diagram highlights the different loops generated by a 5 nodes representation of the plant.
1.2.1 The GFR Core

In this section we describe the core of the GFR as it behaves during an accident. A simple set of differential equations describes the temporal evolution of the main variables interesting from a dynamic point of view. During any accident or operational scenario the two dominant conditions for the core are to keep the pressure drop low in order to favor natural circulation capability, and, to keep the maximum temperature of the core below the limit imposed by the cladding material.

Core geometry is optimized in order to get high power density, and the composition of fuel is chosen in order to achieve self-sustainable breeding. However, because the interest here is accident conditions, we won’t deal with these features, but will focus on those parameters which are significant in the accident scenario evolution considered and whose uncertainties could play a major role to the final core’s temperature determination: the decay heat curve, the core heat capacity, the core hot spot or peaking factor and the overall heat transfer coefficient. In addition, initial conditions will be considered as a major source of uncertainty as due to specific choice of selecting a particular time frame within the accidental sequence. This last observation will be clarified later in this work.

The set of equations which describes the transfer of heat from the fuel to the gas flowing in the core and then successively extracted from the core is given by the set of equations reported in the next page. The set of equations provides the temporal evolution of the main heat rates and temperatures within the core. The measured output is the difference between the temperature of the cladding and the maximum allowable for the material composing it. The difference measures the safety margins for the considered system.

The description of the system which follows is given in light of the representation of the final system which is going to be seen as the sum of closed circles, or loops, connected to the core. This representation is typical of System Dynamics systems which are characterized by the presence of feedbacks among the main variables. This representation revealed to be suitable also from a thermal hydraulic perspective and helped in the conceptualization of the final plant’s system, seen as the sum of two main T-H circuits passing through the core (The core can be visualized as the 1→2 segment of Figure 2 on the overall loop connecting it to the ECCS...
system). The final representation obtained is coherent with the schematizations used to model the T-H of nuclear power plants [Kazimi and Todreas, 1990].

\[
\frac{dQ_{co}}{dt} = \dot{Q}_{co} = c_{p,co} \cdot M_{co} \frac{dT_{co}}{dt} = \dot{Q}_d - \dot{Q}_{g,co}
\]

Eq. 1

\[
\dot{Q}_{g,co} - \dot{Q}_{g,extr} \equiv c_{p-g,co} \cdot M_{g,co} \cdot (dT_{g,co} / dt)
\]

Eq. 2

\[
\dot{Q}_{g,co} = U_{co-g} \cdot A_{co} \cdot (T_{co} - T_{g,co})
\]

Eq. 3

\[
\dot{Q}_{g,extr} = m_{g,co} \cdot c_{p,g} \cdot (T_{out} - T_{in}) = m_{co} \cdot c_{p,g} \cdot (T_2 - T_1)
\]

Eq. 4

\[
T_{g,co} = (T_{out} (t - \tau) + T_{in}) / 2 = (T(t - \tau) + T) / 2
\]

Eq. 5

\[
\dot{Q}_d = \dot{Q}_0 \cdot a \cdot t^{-b}
\]

Eq. 6

The set of equations, describes the core of the GFR. More generally it can be seen as a cascade of processes where the heat generated by the decay rate, Qd, which is the heat generated in the fuel’s core, is transferred from one process to the other. The first process, see Eq. 1, is the exchange of this heat with the gas contained in the core and the difference between these two rates is the heat which remain stored in the fuel’s core. The heat then flows outside the core: the difference between the heat which is transferred from fuel to the gas coolant, Qg,co, and the heat extracted from the core, Qg,extr, measures the amount of heat which remains in the core’s channels and this is given by Eq. 2. The rate at which the heat is transferred from fuel to the gas coolant, Qg,co, is driven by the difference between the temperature of the fuel in the core, Tco, and the temperature of the gas in the channel, Tg,co, by men of the proportional factor Uco-gAco. This factor is decomposed in the area of the surface through which the heat is exchanged and the overall heat transfer coefficient, Uco-g, which is going to be analyzed in detail in the next sections, as given by Eq. 3. The last two equations provide respectively, the average temperature of the core, which has been described by mean of the average between the inlet and the outlet temperatures, and the decay heat rate, as given by the ANS 79 decay heat formula. Note that the temperature at the outlet is delayed by the time needed by the gas to travel within the core, and that the decay heat formula is an exogenous input to the model. The next Sections describe how to gather the characteristic parameters contained in the above equations starting from the decay constants of Eq. 6.
**Decay Heat Curves**

Because the decay heat represents the fundamental source of heat which is at the basis of the behavior observed during the LOCA (the system wouldn’t evolve in time without it), attention is required to its definition, and dealing with the uncertainties related to it.

The decay heat curve depends on the composition of the fuel. In case of the GFR under study at MIT, only preliminary studies are available, thus we opted to choose the experimental data coming from the Super-Phenix, SPX, reactor, which is similar in its fuel composition and thus it has comparable fission products. The type of model used to draw the decay curve is the ANS 79 used with the data provided by the Super-Phenix decay heat curve reported in Table 2.

<table>
<thead>
<tr>
<th>time(h)</th>
<th>time(day)</th>
<th>SPX (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.173</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.091</td>
</tr>
<tr>
<td>0.03</td>
<td>0.00</td>
<td>0.048</td>
</tr>
<tr>
<td>0.28</td>
<td>0.01</td>
<td>0.025</td>
</tr>
<tr>
<td>0.50</td>
<td>0.02</td>
<td>0.021</td>
</tr>
<tr>
<td>0.83</td>
<td>0.03</td>
<td>0.018</td>
</tr>
<tr>
<td>1.00</td>
<td>0.04</td>
<td>0.017</td>
</tr>
<tr>
<td>3.00</td>
<td>0.13</td>
<td>0.013</td>
</tr>
<tr>
<td>6.00</td>
<td>0.25</td>
<td>0.011</td>
</tr>
<tr>
<td>9.00</td>
<td>0.38</td>
<td>0.009</td>
</tr>
<tr>
<td>12.00</td>
<td>0.50</td>
<td>0.009</td>
</tr>
<tr>
<td>15.00</td>
<td>0.63</td>
<td>0.008</td>
</tr>
<tr>
<td>18.00</td>
<td>0.75</td>
<td>0.008</td>
</tr>
<tr>
<td>24.00</td>
<td>1.00</td>
<td>0.007</td>
</tr>
</tbody>
</table>

The equation has been reduced to a logarithmic equation\(^1\) which has been optimized with respect to the two main parameters “a” and “b” following the historical SPX dataset. The obtained parameter estimates are reported in table above, while equation 7 reports the formula with the resulting values. Time is given in seconds:

\[
\dot{Q}_{in} \, (\%) = 0.173 \cdot t^{-0.28}.
\]

\(^1\) The reactor operation time is assumed to be infinite and the decay heat is mainly that due to the fission products.
Table 3: Confidence intervals on the parameter estimation of the ANS Curve for the Super Phoenix Reactor.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;a&quot; parameter =0.173</td>
<td>0.180</td>
<td>0.166</td>
</tr>
<tr>
<td>variation from predicted</td>
<td>-0.038</td>
<td>0.040</td>
</tr>
<tr>
<td>variation from predicted [%]</td>
<td>-3.85%</td>
<td>4.00%</td>
</tr>
<tr>
<td>&quot;b&quot; parameter =0.28</td>
<td>0.276</td>
<td>0.285</td>
</tr>
<tr>
<td>variation from predicted</td>
<td>0.017</td>
<td>-0.016</td>
</tr>
<tr>
<td>variation from predicted [%]</td>
<td>1.65%</td>
<td>-1.60%</td>
</tr>
</tbody>
</table>

Figure 3: Results from the optimization on the dataset of the Super Phoenix reactor compared to a LWR.

Figure 3 shows the power rate curve obtained from the SPX dataset with the ANS formula and compares it to the corresponding values obtained for a Light Water Reactor, LWR. The two resulting curves are shifted as due to the different fuel composition. Note that this calculation has been performed in order to reduce the uncertainties on the input parameters of the model of the core and therefore avoid the decay heat parameters as possible sources of uncertainty to include in the propagation analysis which will be defined in the next Chapter.
Core Heat Capacity

The calculation for the specific heat at constant pressure is fundamental to determine the core heat capacity which figures in Equation 1. The fast gas reactor in the current design at MIT has a fuel which composes the core according to the table below.

<table>
<thead>
<tr>
<th>Table 4: Tube-In-Duct (TID) Type Core (11ringrev2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Core (No Control Rods Inserted)</td>
</tr>
<tr>
<td>Core Material</td>
</tr>
<tr>
<td>TRU</td>
</tr>
<tr>
<td>BeO</td>
</tr>
<tr>
<td>ODS Ma 956 (Cladding/Support - Steel)</td>
</tr>
<tr>
<td>S-CO2 Coolant</td>
</tr>
</tbody>
</table>

By means of the values reported in the table and by defining the two coefficients x and y, the specific heat of the core, on a heavy metal bases, can be given as:

\[
\bar{\sigma}_p = \frac{270}{238} \left[ c_{p_x} + c_{p_y} \left( \frac{y}{x} \left( \frac{\rho_y}{\rho_z} \right) \right) + c_{p_z} \left( \frac{z}{x} \left( \frac{\rho_z}{\rho_{s}} \right) \right) \right] = 0.536 \text{ KJ/kg core C},
\]

Eq. 8

where:

x = volume fraction UO2

y = volume fraction BeO

z = volume fraction of steel (Fe)

1-(x+y+z) = coolant (CO2)

Then, by taking into account the different uncertainties, based on expert judgments, the value of \( \bar{c}_p \) averaged over the core is:

\[
\bar{\sigma}_p = 0.38 - 0.56 \text{ [KJ/kg core C]},
\]

Eq. 9
Core Hot Spot and Core Average Specific Power

The specific power, for the GFR core is within the range of values reported below:

\[ \dot{Q}_0 = 18\text{-}25 \text{ KW/kg\_core} \]  \hspace{1cm} \text{Eq. 10}

On the basis of the most recent Tube-In-Duct (TID) type core (11RINGREV2), the core average specific power to be equal to 20.7412 KW/Kg-core [C.S. Handwerk at al, 2007].

The core is modeled as a single averaged channel. In order to simulate the hot channel a peaking factor has been introduced. In order to protect the core hot spot we introduce a total peaking factor equal to:

\[ F = F_r \cdot F_z = 1.15 \cdot 1.3 = 1.495 \]  \hspace{1cm} \text{Eq. 11}

This is in turn, determines the maximum temperature of the fuel that allows us to monitor the overall averaged temperature of the cladding and structures within the core.

By putting together all the equations from 1 to 9 and by means of the values provided in the previous sections, it is obtained the model of the core which has been used to run the simulations on which is based the first application of the methodology provided in the next chapter. The only parameter which due to its importance deserves further explanations is given by the overall heat transfer coefficient. With its description concludes this brief section dedicated to illustrate the fundamental features of the core’s model which has been used in this work.
1.2.1.1 Calculation of Temperatures in Cladding and Fuel for a Lumped Core

In this Section, the methodology used to calculate the overall thermal resistance over a lumped core is described. The methodology and assumptions used in calculating the temperature rise across the various materials which constitute the single core channel were adopted from a FORTRAN code called FLOWSPLIT, while the lumping of the core into a unique fuel element was elaborated by following classical geometrical considerations. The FLOWSPLIT code was written by Pavel Hejzlar at MIT in order to provide convenient inputs to the RELAP code. We here limit the discussion to report the set of equations and their relative core parameters. The overall heat transfer coefficient, $U_{co-g}$, used to relate the average temperature of the core to the average temperature of the coolant, as shown in Eq. 3, as taken from this code. Further explanations of the methodology can be found in [M. Pope, 2006] which describes in detail the phenomena and the variables involved in the diagram of a unit TID cell.

Figure 4: Schematization and fundamental geometrical parameters for the core channel.
**Overall Heat Transfer Coefficient**

The overall heat transfer coefficient, \( U_{co-g} \), or inverse of the thermal resistance, relates the linear power of an annular fuel element to the temperature drop \( T_{co} - T_{g,co} \) by considering the series of thermal resistances posed by the fuel, the gap, the cladding and the cooling gas. The coefficient reported below has been calculated assuming a thick clad, neglecting the irradiation term in the gap and assuming the thermal conductivity of the solid annular fuel pellet to be equal to its average value in the temperature range of interest. Also gap size and effect of initial heatup are not considered. The impact of the approximation introduced by these hypotheses is not essential for the purposes of the present work because they do not substantially modify the behavior of the main T-H variables but add uncertainties in the model or parameters which should be taken account of in the successive reliability quantifications.

\[
\bar{T}_{co} - \bar{T}_{g,co} = P_0 \cdot U_{co-g} = P_0 \left[ \frac{K_a}{2\pi k_f L_{tot}} + \frac{1}{2\pi R_{gap} h_{g} L_{tot}} + \frac{\ln(R_{cl} / R_{co})}{2\pi k_e L_{tot}} + \frac{1}{h_{co-g} S_{tot}} \right]
\]

Eq. 12

The methodology followed in order to obtain the four thermal conductivity terms in the above equations are summarized by several considerations. These are explained in the next four sections.

**Thermal Conductivity in the Fuel**

Determination of fuel temperatures is important in order to ensure that no melting occurs during normal operation and to predict fission gas release. An estimate of the thermal conductivity of the fuel must be made in order to determine these temperatures. Usually, for conventional fuels, a great amount of data is available pertaining to the thermal conductivity. In case of the GFR design, however, the specified fuel geometry and composition has not been irradiated in a prototypical experiment and thus some approximations have made [M. Pope, 2006] in order to estimate temperatures during operation.

In estimating the thermal conductivity of the BeO/(U,TRU)O2 fuel, the thermal conductivity of stoichiometric (oxygen/metal atom ratio = 2.00) mixed oxide fuel (MOX) of
20% plutonium and 80% uranium at 95% theoretical density is used as a starting point. Two factors are used to modify this conductivity; one to account for the presence of BeO, and one to account for the change of the conductivity resulting from burnup by means of the major assumption of the separability of these two effects. Further details are once again provided by [M. Pope, 2006] from which we report the fundamental results for the resulting $K_f$ shown in Figure 5.

\[
K_f = 1.5 \cdot F_{bu} \left[ \frac{1}{0.042 + 2.71 \cdot 10^{-4} T_{co}} + 6.9 \cdot 10^{-11} T_{co}^3 \right]
\]

Eq. 13

where $k$ is thermal conductivity in W/m·K and $T_{co}$ is the temperature of the lumped core in °K. The factor of 1.5 accounts for the 50% increase in conductivity from the presence of the BeO diluent, and the burnup degradation factor $F_{bu}$ is evaluated at 120 MWd/kg by means of a polynomial expression which is a function of the temperature.

![Figure 5: Estimated thermal conductivity of BeO/(U,TRU)O2 fuel. Source of the figure: M. Pope, 2006 page 59 [1].](image)

The $K_a$ factor arises from the solution to the one-dimensional Fourier’s heat transfer equation for an annulus of fuel, and introduces a geometrical dependence which is at the basis of the
adopted geometry because it benefits the thermal conductance within the element:

$$K_a = \left[ \frac{(1 - (R_{f1} / R_{ea})^2)}{4} - \frac{\ln(R_{ea} / R_{f1})}{2} \right]$$

Eq. 14

**Thermal Conductivity in the Cladding**

Since the cladding wall thickness is relatively small compared to its diameter and the temperature change across the cladding is small relative to the total temperature difference between coolant and peak fuel, heat transfer through the cladding can be treated with reasonable accuracy as one-dimensional heat transfer through a constant-conductivity medium. This temperature rise across the cladding is estimated for a lumped fuel element by the third term of Equation 3. The material chosen for the cladding is ODS MA956 [Special metals, 2004].

**Thermal conductivity across the Fuel-Clad Gap**

The $h_{gap}$ contained in the second term of Equation 3 is the conductance at the fuel-clad interface which is approximately given by $h_{gap} = k_{gas} / t_{gap}$ where $k_{gas}$ is the thermal conductivity of the gas filling the gap that here is assumed to be CO$_2$ at constant pressure and, $t_{gap}$ is the fuel-clad gap thickness. The approximation introduced by the formula is due to the neglect of a second term, not shown here, accounting for radiation. Previous studies [M. Pope, 2006] showed that for values of the parameters $k_{gas} = 0.08$ W/mK and $t_{gap} = 70$ µm, the gap conductance resulting from conduction alone has a value of 1000 W/m$^2$K and thus, the conduction term tends to be larger than the radiation term, and total gap conductance values in the range from 1000 to 1500 W/m$^2$K can be expected.
**Geometrical Lumping Parameters**

The basic relationship existing between the linear power and the total thermal power explains the presence of $L_{\text{tot}}$ in the first three terms of Eq. 3, which groups the fuel heights of the active channels in a unique parameter given by $L_{\text{tot}} = n_{c} \cdot H_{c}$. The fourth term of the equation make use of $S_{\text{tot}}$ which is the total surface of the channels in contact with the coolant, it is given by $S_{\text{tot}} = n_{c} \cdot (\pi R_{c}^2)$. The linear and the volumetric power are averaged over the axial dimension and thus they represent the average of the co-sinusoidal flux distribution in this direction. Finally note that formally these simplifications are valid in a steady state regime, and constitute a valid approximation in the quasi-steady state transient analyzed in this work.

\[ q' L_{\text{tot}} = q'' \pi R_{c}^2 L_{\text{tot}} \cong P_0 \]  

Eq. 15

In the table below we report the main geometrical values and the core parameters described in this Section. Note that the values are consistent with previous RELAP calculations and with the geometrical parameters of the selected design in this work.
<table>
<thead>
<tr>
<th>Table 5: Parameters and geometrical dimensions used for the GFRDYN Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design</strong></td>
</tr>
<tr>
<td>Thermal power (MW&lt;sub&gt;th&lt;/sub&gt;)</td>
</tr>
<tr>
<td>Axial power shape</td>
</tr>
<tr>
<td>Axial peaking factor</td>
</tr>
<tr>
<td>Radial peaking factor</td>
</tr>
<tr>
<td>Channel Pitch</td>
</tr>
<tr>
<td><strong>CLAD</strong></td>
</tr>
<tr>
<td>Clad thickness (m)</td>
</tr>
<tr>
<td>Clad thermal conductivity (W/m·K)</td>
</tr>
<tr>
<td>Inner Cladding Radius R&lt;sub&gt;ci&lt;/sub&gt; (m)</td>
</tr>
<tr>
<td>Outer Cladding Radius R&lt;sub&gt;co&lt;/sub&gt; (m)</td>
</tr>
<tr>
<td><strong>GAP</strong></td>
</tr>
<tr>
<td>Average Radius Gap</td>
</tr>
<tr>
<td>Gap CO&lt;sub&gt;2&lt;/sub&gt; conductivity (W/m·K)</td>
</tr>
<tr>
<td>Hot gap thickness (m)</td>
</tr>
<tr>
<td><strong>FUEL</strong></td>
</tr>
<tr>
<td>Fuel conductivity (W/m·K)</td>
</tr>
<tr>
<td>Internal Radius Channel RFI</td>
</tr>
<tr>
<td><strong>CORE</strong></td>
</tr>
<tr>
<td>Channels in the core</td>
</tr>
<tr>
<td>External Equivalent Channel Radius Rea (m)</td>
</tr>
<tr>
<td>Core's channel Active Length (m)</td>
</tr>
<tr>
<td>Heated Core Area = lumped core surface (m&lt;sup&gt;2&lt;/sup&gt;)</td>
</tr>
<tr>
<td>Active Core Volume(m&lt;sup&gt;3&lt;/sup&gt;)</td>
</tr>
<tr>
<td>Lumped Core's Length= Active Length*Channels in the core(m)</td>
</tr>
<tr>
<td>Heat transfer coefficient channel gas interface (W/m&lt;sup&gt;2&lt;/sup&gt;·K)</td>
</tr>
<tr>
<td>Channel surface roughness</td>
</tr>
</tbody>
</table>

The values reported in the table for the different annuli constituting the single fuel channel are calculated at hot conditions. Finally note that in the current model the value of the average Kf(T) has been chosen, and that the heat transfer coefficient value as well as all the other quantities composing the overall heat transfer coefficient were calculated, under steady state conditions so they are subject to uncertainties which could be later addressed into reliability calculations.
1.2.2 The Decay Heat Removal System

By definition, passive system functionality does not rely on an external source of energy. The advantage of such systems is that they are always available, yet they rely on energy sources which are weak compared to those of the more traditional active systems [L. Pagani, 2007]. For this reason, in the present work, a limited time window within the LOCA’s evolution has been selected during which the force generated by the system is at its lowest value. Incidentally, this window corresponds to the quasi-steady state transient - period that goes from the time the pressure within the different volumes of the plant reaches equilibrium, to the time the decay heat goes below a pre-defined cutoff value. During this time frame – as the force decreases – the exposure to uncertainties will be higher.

Furthermore, the reference time window assumed, allows excluding the uncertainties related to the time of the intervention of the system and thus excluding from the modeling active pumps and diesel generator system which are typically working in the first few minutes following the scram [JC. Gamier et al., 2006].

In fact, the Decay Heat Removal Systems used to remove heat from the GFR plant are two kinds: passive and active. Hybrid functioning of the two is also considered and details about this can be found in [L. Pagani, 2007]. The two modes of operation are described below:

Active mode: In this mode, an electric motor, EM, obtains electricity from offsite or onsite power sources, and rotates the blower, B. The hot coolant (coming from the core) flows through a check-valve, CV, that it kept closed during full power operation (to prevent backflow from station 3 to 2) and opens for emergency cooling. Then the coolant flows through the Heatric TM Heat Exchanger (HCHX), where the heat is removed (see heat-sink).

Passive mode: The blower is supposed to be inoperable (failure of the blower or of the power sources). Passive flow is induced by the temperature difference between the heat source (core) and the cold source (HCHX, which is located higher than the core).

The analyses performed in the next sections refer to the passive system DHR or to the core connected to it. As stated above, by mean of a selected time-window, the time in which the active system intervenes is excluded from calculations.
1.2.3 The Selected Time-Frame

Among the entire sequence of events characterizing a loss of coolant accident, two particular time are of major interest for our calculations:

- The time at which the equilibrium in pressure (on average) is reached within the whole system composed by the core, the primary circuit and the containment building.
- The time at which the heat produced by the fission products is reduced to 1 MWth and thus, reduced to a cut-off value under which the heat absorption capability of the core is sufficient to remove the heat from the system without producing any further and significant thermodynamic change on the system.

These two major events concur in defining the time frame or, reference time window, we used in our simulations. The reasons which underlying selection of this specific time frame are as follows:

1. Passive systems could be described as systems acting due to weak forces. Among these forces, gravity is the most often used to let the system automatically intervene without resorting to any operator or externally induced force. The decay heat removal system of the GFR is designed in such a way that a gravity induced force, the buoyancy head, drives the gas through a cooling loop. The regime within the loop is dominated by natural convection characterized by a mass flow rate which is lower when compared to the value that could be obtained by actively pumping the gas. The objective of a rigorous safety analysis of the loop then must consider this weak behavior of the system, and the locations and or time-frames in which weakest behaviors can be more stressed. In case of a LOCA, the blow-down time is characterized by high pressures and velocities, thus represented by strong forces imposed on the system over a limited time. The time frame which experiences the weakest behavior is then after the blow-down peak, when the passive DHR has already started removing heat and which has larger temporal horizons and the weakest forces involved (natural convection mechanism in the DHR and containment).

The time window selected initiates starts with reaching an overall equilibrium of the gas among all the possible control volumes connected to the primary system, and ends with a definitive completion of the heat removal from the gas.
2. The phase in which the discharge of the gas into the containment is completely over (about 30 seconds) should be sufficient to discharge a gas with that mole fraction, starting from a 20 MPa pressure regardless of the size of the break considered, and taken into account any uncertainty related to the dimensions of the volume into which it expands at the time when the gas reaches its equilibrium pressure. This avoids including the critical mass flow rate within the model with the benefit of not considering any uncertainty related to the break size, possible changes of phase of the super-critical CO₂, and some components such as active blowers and valves (which is in general few milliseconds if operated automatically, and minutes if manually started), which of course play a role in the initial instants of the accident or post-scram sequence.

3. At the considered initial time (equal to 550 seconds as provided by the RELAP-3D² code) the gas has already reached the so called back-up pressure, which is fundamental in supporting the natural convection mechanism within the ECCS.

4. At this time, local phenomena such as possible stratification of the gas within the containment, or, inhomogeneous local accumulations of the gas can be excluded.

5. Finally, the time at which we stop our simulations, is the time for the decay heat curve to reach the cut-off value of 1 MWth.

6. *In ultimis*, by selecting a limited time frame, we save computational time.

---

² J.Lee has run the code for a LOCA with this set of assumptions which, been used as initial conditions for the model built in order to simulate the core dynamics, namely the GFRDYN_LUMPED Model [Lee, J.I., et al., 2007].
1.2.4 The Reference Scenario

The proposed simulation considers the dynamic evolution of the main thermo-dynamic properties during a LOCA in a selected time-window. The simulation starts after that the blow-down occurs and therefore when the gas is completely discharged from the primary system to the containment building. The energy in the gas, in case of the Fast Gas Reactor under study at MIT is the supercritical CO2, is being removed by the designated auxiliary system based on natural recirculation, the ECCS (Emergency Core Cooling System).

Previous RELAP5-3D simulations show that the timing of the fuel release depends on the size of the break on the primary coaxial duct of the GFR. For a 70,000 metric cubic feet containment volume, with a 100 inch² break on the coaxial duct, the full release completes at around 550 seconds after the accident occurs.

This result is in accordance with calculations of the critical flow reached by the gas when simulated as perfect and single phased. The achieved time value of 550 seconds has been set as the initial time for simulations, and the details behind this choice were explained in the previous Section. The final time for all simulations is set to be at 10,000 seconds, which is when the gas is expected to reach steady-state conditions. In this interval of time, the gas discharged in the first 550 seconds is expected to gradually reach steady-state equilibrium among the different sub-volumes containing it. The observed behavior is usually called quasi-static equilibrium and mathematically expressed by introducing negligible inertia values in the equations of conservation of moment applied along the circuit in which heat is removed from the core (the passive recirculation system).

The figure below reports a classical sequence of the proposed accident in the range 550-10,000 seconds, for the temperature of the core. The temperature of the core starts from an initial value, $T_{\text{fuel} \ 0}$, which is provided by previous RELAP5-3D simulations and thus equal to 600 K. The observed peaking behavior for the reference case is due to both the initial value of the mass flow rate, Gas Mass Flow Rate 0, and to the still high value of the decay heat rate. As time passes the mass flow rate within the ECCS loop increases and the decay heat rate diminishes, therefore increasing the rate of extraction of heat from the core, which in turn leads to the progressive decrease of temperature shown by Figure 6.
Figure 6: Reference case of the fuel temperature, as given by the GFRDYN_LUMPED Model.

In terms of safety margins it can be observed that the temperature gradients within the core (temperature of the cladding and temperature of the fuel) are respected and thus that their nominal values (mean values or best estimated values as predicted from the code) are well below the two lines of Figure 7 defining the failure criteria of the materials in use.

Figure 7: Reference case simulations of the cladding and fuel margins. Blue line represents the temperature transient while the red straight lines show the limits imposed by the materials.
1.3 Modeling Techniques

In this section we present a brief overview of the main modeling techniques used to perform the calculations. The reasons behind their choice are provided together with some of the fundamental features. Specifically, the first section provides insights about the applicability of the System Dynamics paradigm to technical problems. The second part gives an overview of the supervised learning methods which are going to be used to interpret the data obtained from the modeling phase, and provides basic elements of their applicability.

1.3.1 Systems Dynamics Modeling

*System Dynamics to Model Transients in T-H Plants*

System dynamics is a methodology for studying and managing complex feedback systems, such as one finds in business and other social systems. In fact it has been used to address practically every sort of feedback system.

The system and the accident sequence considered are described dynamically by means of Causal-Loop Diagrams and then Stock-Flow Diagrams taking into account all important relations and variables which describe the behavior of the system. This modeling approach, traditionally known under the name of System Dynamics, has been considered suitable to model T-H systems because of the motivations that follow:

*Lumped Modeling versus Traditional T-H Codes*

The choice of a lumped parameter model favors the intuitive understanding of uncertainties propagation, it is easy to implement, provide results that are within the range of validity of traditional codes. Although Systems Dynamics modeling yields a level of accuracy inferior to that of traditional RELAP codes, it is used because it constructively enables us to
catch non-linear feedbacks between the activity occurring in the core and the phenomena taking place in other regions of the plant.

Passive systems in general and especially passive systems based on natural recirculation, are exposed to great changes in their state as due to the variation of initial and boundary conditions. This particularly high sensitivity to the initial condition of the system is conventionally well known and captured by system dynamic models. SD representation reveals particularly appealing in managing the propagation of uncertainties affecting the value of the parameters at initial times.

*Respect of SD Software Capabilities during the Selected Simulation Time-Frame*

A pre-defined time window has been chosen as a privileged interval to check the capabilities of the passive recirculation system under the weakest operating condition. Incidentally, this window corresponds to the quasi-steady state transient period that goes from the time the pressure within the different volumes of the plant reaches equilibrium, to the time the decay heat goes below a pre-defined cutoff value. The reasons beyond the choice of the specified time interval are mainly of physical nature and intimately related to the definition of the safety margins defined for the passive system. So despite the choice to analyze the selected time interval has no foundation in the SD field, the use of SD to define a quasi-steady state transient reveals to be appropriate and actually less computationally expensive than traditional T-H codes.

*Improvements in the Computational Capability to Explore Transients*

The desired accuracy provided by the SD code is preserved and actually allows analyzing in detail – more simulations for each fundamental parameter variation – the scenarios of interest. RELAP5 simulation of analogous system [F. MacKay, 2007] produced 128 runs in about 12 hours considering the variation of six fundamental parameters at a time. Most of the time was attributable to the higher plant’s representation but the response was not always statistically representative of the sample analyzed.
Other General Reasons

System Dynamics as it is known today is not practiced in physics very much. It is used in social sciences, management applications and in general to analyze complex systems.

Although its use is not very diffused to analyze T-H systems, system dynamics tools have been used for modeling dynamical physical processes before: an in-depth study of the dynamical structure of thermal processes, including a discussion of the unified view of physical processes, can be found in [H.U. Fuchs, 2001].

In the specific case of the lumped model provided here basic constitutive equations were used and the whole system can be built by means of mass, and momentum conservation and energy. Energy has three fundamental properties, apart from the fact that it is conserved: it can be stored, it can flow, and it can be released (or bound). In other words, if the rate at which energy is released in one process can be calculated, we can equate it to the rate at which energy is bound in the follow up process. The energy principle and energy properties can be easily used to relate different processes to one another and thus to show the different interconnection between different parts of the plant.

All the presented features made possible the use of System Dynamics model as a valid model to represent the GFR plant. A visual layout of the plant in form of a Vensim Sketch is provided in the Figure below.

**Figure 8:** Sketch illustrates the main variables and parameters modeled by the GFRDYN Model.
1.3.2 Supervised Learning Modeling

Supervised learning is a machine learning technique for creating a function from training data. The training data consist of pairs of input objects, and desired outputs. The output of the function can be a continuous value, or can predict a class label of the input object. The task of the supervised learner is to predict the value of the function for any valid input object after having seen a number of training examples. To achieve this, the learner has to generalize from the presented data to unseen situations in a “reasonable” way.

Supervised learning can generate models of two types. Most commonly, supervised learning generates a global model that maps input objects to desired outputs. In some cases, however, the map is implemented as a set of local models (such as in case-based reasoning or the nearest neighbor algorithm use in cluster analyses).

In order to solve a given problem of supervised learning one has to consider various steps:

- Determine the type of training examples. Before doing anything else, the engineer should decide what kind of data is to be used as an example. For instance, this might be a single handwritten character, an entire handwritten word, or an entire line of handwriting.
- Gathering a training set. The training set needs to be characteristic of the real-world use of the function. Thus, a set of input objects is gathered and corresponding outputs are also gathered, either from human experts or from measurements.
- Determine the input feature representation of the learned function. The accuracy of the learned function depends strongly on how the input object is represented. Typically, the input object is transformed into a feature vector, which contains a number of features that are descriptive of the object. The number of features should not be too large, because of the curse of dimensionality; but should be large enough to accurately predict the output.
- Determine the structure of the learned function and corresponding learning algorithm. For example, the engineer may choose to use artificial neural networks or decision trees.
• Complete the design. The engineer then runs the learning algorithm on the gathered training set. Parameters of the learning algorithm may be adjusted by optimizing performance on a subset (the so called validation set) of the training set, or via cross-validation. After parameter adjustment and learning, the performance of the algorithm may be measured on a test set that is separate from the training set.

Another term for supervised learning is classification. A wide range of classifiers are available, each with its strengths and weaknesses. Classifier performance depends greatly on the characteristics of the data to be classified. There is no single classifier that works best on all given problems. Various empirical tests have been performed to compare classifier performance and to find the characteristics of data that determine classifier performance. Determining a suitable classifier for a given problem is not always an easy operation. For this work, two very widespread techniques have been selected, the Logistic Regression classifier and the Classification Tree. Both these techniques are linear with regard of the fundamental parameters concurring to form the desired output. Despite that, their use has been widely extended to non-linear models [Marquez et al., 2005]. The two methodologies in detail are explained in Section 2 and along with the examples of Section 3.
2. THE METHODOLOGY

2.1 About Passive Systems

The issue of a correct analytic approach to Passive System’ Reliability (PSR), is a recurring one in the design of future generation power plants. One of the principal goals of the new- generation nuclear power plants is, in fact, to address safety in a more efficient and consistent manner. A convenient way to pursue this goal is by either adding, coupling or substituting the traditional active systems with systems that are able to react spontaneously, like those based on ‘natural’ mechanisms. These systems, which are termed Passive, are originally conceived as capable of automatic intervention while being disconnected from all auxiliary sources of energy and they are more specifically designed to be independent from common causes of failure - first among all, human error. Their efficacy though, is yet to be proven, since at the present time, they are still the object of an ongoing debate concerning their possible introduction in nuclear power plants. Among the concerns raised by this debate are the difficulties that would arise during the necessary phases of their testing and repairing, their necessary dependence to contiguous parts of the plant, and the complexities associated with their inclusion in standard PRAs.

This part of the work aims to investigate in detail the motivations behind the latter of these concerns - the inclusion of passive systems in standard PRAs. The methodology that was used in order to achieve this will be explained in this section following a brief account of the issues that emerge during the analysis performed to assess the reliability of Passive Systems.

The development of any proper methodology, aimed at assessing the reliability of a Passive System, is characterized by the following four fundamental steps:

I. Identification and quantification of the sources of uncertainties;
II. Propagation of the uncertainties through a T-H model;
III. Evaluation of the T-H system reliability;
IV. Inclusion of the T-H reliability in the accident sequence.
These milestones are specially indicated for those passive systems characterized by a working and moving fluid (termed Thermo-Hydraulic, T-H). This work focuses mainly on the first 3 steps of the checklist. The first milestone, relating to the identification and quantification of the sources of uncertainties, follows this introduction. The next point, which relates to the propagation of such uncertainties throughout a T-H model, will be argued later on in this section, as part of the introduction to the techniques and tools, in order to better explain the algorithm used to simulate this behavior. The third step, concerning the evaluation of a T-H system in light of the previous analyses, will be expounded throughout the formal discussion of the method.

T-H systems are of great importance and fall under the IAEA Categories B and C. Generally their specific mission is associated with a cooling function and in most of the cases their behavior is difficult to predict, especially if solicited by external or internal perturbations. This last aspect is of fundamental importance, because these systems usually rely solely on forces such as gravity, which by nature are weak. This implies that, as opposed to traditional active turbo-machineries, in order to accomplish the same task - for instance that of removing the decay heat- a passive system usually takes longer and it requires specific environmental conditions to operate. As a result, the conventional uncertainties which any T-H model is exposed to, combine with those arising from the use of this particular type of system. This aspect makes Passive Systems particularly vulnerable to those perturbations which are hardest to predict, and most likely to occur, in the event of an accident. This possible vulnerability to external factors was taken into account by concentrating, during the analysis of the accidental scenario, on studying the specific time of the transient under which the convection mechanism is less stable (since it’s just across different T-H regimes). Example I of Section 3 analyses a component of the passive system and subjected to a thermal shock.

These and other fundamental aspects of PSR would be clarified later in this work. The design of a methodology to help conduct the analyses of milestones I, II and III will be the next objective within this discussion. Finally the tools used to solve the problem with the new approach, will be given at the end of this Section.

Though the focus of this study remains on the relevant parameters - taken as input to the models used to characterize the plant, or portions thereof, during the selected scenario – it is nonetheless important to know the whole gamma of uncertainties that could affect the operation
of the passive system. Following is a list of most of the sources of uncertainties characterizing such systems:

- **Initial and boundary conditions.** This is particularly observable in Example I of this work: it focuses on a specific time-frame of the accident by neglecting the initial discharge of the gas within the containment, and thus the initial conditions remain uncertain (even if provided by separate analyses). Another typical example of this kind of uncertainty is the initial time setting, meaning the selection of the moment during which the system begins to intervene.

- **Phenomenological uncertainties,** as due to the possible unknown phenomena occurring in the working fluid. Typically these are consequence of approximations made in modeling the physical processes such as correlations or phenomena like the stratification of the gas. This aspect is very crucial because as explained in the previous chapter, it is cause for deviations from the expected normal operational state of the system.

- **Parametrical uncertainties,** due to the lack of data, as is the case for innovative technologies such as the GFR. Approximations in the system’s geometry often take place. An apparent example of this is the chosen height at which the Heat Exchanger of the ECCS is placed.

- **Uncertainties deriving from modeling aspects which are unknown because still untested and unobserved.** This is the case of the passive systems for which only few prototypes are available, and thus little to no data from experiments exist.

- **Uncertainties which are a consequence of an inaccurate determination of failure/success criteria.** An example of this could be the choice of an epistemic failure limit for the temperature of the cladding, as it is for the material used in the GFR, which are not tested yet under real operating conditions (radiation, high temperature, and so on). In this work, the consequences of multiple failure modes, or failure criteria, are highlighted in Example III.

- **Internal uncertainties improperly evaluated by the T-H codes.** This aspect is not stressed through this work, but it is important for a comprehensive assessment of the accuracy provided by the adopted codes with regards to the overall accuracy in gauging the above mentioned sources of uncertainty.
2.2. Available Resources

The PSR debate is major, thus it is important to define a methodology because of the pressing need to include such systems in the next generation of nuclear power plants. In view of this, there have been a number of diverse approaches to fulfill this task. Most of them, given that they refer to T-H applications, follow the milestones mentioned in the previous chapter. There is no evidence as of yet of a commitment to a specific technique for the identification of the fundamental parameters, yet there is a collective agreement that such techniques should not be based on a single method but instead on the comparison of multiple ones [J.P.C. Kleijnen, J.C. Helton, 1999; Marques et al., 2005]. A good portion of the relevant available literature [Zio et al., 2003; Marques et al., 2005] justifies the need for a method. Most of these efforts recommend solving this problem by adopting the Analytical Hierarchy Process, AHP, proposed to help identify the important driving parameters then rank them by building a hierarchy based on pairwise comparisons [Saaty, 1980]. However, the effectiveness of this method is limited by its qualitative nature, together with the sheer manpower needed to carry it out (it requires to have multiple analysts’ meticulous judgments).

An equally important matter centers on clarifying the need, or lack thereof, of iteration loops between the tasks required to evaluate the PSR. A number of studies support the use of iteration loops [Ricotti. et al., 2002; Marques et al., 2005], while other more recent works maintain that the focus should be on simplifying the conceptual development of a method, thus sparing iterative processes and further initial sensitivity analyses [J. Mackay et al., 2007]. These and other issues will be dealt with later on in this document, while discussing the development of the methodology proposed here.

Additional sources that have been researched for the purposes of the present work are efforts that do not specifically refer to safety systems application [Wilson et al., 1998], or even to nuclear power plants [J.C. Helton et al., 2006], yet they were taken into consideration for their approach to specific obstacles common to PSR.
2.3 Possible Applications and Main Objectives

The method proposed here, aims to improve the existing PSR methodologies specifically by including a quantitative evaluation of the fundamental parameters. Given the nature of the problem, this is in truth a multi-objective task, thus, it was elaborated by following the goals listed below. Further evidence of their impact will be given in the examples discussed by the next Section.

A. Distinguish among the available parameters those that significantly contribute to the failure of the single-systems composing the nuclear power plant. The measure selected as acceptance criterion for the parameters is the statistical significance;
B. Define a hierarchy of the parameters concurring to the failure of the system. The parameters can be ranked with respect to their ability to lead the system to failure, following a classification based, again, on their statistical significance;
C. Prioritize the sub-systems composing the plant, on the basis of the contribution given by the single parameters which lead the system to failure;
D. Provide possible combinations/clusters/parameter sets which can lead the plant to anomalous behaviors (failure or semi-failure), thus enabling engineers to avoid certain configurations (combination of parameters);
E. Interpret the statistical significance of the results obtained from the code;
F. Predict the behavior of the plant, and thus reduce the computational time which is required to run further sensitivity tests;

Note that this type of methodology can be applied to steady state problems as well as to dynamic ones. This study focuses on the latter type of systems so those undergoing transient conditions and quasi-static states. The benefit of this type of analysis is that it makes it possible to detect failures of the systems instantaneously and simultaneously (time dependence failure progression can also be taken into account). Note that, the need of an application of a method for PSR to transient problems, was pointed out by previous analyses [Pagani et al., 2005]; its relevance to steady state problem is discussed with Example III in Section 3.
Some of the main advantages in providing a methodology as that described in the next chapter are the following:

- **Pre-design risk-informed**: test different design strategies at their initial stage. i.e. Containment volume can be used as a proxy for the different pressures and the probability of failure can be measured for different values of the system or containment pressure;
- **Regulatory aspects**: establish design sequences and iterative risk-informed evaluations of the plant;
- **Optimizations check**: reliability versus economical feasibility, actual state of technology and flexibility of design.

These three driving benefits, among the many possible, were selected because of their relation to risk-informed procedures. The latter are also part of the ongoing debate about PSR. Their inclusion could be part of future developments of this study and it is discussed in the last Section of this document. The next chapter will venture into the formal description of the method proposed in this document, clarifications on the statistical tools, sample techniques and software used in this effort, will also be provided later on.

### 2.4 The Methodology

In traditional Thermal-Hydraulic (T-H) codes, the computational time required to have a satisfactory number of runs is usually high. In addition, the high number of parameters involved in the sensitivity analyses calls for a ranking of their importance based on their relative contribution to the system failure. Therefore, a method based on data interpretation techniques is thus proposed in order to provide guidance in the inclusion of passive systems in PSA.

The following diagram provides a possible roadmap for the determination of the PSR. The single steps of the methodology are described below by highlighting the contribution given
by the appliance of statistical techniques in order to address some of the fundamental steps associated with that.

Each step of the technique has been validated and tested through the results obtained in the course of three different examples illustrated in the next Section. Note that the comprehensive analysis presented here has been first elaborated at MIT\textsuperscript{3} and then tested in a preliminary manner via the present work.

The conceptual development of the methodology is here presented following the tasks shown in Figure 9. Note that the check-list is expanded to catch more aspects than in the above step.

\textsuperscript{3} The methodology proposed in this work has been implemented at the 15.077/ESD.753 - Statistical Learning & Data Mining – class held by Prof. R. Welsch whose contributions to this work are much appreciated.
sketch, and therefore does not exactly match with the diagrams in the figure which have been kept compact in order to maintain its original illustrative purpose.

1. Build the model of the plant;
2. Build the model of the accident;
3. Choose an initial set of parameters and attribute them a distribution (PDF);
4. Define the failure criteria for the plant;
5. Determine the functional failure states of the plant;
6. Determine the overall failure of the plant;
7. Determine the relevant parameters concurring to the single failure of the plant;
   7a. Run a sensitivity analysis as a function of the initial pre-defined set of parameters;
   7b. Use the predicted realizations as dependent variables of a logistic regression;
   7c. Exclude from the dataset those parameters which are not statistically significant in leading the system to failure (p-statistic above 0.05) and rank the remaining ones;
   7d. Verify that the new overall failure of the system does not change when the non-influential parameters are removed. If it change an interaction analysis is required;
8. Interpret Failures through the pair-wise comparison of the statistically significant parameters through the scope of the scatterplots;
   8a. Use a classification tree to determine the value of the parameters which lead the plant to failure and cross validate the importance of the branch of the tree with the values of the p-statistic originated by the logistic model;
9. Go back to point 1 and re-design the system in light of the new information provided by the classifiers.

Asides from the nine points through this table, it is important, for the sake of completion, to add two more remarks to those already made. The first is that the core of the present analysis is given in step 7 where it is pointed out the way the effective Parameters Evaluation for Passive System’s Reliability, PEPSR. The second is an amendment to step 7d: there should be some interaction among the fundamental parameters used as a input for the model, when that occurs, removing the parameters with high p-values could lead to significant changes in the measured
Failure Probability, PF and in the re-obtained –values of the parameters concurring to it. Explicit techniques may be used, if such limitations are known a priori from the analyst that the states of knowledge of the parameters are dependent: the adoption of multivariate distributions or conditional probability distributions would ‘remove’ the presence of common contributors to the uncertainty of the parameters. Instances of such limitations need to be identified and addressed by showing them in specific simulations which are not given in the present work.

A final observation is about the overall methodology compared to similar milestones: the method developed is recursive in the determination of the fundamental parameters participating to the final output PF and avoid the use of preliminary or parallel sensitivity analyses in addition to the fundamental analysis depicted in Figure 10.

Further aspects of the developed methodology are differed to the practical application offered by the exercises contained in Section 3.

2.5 Tools and Computational Methods

In order to address to test the methodology illustrated in the previous chapters, a combination of techniques has been chosen. The interconnection existing between the different modules which compute the nine steps of the milestones’ list is depicted in Figure 10. The diagram presented in the figure can be divided into three main computational zones:

I. Input Analysis and LHS sampling;
II. Model of the world processing and output;
III. Data mining analysis and interpretation of the results;

The three zones, as shown in the figure, represent different processes in cascade: first we select the input, define their distributions, then we process them into the selected model of the world (GFRDYN_LUMPED Model or RELAP5-3D code), then, the obtained output (probability of failure) is processed along with the dataset created by LHS in a unique table that the classification models (logistic regression and classification tree) read to train and validate before
providing their responses. Last step is the reciprocal comparison between LR and CT via agreement tables and lift charts. A brief presentation of the tools is given in the next sections.

Figure 10: Visual of the scheme used to create the dataset which is used as input for the supervising models.

2.5.1 Classification Techniques

In order to address the objectives specified in the previous chapters, a combination of two popular prediction techniques, the classification tree and the logistic regression, was adopted. This choice is motivated chiefly by the simplicity in the application of the two methods, and to a lesser extent because of the effortless in the understanding of their results.

Note that the adoption of the two techniques in parallel guarantees an extra check on the accuracy of the obtained results [J.P.C. Kleijnen ET AL., 1999]. Furthermore, the cross validation between the two techniques, increases the confidence in the overall method and provides higher robustness and credibility to each of the two techniques.

Lastly, classification techniques allow for the verification of the basic assumptions made on the parameters we input in the sensitivity analyses i.e. their independence.
The next section provides details of these two techniques by specifying their comparative features and by illustrating them in a first example.

The use of classification models to compute predictions for a discrete or categorical dependent variable is widespread and well recognized. Examples of dependent variables in this type of model are binary variables. In these, as the name implies, there are exactly two levels - or possible states - this makes them particularly suitable to model the failure state of a plant. The variables in the model that determine the predictions are called the independent variables, which, if the goal is to describe the plant’s reliability, represent the possible contributors to the final failure states.

**Classification Trees**

Classification trees, CT, are tree-based models that provide a simple yet powerful way to predict a categorical response based on a collection of predictor variables. The data are recursively partitioned into two groups based on predictor (independent) variables. This is repeated until the response (dependent) variable is homogenous. The sequence of splits of the predictor variables can be displayed as a binary tree, hence the name.

A classification tree can be described as a series of rules. For a response $y$ and given a set of predictors $x_1, x_2, ..., x$, a classification tree rule would be of the form:

\[ \text{If } x_1 < 23 \text{ and } x_2 \in \{A, E\} \text{ then } y \text{ is most likely category 2} \]

The simplicity of the model display and prediction rules makes classification trees an attractive data mining tool. Other advantages of tree models include:

- Invariance to monotone re-expression of the predictor variables;
- Can easily capture nonlinear behavior in a predictor as well as interactions among predictors;
- Unlike logistic regression, can model categorical response variables with more than two levels.
Logistic Regression Classifiers

The logistic regression, LR, allows for the modeling of the probability of a binary event occurring as a linear function of a set of independent variables. Logistic regression models are a special type of linear model in which the dependent variable is categorical and has exactly two levels or possible outcomes; because of this they are particularly suitable for predicting the probability of failure.

A linear model provides a way of estimating a dependent variable, conditional on a linear function of a set of independent variables, $X_1, X_2 \ldots X$. In classical notation this can be expressed as:

$$Y = \beta_0 + \sum_{i=1}^{p} \beta_i X_i + \epsilon.$$  

Eq. 16

In this equation, the $\beta_i$ terms are the coefficients of the linear model; the intercept of the model is $\beta_0$ and $\epsilon$ is the residual. Estimates of the coefficients, $\hat{\beta}_i$, are computed from the training data from which an estimate of the dependent variable, $\hat{Y}$ is calculated by substituting the estimated coefficients in Eq. 16. An estimate of the residual, $\hat{\epsilon}$, is obtained by computing the difference between the observed dependent variable and its estimate.

As mentioned earlier, in a logistic regression model the dependent variable is binary, and each of the two class levels are coded as either a 0 or 1. The estimated dependent variable, $\hat{Y}$, is a guess of the probability of the level coded as a 1. The logistic regression model uses the logistic function to express $\hat{Y}$ as a linear function of the set of independent variables. In its classical notation the equation reads as:

$$g(\hat{Y}) = \log \left( \frac{\hat{Y}}{1-\hat{Y}} \right) = \hat{\beta}_0 + \sum_{i=1}^{p} \hat{\beta}_i X_i + \hat{\epsilon}.$$  

Eq. 17
2.5.2 Latin Hypercube for Sensitivity Analyses

Sensitivity analysis is widely recognized as a fundamental component of studies based on mathematical modeling and because of that is used to accomplish uncertainty analyses for PRA and specifically to this work for PSR evaluations. Although sensitivity analysis is intimately tied to uncertainty analysis, it has a tendency to be a more complex undertaking due to both the variety of possible measures of sensitivity and the additional computational procedures required to evaluate these measures. LHS is recognized to be a suitable technique to improve the computational time of simulators and it constitutes part of the Monte Carlo procedure used for the propagation of uncertainty. Monte Carlo is not the only option for this, in fact there exist a number of other procedures that are also conventionally utilized for the propagation of uncertainty, such as differential analysis, response surface methodology, RSM, the Fourier amplitude sensitivity test, FAST, and the closely related Sobol variance decomposition, and fast probability integration, FPI.

Amongst all, the Monte Carlo technique constitutes the most widely used approach to sensitivity analysis in PRA studies, and it was adopted for this work because of its conceptual simplicity.

When viewed in a form of a vector, a Monte Carlo sensitivity study involves a vector of uncertain model inputs:
\[
\tilde{x} = [x_1, x_2, ..., x_n], \quad \text{Eq. 18}
\]
where each \(x_i\) is an uncertain input and \(n_i\) is the number of such inputs, and a vector of model predictions:
\[
\tilde{y} = \tilde{f}(\tilde{x}) = [y_1, y_2, ..., y_n], \quad \text{Eq. 19}
\]
where \(\tilde{f}\) is a function used to represent the model under consideration, each \(y_i\) is an outcome of evaluating the model with the input \(\tilde{x}\), and \(n_o\) is the number of the possible outcomes, which for our purposes will be equal to the number of systems which could lead the plant to failure.
To carry out uncertainty and sensitivity analyses, the uncertainty in the elements of $x$ must be characterized. For this work, the uncertainty in the elements of $\tilde{x}$ has been shown it is assumed to be characterized by a sequence of input distributions

$$ID_i = 1, 2, ..., n_i,$$

Eq. 20

which are used to characterize the uncertainty in each input $x_i$, where ID$_i$ is the distribution assigned to $x_i$, these distributions, as already mentioned in this Section, are most of the time developed through an expert review process.

The LHS procedure is used to extract a sample:

$$\tilde{x}_k = [x_{1k}, x_{2k}, ..., x_{nk}] \quad k = 1, 2, ..., M,$$

Eq. 21

from the population of $x$'s with the distributions in Eq. 19, where $M$ is the size of the sample. Evaluation of the model under consideration with the sample elements $\tilde{x}_k$ in Eq. 21 then creates a sequence of outputs of this type

$$\tilde{y}_k = \tilde{f}(\tilde{x}_k) = [y_{1k}, y_{2k}, ..., y_{nk}] \quad k = 1, 2, ..., M,$$

Eq. 22

where each $y_{jk}$ is a particular outcome of evaluating the model with $\tilde{x}_k$. The pairs

$$(\tilde{x}_k, \tilde{y}_k) \quad k = 1, 2, ..., M,$$

Eq. 23

constitute a mapping from model input $\tilde{x}_k$ to model output $\tilde{y}_k$ that can be explored with various sensitivity analysis techniques to determine how the individual analysis inputs contained in $\tilde{x}$ (i.e., the different fundamental parameters involved in the analysis) affect the individual analysis outcomes contained in $y$ (i.e., the single systems' failure probabilities). In this work these analyses are the logistic regression, the classification trees, and examination of scatterplots.

Although techniques based on regression analysis and correlation analysis are often successful in identifying the relationships between model input and output embedded in the mapping in Eq.23, it must be pointed out that they may fail to identify well-defined, but nonlinear, relationships. In these cases, other sensitivity analysis techniques would be needed to identify the desired patterns in the mapping of Eq. 23. These other techniques are not discussed here but their inclusion in the present methodology could actually be the objective of further developments of this work.
2.5.3 Statistical Indicators

This study has attempted to maintain a statistical approach throughout the whole methodology. A natural objection to this type of analysis, of using the same method during the entire process, is that the observations deriving from its applicability could be flawed or partial. In reality, the method retains its fairness thanks to the inclusion of different statistical techniques, which are in turn based on various tests or algorithms; combined, these features assure that results be verified through cross-validation. This is accomplished by means of two statistical validation techniques applied in LR mode to classifier outputs: p-value and Wald value. Following is a brief description of their main functions.

**P-value**

The p-value test, on the other hand, is handy for its ability to evaluate the probability of a certain event, or of an extreme set of values. In classical statistics, it is used to calculate the probability of the null hypothesis being true. Generally, one rejects the null hypothesis if the p-value is smaller than or equal to the significance level, often represented by the Greek letter $\alpha$ (alpha). If the level is 0.05, then the results are only 5% likely to be as extraordinary as just seen, given that the null hypothesis is true.

**Wald Statistics**

In order to test the relevance of the results, statistic techniques such as the Wald and p-value tests, where employed. The likelihood ratio test can also be used to test whether an effect exists or not. Usually the Wald test and the likelihood ratio test give very similar conclusions (as they are asymptotically equivalent), but very rarely, they disagree enough to lead to different conclusions: this is because the statistical significance is always somewhat arbitrary, as it depends on an arbitrarily chosen significance level. The other reason is that the Wald test uses two approximations (that we know the standard error, and that the distribution is chi-squared), whereas the likelihood ratio test uses one approximation (that the distribution is chi-squared) [F. Harrell, 2001].
2.5.4 The Software Used in the Simulations

Though the method is designed to be computationally easy, it was nonetheless necessary to rely on software tools in order to carry out calculations and simulations.

All models of the world were developed with Vensim®, GoldSim® and RELAP5-3D®. The combined use of Vensim®, GoldSim® for Examples I and III, was accomplished in order to further validate their respective outputs. This was done in order to reduce uncertainties intrinsic to the code.

As previously mentioned, LHS was used in order to select samples from the available simulations. For the case of Example I, this operation was accomplished through the use of Vensim® and GoldsSim® both of which have an LHS routine incorporated in their code.

Example II on the other hand, involved the use of a dataset provided by a specific T-H code, RELAP5-3D® which is the latest in the RELAP® code series developed at the Idaho National Laboratory (INL). This resulted in the necessity of using a subroutine in MatLab® specifically designed [by F. Mackay, 2007] to create LHS dataset suitable for RELAP5-3D®.

The outputs thus generated by the above codes were statistically analyzed with the aid of Insightful Miner®.
2.5.5 Data Interpretation and Final Remarks

The interpretation of the results of the last module in Figure 10 is very intuitive. If the model reaches an acceptable accuracy, between 80 and 100 percent, it means that the model’s responses can be trusted. This manner of determining the effectiveness of the overall method reflects on the credibility of the indexes used to establish the importance of the single parameters. The significance of each parameter when considered as an input is determined by its p-value: those parameters identified by values below the threshold value of 0.05 are to be considered important with regards to the final measured output, all others are to be excluded from the ensuing analyses. The Wald Statistics’ Test is then used as an additional validation for these conclusions. The next step is to rank the parameters on the basis of the absolute values given by the p-statistic. Finally the threshold zones, as given from the classification tree, are to be identified. The parameters are then linked back to the PEPSRs, so as to guide the designers into the process of reviewing the single features, and attempt to move those functioning points away from those identified, by the threshold analysis, as critical. In this stage, the visualization of the scatterplots could provide further insights and help in the interpretation of the results obtained.

The study conducted for this document, attempts to propose a method distinguished by: a coherent approach to the problem, given by the steady use of statistical tools throughout the analyses; a robust output, promoted by the continuous cross validation of each step (LR vs CT, Vensim vs Goldsim, p-value test vs Wald Test, etc); a simple interpretation of its results achieved through the extensive use of graphs and scatter-plots; its computational ease as perceived by the relatively small processing time (1~2 hours) as opposed to pre-existing methods (~16 hours).

The use of this methodology has been verified and illustrated by mean of the three examples reported in the next Section. Note that the iterations loops given in the diagram, which aim to improve the efficiency of the method, were not tested in this work.
Following are two examples which were used in order to both, illustrate the methodology and to prove its efficacy when applied to reliability studies.

The first example is purely illustrative and shows the GFR's core failure under an induced thermal shock. Despite the low occurrence of this accident, - due to the unlikelihood of a loss in the capacity to transfer heat from the fuel to the coolant - the example remains credible and provides interesting insights to the combined CT – LR technique. In particular, deciphering the scatter-plots resulting from sensitivity analysis provides relevant information which shall be discussed in detail.

The second example uses a high-level code, RELAP5-3D, to model a more realistic scenario: the transient phase of a Loss of Coolant Accident (LOCA) in a GFR. The apparent advantages of studying a more plausible accident are tempered by the necessarily reduced number of available samples to analyze.

The third example provides insights and clarifications on the Latin Hypercube Sampling method (LHS) used to collect the data, and it investigates the relationship existing between the results provided by the predictors and the way the dataset has been created for the LHS.
3. RESULTS

EXAMPLE I: Core Failure of the GFR as due to a Negative Heat Pulse

The model

The example is based on the GFRDYN_LUMPED model presented in the first part of this work. It will now be tested to check for variations in the number of failures as due to uncertainties in the fundamental parameters constituting the model.

In order to save computational time and to better illustrate the methodology, only a portion, or module, of the original model’s structure was selected (the core sub-model) and it is shown in form of a Vensim sketch in Figure 11. The core’s sub-model shown in the figure has been “disconnected” from the remaining ECCS-HXC sub-model and therefore it is exogenous in mass flow rate and inlet temperature of the core.

![Vensim Sketch of the molecule describing the GFR core under the selected transient.](image-url)
The Accident and the Computational Time Required Running the Scenarios

A pulse heat is set to occur for an interval of 1000 seconds at the 450 seconds mark from the selected initial time. The pulse simulates a loss in the capacity of the fuel to transfer heat to the gas within the core. As previously mentioned, the mechanism that would generate the hypothesized loss in the capacity of heat transmission, is not provided here; the aim is to present an accidental scenario, which would be capable of calculating the probability of failure for the system, in a reasonable computational time, and with a sufficient number of realizations. The number of runs provided for this case has been settled to be 10,000, so around 10 times more than those completed for Example 2. Note that the computational time required to obtain the 10,000 realizations in this example is of approx 1 hour, while by running the same example with the complete GFRDYN_LUMPED model or with the RELAP5-3D code of Example 2, it would respectively be of 24 hours and 160 hours. The comparison in terms of computational times is possible by taking into account that the two sets of simulations were executed by means of the same sampling algorithm (Latin Hypercube). The final time for both simulations is 10,000 seconds, with a time-step of less than 1/100 second. Figure 12 shows the temperature of the core during one realization.

![Figure 12: Temperature of the GFR core during a LOCA as obtained by the core's sub-model. Reference Case (bottom red line) and Forced Pulse Heat Case (upper blue peaked line).](image)

4 Note that now, the probability of failure of the plant is doubly conditioned (given that a LOCA take place and that an interruption of the transmission between fuel and gas occurs in the gas). Therefore the probability of failure of the cladding provided in this section does not represent the real failure probability of the plant if we do not multiply the result for the two conditioning events.
**Uncertainties Quantification**

The model, shown in Figure 11, has been tested by varying three fundamental parameters, each with a different source of uncertainty. Distributions of Figure 13 were chosen arbitrarily but within acceptable ranges:

1. The specific heat of the carbon dioxide, \( C_p \) \( \text{gas} = \) input uncertainty;
2. The heat transfer coefficient of the core, \( H_T[C\text{ore}] \) = model uncertainty;
3. The initial temperature of the core, \( T_{\text{Fuel 0}} \) = initial / boundary conditions uncertainty.

![Figure 13](image)

**Figure 13:** Input dataset used to model the \( H_T[C\text{ore}] \), \( C_p \) and Initial temperature in the fuel.

The table below summarizes the values for the selected distributions with respect to the reference case (Best Estimate value, BE, or most likely value).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SHAPE</th>
<th>Range</th>
<th>Best Estimate</th>
<th>Mean</th>
<th>Std dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_p ) (J/(K-kg))</td>
<td>TRIANGULAR</td>
<td>200-600</td>
<td>536</td>
<td>445.33</td>
<td>87.717</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>( H_T[C\text{ore}] ) (W/K/m^2)</td>
<td>LOGNORMAL</td>
<td>50-600</td>
<td>100</td>
<td>300</td>
<td>100</td>
<td>1.037</td>
<td>4.9713</td>
</tr>
<tr>
<td>Initial Temp Core ( K )</td>
<td>UNIFORM</td>
<td>450-650</td>
<td>600</td>
<td>550</td>
<td>57.735</td>
<td>0</td>
<td>1.8</td>
</tr>
</tbody>
</table>
Resulting Output from the Dynamic Model

The output resulting from LHS, and expressed in terms of Probability Density Functions (PDF) for the core average and for the hottest channel, are shown in Figure 15. The corresponding realizations, in form of single runs are given in Figure 14. The failure criterion for the plant is to pass the boundary limit given by the maximum allowable temperature of either the cladding, 1473 K - as shown by the upper limit line of Figure 14 - or of the fuel, 2200 K.

Figure 14: Temperature of the cladding in 200 realizations and deterministic cladding limits.

The final probability of failure for the plant would be given by the intersection of the single failure probabilities of the two distinct systems (cladding and fuel). In this example no evidence of the failure of the fuel was found and therefore considered as a single failure problem.

Figure 15: Output distributions of the fuel temperature for the average and hottest channels of the GFR.
Creating the Dataset for the Prediction Models

The results, in terms of failures, obtained from the core of the GFRDYN_LUMPED model, have been transformed into a logic response which equals a 1 when the single realization curve passes the boundaries given by the adopted failure criterion, or a 0 otherwise. In order to create the dataset for the classification models, the obtained response must be coupled with the input deck given by the LHS algorithm. Thus, the second step is to arrange on the same row of a matrix the 3 values sampled from the characteristic distributions of Figure 13, followed by the binary failure response associated with that run. This procedure is repeated for a number of rows equal to the number of realizations sampled from the dynamic model. The so constructed matrix, which in this case is made of 4 columns and 10,000 rows, now constitutes the dataset used as input for the classification models. Table 7 shows the first 20 realizations of the dataset which was obtained by following the explained procedure. The dataset is then going to be processed by the classification model which will associate the input columns to the response column by mean of an algorithm which depends on the nature of the classifier adopted to conduct the analysis.

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Amplitude Pulse</th>
<th>Cp gas</th>
<th>HTC Core</th>
<th>T core 0</th>
<th>T Counter Cladding</th>
<th>response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>547.5</td>
<td>327</td>
<td>281</td>
<td>607.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>508.9</td>
<td>228.6</td>
<td>499.4</td>
<td>554.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>469.7</td>
<td>360.6</td>
<td>373.4</td>
<td>593.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>430.3</td>
<td>505.8</td>
<td>149.4</td>
<td>638.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>446.3</td>
<td>233.8</td>
<td>397.8</td>
<td>605.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>577.9</td>
<td>303.4</td>
<td>411</td>
<td>451.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>563.3</td>
<td>285.8</td>
<td>499.8</td>
<td>519.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>415.7</td>
<td>208.6</td>
<td>162.2</td>
<td>494.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>529.3</td>
<td>553.8</td>
<td>201.4</td>
<td>575.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>521.3</td>
<td>383</td>
<td>423</td>
<td>510.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>502.5</td>
<td>489.8</td>
<td>313</td>
<td>465.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>539.7</td>
<td>599.8</td>
<td>441</td>
<td>609.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>460.3</td>
<td>340.6</td>
<td>263</td>
<td>456.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>578.5</td>
<td>478.6</td>
<td>155.8</td>
<td>632.7</td>
<td>241.208</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>412.1</td>
<td>415.4</td>
<td>330.2</td>
<td>490.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>563.9</td>
<td>535.4</td>
<td>194.2</td>
<td>523.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>595.3</td>
<td>280.6</td>
<td>427</td>
<td>521.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>476.5</td>
<td>393.4</td>
<td>335.4</td>
<td>537.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>538.9</td>
<td>336.6</td>
<td>101.8</td>
<td>498.5</td>
<td>568.579</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>406.5</td>
<td>386.6</td>
<td>341</td>
<td>581.9</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Results from the Classification Models

By entering the dataset acquired from the dynamic simulator into the classification model of Figure 10, we can start exploring the sensitivity of the model to previously selected parameters. For the case of the logistic classifier, the model would now have the form of the equation:

\[ Y = \beta_0 + \sum_{i=1}^{p} \beta_i X_i + \varepsilon = \beta_0 + \beta_{\text{HTC Core}} + \beta_{\text{T Fuel}} T + \beta_{\text{CP Gas}} + \varepsilon. \]  

Eq. 24

The logistic classifier returns the importance of the single parameters concurring to the output. As shown in Table 8, the parameter’s estimates are given in the first column along with their standard errors. The standard error for an estimate is a measure of its variability. If the standard error for a coefficient is small in comparison to the magnitude of the coefficient estimate, then the estimate is fairly precise. The third and the fourth column describe the results from the t-statistic test and their corresponding p-values. The p-value for each t-statistic indicates whether the corresponding coefficient is significant to the model. In general, if the p-value is less than 0.05 the t-statistic is greater than 1.96; this result suggests that the coefficient is indeed significant. The Term Importance of Table 8 shows the weight of each variable in the model. In general, this type of table is used to confirm and validate the results coming from LR estimation. For categorical variables the Wald statistic provides another measure of how much the selected variables contribute to the model. The probability of the Wald statistic is based on a Chi-squared approximation which usually is in agreement with the results provided by the p-test (see also section 2.5.3).

Based on the above explanations, Table 8 confirms the expected absolute importance of the first parameter, HTC [Core], which dominates the dependent binary variable, failure of the cladding. The other parameters are clearly above the p-test cut-off value of 0.05 and thus confirm their total lack of influence on the failure of the cladding during the analyzed transient. The result is confirmed by the value of the Wald statistic for the HTC [Core] parameter, which is two orders of magnitude higher than any of the remaining fundamental parameters.

---

5 The Wald statistic for a variable is the square of the t-statistic. The Wald statistic is useful for categorical independent variables where there is a coefficient for each class level and the coefficient t-statistic shows the significance of each class level, making it difficult to assess how the variable contributes to the model as a whole.
The next step, as given from step 7.c of the milestone list of section 2.4, is to exclude from the analysis the parameters which are not statistically significant and rank the remaining ones sorted by mean of their p-values.

| Coefficient Estimates | Term Importance |
|-----------------------|----------------
| Variable | Estimate | Std.Err. | t-Statistic | Pr(|t|) | Wald Statistic | DF | Pr |
| (Intercept) | 65.8 | 6.05 | 10.88 | 0 | - | - | - |
| HTC[Core] | -0.42 | 0.04 | -11.99 | 1.17E-32 | 143.67 | 1 | 0 |
| T fuel 0 | 3.70E-04 | 3.73E-03 | 0.1 | 0.92 | 2.3 | 1 | 0.13 |
| Cp gas | 0.007 | 2.42E-03 | -1.52 | 0.13 | 0.01 | 1 | 0.92 |

Table 9 provides the misclassification matrices for both the methodologies and the output agreements between the two classification techniques of LR and CT. The logistic predictor succeeds in capturing the failures - labeled as 1’s - with an agreement equal to 148/155 = 95.5 % while the classification tree shows an even higher agreement of 98.7%.

| Table 8: Summary Statistics from the Logistic Classifier |
|---------------------------------|----------------
| Coefficient Estimates | Term Importance |
| Variable | Estimate | Std.Err. | t-Statistic | Pr(|t|) | Wald Statistic | DF | Pr |
| (Intercept) | 65.8 | 6.05 | 10.88 | 0 | - | - | - |
| HTC[Core] | -0.42 | 0.04 | -11.99 | 1.17E-32 | 143.67 | 1 | 0 |
| T fuel 0 | 3.70E-04 | 3.73E-03 | 0.1 | 0.92 | 2.3 | 1 | 0.13 |
| Cp gas | 0.007 | 2.42E-03 | -1.52 | 0.13 | 0.01 | 1 | 0.92 |

| Table 9: Classification Agreement Table |
|---------------------------------|----------------
| Input Node - Logistic Prediction (8) | Input Node - Classification Tree Prediction (9) |
| Misclassification Matrix | Misclassification Matrix |
| Predicted | Totals | Predicted | Totals |
| Observed | 0 | 1 | 4845 | 0 | 1 | 4845 |
| 1 | 7 | 148 | 155 | 1 | 2 | 153 | 155 |
| Totals | 4850 | 150 | 5000 | Totals | 4842 | 158 | 5000 |
| % Agree | 100.00% | 95.50% | 99.80% | % Agree | 99.90% | 98.70% | 99.90% |
| Positive Category - 1 | Positive Category - 1 |
| Recall | Precision | F-Measure | Recall | Precision | F-Measure |
| 95.50% | 98.70% | 97.00% | 98.70% | 96.80% | 97.80% |
The precision in the two measures is deemed as accurate, because higher than 95% in both the models and for each of the tests executed (lower tables of Table 9). Therefore, any conclusion based on the predictions returned by the CT or LR classifiers are to be considered reliable. Similar conclusions about the role played by the single parameters can be argued by looking at the results coming from the misclassification matrix of the classification tree in Figure 16. However, the classification tree provides extra information on the values around which the parameters - deemed important by the logistic classifier - start to show their influence on the dependent variable ‘Failure of the Cladding’. Figure 16 shows the structure of the classification tree: the splits in the tree structure provide the threshold values for the parameters in regards to the failure of the cladding. The nodes in the hierarchical view show the distribution of the classes in that node in a colored rectangle. The bottom right panel of Figure 16 reports the expanded, hierarchical view of the tree. The top panel displays the tree in form of a dendrogram: the tree is drawn such that the depth of the each branch is proportional to the change in the fitting criteria between the node and the sum of the two children nodes. This provides a quick visual of the importance of each split.

In this case, the CT returns, as expected from the results of the LR, that the only significant parameter is ‘HTC[Core]’ but it also provides the extra-information of the value around which ‘HTC[Core]’ is going to dominate the failure mechanism. The threshold values which lead the cladding to fail, are in this case two, 157.14 and 153.84 W/(K*m²), and they are
identified directly by the tree structure and can be easily visualized. The tree in this case confirms the importance of ‘HTC’ and thus validate the results coming from the LR classifier.

Analysis of the Scatter Plots

Finally the scatter plots of Figure 17, Figure 18 and Figure 19, provide insight on the potential of the method described here; the parameters identified as important because relevant from a statistical point of view, have merits in leading the plant to failure. On the other hand, the parameters which are non-relevant do not show any particular pattern in the scatter plots. This can be easily observed in Figure 17 where by plotting the two “irrelevant” parameters ‘Cp gas’ and ‘T fuel 0’ we do not find evidence of patterns for the failures occurring in the plant. More interesting conclusions can be deduced by observing Figure 18 where the scatter plots of ‘Cp gas’ versus ‘HTC[Core]’ show a lateral left stripe where failures are definitely most likely to occur. Figure 19 combines the information gained by the tree with the thresholds obtained through use of a CT and identifies the values which confine the “failure strip” from the remaining zones of the scatter plot: the two vertical bold lines are delimiting the two values of 157.14 and 153.84 measured in W/(K*m²).

![Figure 17: Scatter plot of the Specific heat for CO₂ versus Initial Temperature of the fuel.](image-url)
Figure 18: Scatter plot of the specific heat for CO₂ versus Heat Transfer Coefficient in the core.

Figure 19: Scatter plot of the initial temperature of the fuel versus the heat transfer coefficient in the core.
Findings

In conclusion, it must be pointed out, that despite the proposed example did not seem to be particularly engaging because of its easy nature, the proposed analysis allowed to:

- Efficiently identify the parameters which weighted the most in the determining of the final failure probability of the plant. This was done by using as a criterion of classification the individual statistical significance of the parameters through the response of a logistic model.

- Evaluate the statistical significance of the dataset in use as obtained by both the classification techniques through the evaluation of the capability of prediction of the models given by their misclassification matrixes. In this case, in fact, both models provided evidence of the absolute importance played by the fundamental parameter ‘HTC[Core]’.

- Identify and visually present the zones in which the parameters are most likely leading the plant to failure. This was done by means of a classification tree which determines the thresholds lines used to mark the failure regions on the scatter-plots returned by LHS.

Note that the number of possible scatterplots obtainable from LHS is equal to the number of pairs of the dependent variables in the model. In our example we selected three dependent variables which returned only three scatterplots. As the complexity of the model passes to a higher number of parameters, the number of fundamental parameters which can affect the measured output, in turn increases too. Therefore realistic models, with higher degree of complexity, could lead to the examination of huge number of scatterplots. A way to discriminate and select among all the possible scatterplots is to choose only those scatterplots which consider two statistically, or at least one, significant parameters. This could be done, in the current example, by excluding Figure 10 from the analysis. This aspect will be remarked in Example II.

This final observation leads to conclude that the p-test measure by the LR model also:

- Provides a criterion to select those scatterplots which most likely are going to contain information useful to individuate patterns in the parameters leading the system to failure. The criterion is based on the p-test responses of the LR model.

The following example will check the above conclusion with a more realistic case and extending the analysis from the core to the entire GFR plant.
EXAMPLE II: Leakage from a Check Valve in a Helium Cooled GFR

The Model

This example provides a more realistic application of the methodology explained in this Section. The dataset in use has been adopted from a recent work developed at MIT on a 2-DHR-loop GFR helium cooled plant [F. MacKay et al., 2007]. The simulations presented here were carried out with the RELAP5-3D code. A set of six input parameters were selected to be propagated throughout the model. The selection was made based on sensitivity analyses done on several parameters and on expert opinion. Table 10 reports these parameters along with their probability distribution functions.

<table>
<thead>
<tr>
<th>Input Parameter</th>
<th>Distribution</th>
<th>Extreme Values</th>
<th>Units</th>
<th>Extreme Values Correspond To Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td></td>
</tr>
<tr>
<td>Core Roughness</td>
<td>Lognormal</td>
<td>1.00E-05</td>
<td>1.00E-04</td>
<td>[m]</td>
</tr>
<tr>
<td>DHR1 Check Valve Leakage</td>
<td>Exponential</td>
<td>-</td>
<td>0.5</td>
<td>[kg/s]</td>
</tr>
<tr>
<td>DHR2 Check Valve Leakage</td>
<td>Exponential</td>
<td>-</td>
<td>0.5</td>
<td>[kg/s]</td>
</tr>
<tr>
<td>Heat Transfer Coefficient, Containment Structures</td>
<td>Lognormal</td>
<td>0.5</td>
<td>2</td>
<td>[Dmnl]</td>
</tr>
<tr>
<td>Moment of Inertia of the Shaft</td>
<td>Normal</td>
<td>656.88</td>
<td>802.86</td>
<td>[kg*m2]</td>
</tr>
<tr>
<td>Core Heat Transfer Coefficient</td>
<td>Lognormal</td>
<td>0.4</td>
<td>1.6</td>
<td>[Dmnl]</td>
</tr>
</tbody>
</table>

The Selected Accident

A LOCA on the cold leg of the PCU was chosen as the initiating event of the analysis. The system, originally designed to survive a 500 cm² break, was evaluated for a small break of 5 cm² because of its higher incidence (frequency: ~10⁻³ per year) over a larger one.
**Failure Modes and Sampling Techniques**

After reviewing the parameters of the system at various points, two failure modes were selected for the reliability evaluation of the system: cladding damage in the hot channel and structural failure in the hot legs of the DHR loops. The analysis reported in this example refers to the second failure mode, while the first failure mode is not analyzed here because of the low number of failures relative to the number of trials available (the original dataset had in fact only 13 failures of the cladding which were independent and thus not correlated to the second failure mode). A limit of 1123 K for the maximum temperature in the hot legs (MTHL) was set as the failure criterion for this mode.

In order to reduce the number of trials, also in this example Latin Hypercube Sampling (LHS) was employed to propagate the uncertainties in the input parameters throughout the model. Figure 20 shows two sets of 128 realizations for each of the two DHR loops analyzed, for a total of 256 realizations. The horizontal dashed line corresponds to the DHR temperature limit for the materials used in the GFR helium plant. There were 26 curves that exceeded the limit, so the overall probability for this failure mode, conditional on the LOCA, is 26/128 = 0.2.

The decreasing curves correspond to the DHR loop with the check valve closed. In the initial conditions of the simulation, the MTHL is set as if there was no leakage in either loop. This is the worst possible initial condition because the MTHL is smaller with leakage than without. This is due to the fact that the leaking helium comes from the downcomer and passes through the water-cooled heat exchanger where it cools down substantially. Additional explanations for the behavior observed in the graph can be found in [F. MacKay et al., 2007].
Figure 20: Realizations for the Maximum Temperature in the Hot Leg of DHR Loops [F. MacKay et al., 2007].

Results from the Classification Models

The dataset was used as an input to the logistic and tree classification models. The two models were tried with 70% of the 128 simulations available. The results are listed in Table 11 and Table 12 where the summary statistics and the compared classification agreements are reported. From Table 11 it can be inferred that the only two parameters leading the plant to failure are given by ‘Max leakage’ and ‘HTCS’ as provided by the value of their p-test. The values of the Wald statistics are in agreement with the results coming from the p-test.

| Variable              | Estimate | Std.Err. | t-Statistic | Pr(|t|) | Wald Statistic | DF | Pr  |
|-----------------------|----------|----------|-------------|-------|----------------|----|-----|
| (Intercept)           | 1.56     | 4.34     | 0.36        | 0.72  | -              | -  | -   |
| Roughness             | 522.99   | 7,571.49 | 0.07        | 0.95  | 7.58           | 1  | 0.01|
| Max leakage           | -6.61    | 2.4      | -2.75       | 0.01  | 4.09           | 1  | 0.04|
| HTCS                  | 0.49     | 0.24     | 2.02        | 0.05  | 1.09           | 1  | 0.3 |
| HTCC                  | 0.23     | 0.31     | 0.74        | 0.46  | 0.55           | 1  | 0.46|
| Rupture Form Factor   | -2.96    | 2.84     | -1.04       | 0.3   | 4.77E-03       | 1  | 0.94|
| Shaft Moment Inertia  | 3.08E-04 | 4.52E-03 | 0.07        | 0.95  | 4.65E-03       | 1  | 0.95|
Table 12 does not deny this result, but instead highlights the very low capacity of the model to predict failures as documented by the values of agreement: 27.3% and 45.5% respectively for the logistic and the tree model. This result is reflected in the uncertainty of the obtained failure probability for MTHL: the 26 curves indicating failures shown in Figure 20 are subjected to an uncertainty which is proportional to the observed percentage of agreement of the predictor model in use. The low values observed in the percentage to predict failures from both models is because predictions are based on the values used to test the dataset and thus equal to just 30% of the total samplings. Therefore the numbers of failures analyzed by the LR and CT models are 11 out of a total of 38 samplings used to validate the model. This problem could be solved by either, increasing the number of samples, or by increasing the validation dataset. The former solution stresses the need of a higher number of realizations in order to identify possible patterns in the parameters leading the plant to failure. The latter solution indirectly lead to the same conclusion, because in order to avoid oversampling effects (which is the loss of the capability to predict patterns which are different from those given by the training dataset) it is required to increase the number of simulations in both the validation and the training sets.

<table>
<thead>
<tr>
<th>Table 12: Classification Agreement Table</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input Node - Logistic Prediction (8)</strong></td>
</tr>
<tr>
<td>Predicted</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Totals</td>
</tr>
</tbody>
</table>

| **Observed** | **Overall** | **Observed** | **Overall** |
| 0 | 1 | 0 | 1 |
| % Agree | 96.30% | 27.30% | 76.30% | % Agree | 92.60% | 45.50% | 78.90% |

| **Positive Category - 1** | **Positive Category - 1** |
| Recall | Precision | F-Measure | Recall | Precision | F-Measure |
| 27.30% | 75.00% | 40.00% | 45.50% | 71.40% | 55.60% |

Despite the mentioned scarcity of points (measures), the classification dendrogram portrayed in Figure 21 gives the information needed to confine the zones where failures are most
likely to occur. The tree is obtained by considering the statistically significant parameters obtained by the logistic regression. Therefore Figure 22 is acquired by plotting the dependent output variable, failure of the hot leg, distinguished by its binary connotation, on a scatter-plot as a function of the two variables identified as statistically significant from the p-test. The vertical line corresponds to the threshold value of 0.63 for the maximum leakage as given by the second split of the dendrogram of Figure 21. The horizontal line represents the threshold limit for the HTCS parameter as obtained from the first split of Figure 21 and equals to 0.93. It can also be noticed that for values of ‘max leakage’ below 0.63, the tree splits into two distinct clusters of failures identified by the threshold value of 0.22 for ‘max leakage’. Due to the low number of samples available, it is not given to know if these two separate zones could merge into a unique failure region or remain divided into two different clusters. It is then spontaneous to ask if the results coming from a 128 datasets could be considered definitive or not for the CT classifier. This will be seen in Example III which reports a series of investigations aimed to determine the dynamics of the failures as obtained by increasing the number of realizations along with other information regarding the uncertainty in the measured probability of failure. For the moment, the conclusion here is that, once again the sampling size reveals to be of fundamental importance to substantiate some of the behaviors identified by the combined use of LR and CT models.

Figure 21: Classification tree resulting from LHS sampling of six parameters in the Helium GFR.
Findings

As before, this example strengthens with evidence the competence of the proposed method when dealing with responses coming from sensitivity analyses computed by varying multiple parameters and then measuring the effects of their uncertainty propagation during a transient.

- The possibility of its extension to a problem with a higher degree of difficulty has been shown by providing an efficient and fast way to screen the output as identified from the selected failure mode.

- In particular, the method showed through clear evidence that classification models can efficiently be applied to predict the probability of failure of a given system only when a sufficient number of samplings is available.

- The logistic regression model revealed to be a suitable tool to discriminate the parameters concurring to the final probability of failure of the system analyzed.
In addition, the visual representation, in form of scatterplot of the variables screened by the p-test criterion, agreed with the results coming from the classification tree (despite its low prediction capability).

Note that this example is using 6 fundamental parameters and this would result in 15 possible scatterplots, as obtainable form the different pair combinations of these parameters, despite that, just 1 scatterplot has been selected; this has been done by following the p-test criterion identified in the previous example and thus, in this case, by including in the visual analysis only those combinations arising by coupling the parameters passing the p-test. This criterion offers a more structured visualization of the results than the one that could be obtained, as done in Example I, by portraying in the scatterplots only one significant variable. This is in fact allows to identify two-dimensional patterns of the failures and therefore to capture behaviors which are less intuitive than the ones offered by screening the behavior of a single parameter.

Further research is needed to identify the relationships between the number of realizations and the accuracy of the returned probability of failure, PF, and the resulting patterns of the failures among all the possible system states. Example III illustrates a possible approach to explore these relations by means of a simple explanatory model.
EXAMPLE III: Explanatory Example about LHS

Introduction

The example is introduced in the form of a set of observations commented on a case by case basis. These comments are a consequence of calculations performed on either the LHS dataset obtained for Examples I and II, or on separate numerical examples which also sustain the findings of the first two examples.

To support the considerations reported in the current example, a broad review of the basic definitions shared by authorities in the field, fundamental statistic theorems, and general properties of LHS, has been necessary.

The two previously mentioned analyses – Example I and Example II- differ, besides in their different degree of complexity, mainly in the number of available simulations, 10,000 and 128 respectively. Their relative conclusions, from a statistical perspective, are more or less prejudiced by this factor, and as a consequence, the accuracy of the measured PF is proportional to the value of $M$ too. The following example has, therefore, been named “the explanatory example” because its aim is to provide a correct formalism for the definition of the probability of failure, PF, as it varies with the number of realizations $M$. This last factor as a matter of fact influences not only the credibility of the results, or the uncertainty on the measured PF, but also the capacity to accurately model the pattern followed by the failures when $M$ increases. Evidence of both these effects (and of the relationship between the two), was indirectly shown by comparing the first two examples, and will be investigated here through further analyses. Before venturing into this matter, a basic definition for the PF is required and will be recalled later on in this work. Let $A_i$ be a binary variable that is equal to 1 when the i-th realization leads to failure, and equal to 0 otherwise. If $M$ is the total number of realizations, then the probability of failure of a given system $S$ can be defined as:

$$PF_S = \frac{\sum_{i=1}^{M} A_i}{M}$$

Eq. 25

Some basic definition about the LHS need also to be pointed out and are given along with the first of the observations driving the explanation of the actual exercise.
**Latin Hypercube Properties**

Two among the main theoretical benefits in the use of LHS, are that:

- Its estimates are unbiased. This means that LHS reduces the variance of the results as the number of trials increases.
- that the expected variance is reduced, or asymptotically lower, than that obtainable from canonical Monte Carlo samplings [McKay et al., 1979].

The first statement is well accepted and evidence of it can be found in the literature. But if it is credible and proven that LHS reduces the variance, it remains yet to demonstrate how many trials are necessary for this reduction to be significant. This of course depends on: the nature of the problem, and thus from the presence of non-linearities in the model; the number and shape of the distributions selected; and, ultimately from the number of samplings available. A measure of the variance reduction can be obtained by plotting the relative variations of the PF, as a function of the increasing number of trials. Figure 23 compares this measure for the first 400 runs to the full number of runs, and it clearly shows the reduction of the variability for higher numbers of M. The first 400 runs present steep variations of PF which progressively smooth to zero, roughly, in the first 5 steps, corresponding to $5*20=100$ runs; the result is maintained for the remaining number of runs. Therefore the second statement is confirmed.

![PF Delta Variation - M = 10,000 vs. M = 400](image)

**Figure 23:** Measured PF deltas on the first 400 trials compared to the whole 10,000 sequence.
The first observation recalls Example I and refers to the PF obtained from 10,000 realizations.

**Observation I:** the stratified method of sampling provided by LHS effectively reduces the variance with an increase in the number of realizations and confirms the expected properties of the algorithm as expressed by the second statement.

![PF Measured every 20 Trials for M=400](image)

**Figure 24:** PF of the cladding for 2000 simulations. Data are collected every 20 trials.

The first statement asserts that the LHS is unbiased. As illustrated by Figure 24 and Figure 25, the probability of failure, obtained from the plant configuration of Example I, shows a systematic decrease of PF, even for a significantly high number of trials. This result is in contradiction with the expected convergence of the LHS after a low number of trials.

Despite this, the range of variation is not remarkable and does not affect the outcome (the returned probability of failure). The PF stays in a narrow corridor of values which, as shown in Figure 24, goes from just above 2.00E-02 to 5.00-E02 and that, for the case of a higher number of trials as shown in Figure 25, tries to stabilize around a value of 3.00E-02. Therefore, for a high number of simulations, the value around which the algorithm is going to converge can be inferred, though it cannot be accurately predicted, from the first few hundred trials. Note that these and the remark that follow, refer to the specific case studied in Example I.
Observation II: for a low number of trials the LHS returns an estimate value of the PF which, despite being representative of the actual PF predicted from a higher number of trials, has a low confidence level. Table 13 reports the first 140 samplings used to determine the probability of failure in Example 1 and shows that the first 60 samplings could position the result close to the expected value of 3.00E-02 corresponding to the convergence value of Figure 25.

This result implies that the value provided by a limited number of realizations, as in the case of the 128 runs of Example II, does not correspond to the final definitive value of the PF for that given system. The main finding is, therefore, that for a low number of trials the value returned by LHS is slightly biased and even though this effect has an apparent low impact on the final result of Example I, it has nonetheless some implications. These are reported in the next observations.

Observation II could play a bigger role by relaxing some of the assumptions of Example I, such as the shape and number of the distributions concurring to the final PF of the system, the nature of the problem under analysis, and the single failure criterion identified for the temperature of the cladding. In other words, when the complexity of the problem increases, the low confidence measured on the output can be amplified by those factors neglected by the initial assumptions on the problem. The next observations analyze the single factors in details.
Table 13: Probabilities of failure, PF, obtained from 140 Latin Hypercube samplings.

<table>
<thead>
<tr>
<th>Number of Trials M</th>
<th>Number of failures Nf</th>
<th>PF</th>
<th>delta PF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.00E+00</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>5.00E-02</td>
<td>5.E-02</td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>2.50E-02</td>
<td>-3.E-02</td>
</tr>
<tr>
<td>60</td>
<td>2</td>
<td>3.33E-02</td>
<td>8.E-03</td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>3.75E-02</td>
<td>4.E-03</td>
</tr>
<tr>
<td>100</td>
<td>4</td>
<td>4.00E-02</td>
<td>3.E-03</td>
</tr>
<tr>
<td>120</td>
<td>4</td>
<td>3.33E-02</td>
<td>-7.E-03</td>
</tr>
<tr>
<td>140</td>
<td>5</td>
<td>3.57E-02</td>
<td>2.E-03</td>
</tr>
</tbody>
</table>

**Observation III:** the value of the probability of failure of a system, PF, obtained by LHS sampling is the reference value for a distribution whose shape in turns depends on the distributions of the selected uncertain parameters from which we sample. Thus, the value of the PF as given in Eq. 25 is actually a reference value for a probability density function of PF. That reference value, in general, is not the most likely value or average value of the PF distributions if not under special hypotheses. These hypotheses are summarized in this and in the next two observations.

**Observation IV:** the value of the probability of failure of a system, PF, depends on the number of parameters the LHS is sampling from, relative to the number of trials available.

**Observation V:** all of the above observations depend on the specific nature of the physical problem under analysis as well as from the complexity of the overall model used to describe the system under study.

In order to give evidence of observations from III through V a simple model has been built. In other words, observations III through V are illustrated by means of a set of auxiliary calculations executed on a proposed mathematical model. Although any simulation problem could have been used to provide evidence for these observations, the use of the simple model was preferred because of the lower computational time. Secondly, results can be easily extended to structured designs and real codes.
Model Description

The model is made up by the weighted sum of five parameters. The returned sum is the measured input, which, by construction, is linear with regards to the five fundamental inputs. A second model was obtained by squaring the results coming from the previous model. This results in a non-linear model. LHS is used to sample the five parameters and build the outputs under different set of the hypotheses that were remarked in the last three observations. The results surfacing from the different sets of simulations support the observations.

Common to both models are the input datasets, summarized in the table below, and their weights, set to be all equal to 0.5. Results are given in the form of three different sets of simulations, each of them capturing different properties and connotations of the studied problem.

<table>
<thead>
<tr>
<th>Dependent Variables 1: Dataset Uniform</th>
<th>Dependent Variables 2: Input Non-Uniform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inputs[input1]=RANDOM_UNIFORM[0,10]</td>
<td>Inputs[input5]=RANDOM_UNIFORM[0,10]</td>
</tr>
<tr>
<td>Inputs[input5]=RANDOM_UNIFORM[0,10]</td>
<td>Inputs[input3]=RANDOM_UNIFORM[0,10]</td>
</tr>
<tr>
<td>Inputs[input4]=RANDOM_UNIFORM[0,10]</td>
<td>Inputs[input1]=RANDOM_TRIANGULAR[0,10,50,60,70]</td>
</tr>
<tr>
<td>Inputs[input3]=RANDOM_UNIFORM[0,10]</td>
<td>Inputs[input2]=RANDOM_POISSON[0,10,5,0.1,0.1]</td>
</tr>
<tr>
<td>Inputs[input2]=RANDOM_UNIFORM[0,10]</td>
<td>Inputs[input4]=RANDOM_NORMAL[0,10]</td>
</tr>
</tbody>
</table>

Model 1: Linear

\[ Y = w_1X_1 + w_2X_2 + w_3X_3 + w_4X_4 + w_5X_5 \]

Weights

\[ w_1 = w_2 = w_3 = w_4 = w_5 = 0.5 \]

Model 2: Non-linear

\[ Y = (w_1X_1 + w_2X_2 + w_3X_3 + w_4X_4 + w_5X_5)^2 \]

Weights

\[ w_1 = w_2 = w_3 = w_4 = w_5 = 0.5 \]

The dependent variable \( Y \) reported in the table, it will be shown, it can be imagined as the measured load of a T-H problem in a particular instant of time. By means of the analogy between this load and the loads of the two previous examples (the temperatures of the cladding and of the ECCSs’ structures) some interesting properties of the calculations, used by the established methodology of Section 2, can be derived and inspected.
Results

As a first experiment, LHS was performed on both models by varying the shape of the curve associated with the input distributions and by showing the returned outputs as a function of the number of trials, \( M \). The conclusions illustrated below are in support of Observation III and V.

Table 15: Distribution parameters for two different input datasets: uniform versus non-uniform.

<table>
<thead>
<tr>
<th>Trials</th>
<th>Output Linear</th>
<th>Median</th>
<th>Mean</th>
<th>StDev</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UNIF non UNIF</td>
<td>UNIF</td>
<td>non UNIF</td>
<td>UNIF non UNIF</td>
</tr>
<tr>
<td>20</td>
<td>12.625 10.22035122</td>
<td>10.319</td>
<td>12.5</td>
<td>3.980813 3.980813</td>
</tr>
<tr>
<td>60</td>
<td>12.83333397 10.70901012</td>
<td>10.3179</td>
<td>12.5</td>
<td>3.345966 3.345966</td>
</tr>
<tr>
<td>400</td>
<td>12.5 10.31872845</td>
<td>10.3024</td>
<td>12.5</td>
<td>3.237578 3.237578</td>
</tr>
<tr>
<td>10,000</td>
<td>12.49100018 10.3074913</td>
<td>12.5</td>
<td>10.2983659</td>
<td>3.246991 2.517422</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output Non Linear</th>
<th>Median</th>
<th>Mean</th>
<th>StDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNIF non UNIF</td>
<td>UNIF</td>
<td>non UNIF</td>
<td>UNIF non UNIF</td>
</tr>
<tr>
<td>20</td>
<td>159.390625 104.5316925</td>
<td>172.097</td>
<td>109.6435623</td>
</tr>
<tr>
<td>60</td>
<td>164.7795105 114.6829376</td>
<td>167.445</td>
<td>112.3515396</td>
</tr>
<tr>
<td>400</td>
<td>156.250351 106.4761887</td>
<td>166.732</td>
<td>112.4572372</td>
</tr>
<tr>
<td>10,000</td>
<td>156.0250854 106.2443848</td>
<td>166.793</td>
<td>112.3937607</td>
</tr>
</tbody>
</table>

Figure 26 compares the outputs obtained by the simulations which were carried out using sample size \( M \) values ranging from the very low value of 20 to the extremely high value of 10,000. This array enables to weigh the effect of the sample size on results; however, in practice, the typical values available to work with in standard reliability calculations fall in the lower range, from 60 to 400. The output is expressed in terms of the probability density function (PF) as it is measured at any given time. By recalling the simple mathematical definitions of the outputs given in this example, we treat them as time invariant variables. Therefore, we could translate the obtained curves into a proxy value for the load calculated in a typical T-H problem at steady state. The load is thus calculated first from an initial set of five identically uniformly distributed inputs, and then from generically distributed inputs, chosen as indicated in Table 14. The red curves of Figure 26 represent the loads given by the first set, while the blue curves those from the second set. The left and right columns are differentiated by the type of model used (respectively linear and non-linear).
Two distinct effects are clearly visible and then confirmed by the characteristic values reported in Table 15:

- The shape of the curves changes as a function of the mathematical definition given to the model. In addition, the curves shown in the left of the figure have a shape close to a normal distribution, while the right ones seem closer to a "lognormal" one.
- By comparing the blue and the red graphs in each plot, it can be seen how the curves shift to the right when going from a uniform input set to a generically distributed dataset. This effect is more obvious in the 10,000 samplings case.

Finally, by comparing the upper and the lower plots, it can be seen how the two effects are more marked for the cases with the higher number of simulations than for that with few samples.

The first effect is explainable by recalling the central limit theorem that states that the linear combination of a given set of probability density functions returns a normal distributed PDF regardless of the original shapes of such distributions [Art B. Owen, 1992]. While, for non-linear combinations of the same parameters, the shape of the resulting output is not, a priori, related to any pre-assigned shape.

The second effect is interesting for the implications it has in terms of failure probability. In fact, if we assume, as already mentioned, that the output distribution can be seen as a proxy for the load of a possible T-H problem, and that a deterministic failure criterion is assigned to describe the capacity of the system, the failure probability \(PF\) is then given by the right green areas of Figure 27. As shown in the figures, due to the shift generated in the load PDFs, the resulting probability of failure is increased as passing the non-uniform to the uniform dataset. The validity of the illustrated results is still functional to \(M\) and to the structure of the model.

Figure 27: Load distribution and Capacity for the linear and non-linear cases (left and right graphs) for \(M=10,000\).
In a second experiment, the probability of failure obtained by using identical failure criteria for the linear and non linear case and analyzed by means of the two different input datasets are reported in the following table. Results are reported as a function of the two extreme values of M equal to 20 and 10,000.

Table 16: Probability of failures as a function of M for the four different cases.

<table>
<thead>
<tr>
<th>M</th>
<th>20</th>
<th></th>
<th>10,000</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lin</td>
<td>Non-lin</td>
<td>lin</td>
<td>Non-lin</td>
</tr>
<tr>
<td>Non Uniform = NU</td>
<td>1.50E-01</td>
<td>5.00E-02</td>
<td>7.94E-02</td>
<td>5.00E-02</td>
</tr>
<tr>
<td>Uniform = U</td>
<td>3.50E-01</td>
<td>5.00E-02</td>
<td>3.27E-01</td>
<td>5.00E-02</td>
</tr>
<tr>
<td>total max PF</td>
<td>5.00E-01</td>
<td>1.00E-01</td>
<td>4.06E-01</td>
<td>1.00E-01</td>
</tr>
</tbody>
</table>

Table 16 shows clearly once again, that even in this proxy problem the values of the PFs for the two models (linear, Lin and non-linear, Non-Lin) change according to the shape assigned to the distributions of the fundamental input parameters. The most interesting result of the table is given in the further left column, where, when going from the uniform case, U, to the non-uniform one, NU, in the linear model, we observe that the value of PF varies by one order of magnitude. This result provides further validation for Observation III. Note that, once again, these results are sensitive to the number of trials M.

Along the same lines of the previous analysis, but adding a layer of complexity to the study, the case of a system with two possible failure components (as before, Linear and Non-Linear) is considered. Each of these components now has two different failure criteria which, as shown in Figure 20, are merely deterministic.

Table 17: Total PF for a two component system. Linear versus non-Linear for two different input sets and seeds.

<table>
<thead>
<tr>
<th>M</th>
<th>20</th>
<th></th>
<th>10000</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lin</td>
<td>Non-lin</td>
<td>lin</td>
<td>Non-lin</td>
</tr>
<tr>
<td>Lin-Non Lin Seed</td>
<td>4.00E-01</td>
<td>4.00E-01</td>
<td>2.00E-01</td>
<td>2.00E-01</td>
</tr>
<tr>
<td></td>
<td>4.90E-01</td>
<td>3.40E-01</td>
<td>4.82E-01</td>
<td>3.40E-01</td>
</tr>
<tr>
<td>max</td>
<td>6.00E-01</td>
<td>4.50E-01</td>
<td>3.00E-01</td>
<td>2.50E-01</td>
</tr>
<tr>
<td>min</td>
<td>4.00E-01</td>
<td>3.50E-01</td>
<td>5.00E-02</td>
<td>1.00E-01</td>
</tr>
</tbody>
</table>
Next, the final \( PF \)s of the system, along with their individual variations, is considered. LHS sampling is then used, by sampling with different seeds, in order to take into account the variability of the results as due to the combination proposed by the LHS algorithm. Observations VI and VII are useful in explaining the reasoning behind this approach.

Table 17 shows the final \( PF \) as given by intersection of the failures of the two single systems (recall that they are given by the linear and the non-linear models) minus the correlated failures (this aspect will be explained in detail with Observation VI). Due to the complexity of the table, the focus is to be given to its furthest left part, referred to 20 simulations and then extended and compared with the previous results. The purpose of this calculation is to show that by changing the shape of the initial distributions and the seed of the LHS algorithm, the variability of the measure on probability of failure of the overall system increases along all cases considered. The measure of such variability is given in the lowest part of the table where the maximum and minimum values for all the cases are reported: it can be seen that \( PF \) varies from 4.5E-01 to 5E-02 amongst all the four possible combinations considered and within the different values of \( PF \) provided by LHS. This result validates Observations III and V by, once again, showing that the variability of the measured \( PF \) changes as a function of the hypotheses done on the mathematical model of the world used, on the shape of the initial distributions and also on the criteria used to determine \( PF \). This last observation refers to the multiple criterion used to define the overall \( PF \): as can be seen by comparing Table 17 with Table 16, the range for the final \( PF \) changes and it is amplified by passing from a single to a multiple criteria. This means that small ranges of variability for a single component or system could manifest in wider ranges of variability affecting the overall system. So special care should be taken when studying problems which involve multiple failure criteria. Note finally that, once again, all these considerations are a function of \( M \) as given by extending these considerations to the remaining part of the tables illustrated here.

The next two following observations are used to provide further reasons of the choices behind this second experiment. Specifically, they clarify the motivations of the use of different seeds and of multiple failure criteria.
Observation VI: All of the observations made up to this point refer to the case of a single system with a single failure criterion. The probability of failure, \( PF \), as defined by Equation 25, is in fact an expression of the unreliability of the single system/structure/component/phenomenon, SSCP, measured in the plant and thus obtained from the assumption of there being a unique mode leading to failure. As shown in Figure 28, multiple failure criteria should be used whenever more than one output parameter is instrumental in leading the system to failure. Similarly, single components could fail in more than one mode, each of which associated with different failure criteria. Various failure modes and multiple failures could occur simultaneously or sequentially in a complex plant. This could be simulated by studying the intersection of all the binary responses of the LHS database returning a value of 1. This procedure allows automatically preventing the double counting of those failures which are correlated.

![Figure 28: Schematic of the GFR showing the plant subjected to multiple failures.](image)

Let \( F_i \) be the failure vectors associated to \( N \) failure modes over \( M \) realizations. Then, the mean probability of multiple failures is defined as:
Then, the mean probability of multiple failures is defined as:

$$\bar{PF}_M = \frac{1}{M} \sum_{i=1}^{M} F_i - \text{CorrF}$$

where, a correlated failure is a failure which is common and simultaneous to two, or more, systems of the same plant. This means that one way of obtaining the plant’s failure probability, $PF_M$, is to calculate the probability of failure of each individual component through the definition provided by Equation 25, then sum all the obtained values and finally subtract the probabilities which are derived from their correlated failures.

**Observation VII:** being $Y$ a general simulation estimate, conditioning $Y$ on the $(n!)^k$ equally likely sequences for the $k$ input values, it follows that:

$$\text{Var}(Y) = \sigma_{\text{seq}}^2 + \sigma_R^2,$$

where the first term accounts for the variance component of $Y$ due to sequence variability, while the second term reflects the remaining variance component of $Y$ as due to the set variability conditioned on the $(n!)^k$ sequences. Using LHS, the residual term $\sigma_R^2$ will be of the order of $o(n^{-a})$, with $a \geq 2$, while no matter the sampling methods $\sigma_{\text{seq}}^2$ is always of the order of $o(n^{-1})$.

This implies that, for LHS, the second term will decrease faster than the first, so that for a high number of trials the only variability in LHS will be given by the left part of the equation, thus from the sequence variability intrinsic to the algorithm [E. Saliby, 1997]. This variability, which represents the unavoidable part of variance imputable to the different values obtainable by sampling from the same set of parameters but with different seeds, can be attained by sampling different sets of simulations, each of size $M$, with a different initial seed; by changing the seed, for a low $M$, the variability of the output $Y$, which for our purposes is given by $PF$, can be shown. Thus, if, for example, computational limitations are present, it would be nonetheless possible to measure the uncertainty of $PF$ by sampling four sets of 32 trials each, instead of one single set of 128 trials. With approximately the same computational time it would also be possible to gather information about the uncertainty on the measure of $PF$. These last two observations lead to a reformulation of $PF$. 
Reformulation of the Probability of Failure

Based on the results yielded by these numerical examples, a revision of the formulation of \( PF \), as given in the equation provided at the beginning of this exercise, is necessary. This reformulation originates from the need to provide a measure of the uncertainty affecting \( PF \) in problems where the computational time represents a limit for the observations. Through this operation, it should be possible to obtain a measure of the variance of \( PF \), and therefore monitor it, for a low number of realizations \( M \). The formulation can be naturally extended to the definition of an overall failure probability of the plant, \( PF_M \).

The \( A_i \) variable is still 1 or 0 as it was defined in Equation 25 but the \( PF_s \)s are now given by the sum of the \( A_i \)'s weighted by the number of times they appear in a defined set of \( N \) runs with \( M \) samples each. Therefore, the probability of there being a failure in the system is averaged with the values coming from a set of \( N \) equi-probable multi-seeds realizations:

\[
P F_s = \frac{1}{N} \sum_{j=1}^{M} \left( \frac{A_i}{M} \right)_j
\]

Eq. 28

The probability of failure of the system, is now averaged by the number of runs, \( N \), and it expresses the mean value over them. This definition has a general connotation and thus is to be considered valid regardless of the assumptions made by the specific sampling method (in this case Latin Hypercube). Furthermore it expresses the random uncertainties that characterize the output parameter \( PF \). The variance on the output \( PF \) can be similarly obtained as:

\[
Var(PF_s) = \frac{1}{N} \sum_{j=1}^{M} \left[ PF_s - \left( \frac{A_i}{M} \right)_j \right]^2 = \frac{1}{N} \sum_{j=1}^{M} \left( \frac{N_f - N_{f,j}}{N_f} \right)^2
\]

Eq. 29

where \( N_f = PF_s \) and \( N_{f,j} = \left( \frac{\sum_{i=1}^{M} A_i}{M} \right)_j \)

Eq. 30

This final definition concludes the observations about the LHS method and the variables involved in the determination of its accuracy. It emerged clearly that the number of trials \( M \) remains a critical variable for the determination of \( PF \). To conclude this aspect, the next Section investigates the consequences of different \( M \) on the response of the classification techniques.
Prediction via Classification Techniques: the Patterns of the Failures

One way to reduce the uncertainty in the output, and therefore to save computational time, is to predict the failure via classification techniques. Figure 30 and Figure 31 show the scatter-plots derived from Example I for different sampling sizes (the number of realizations, \( M \), changes from one plot to the other); it can be seen how, as the number of realizations increases, the number of failures around the threshold line accumulates. The scatter plot progresses from two single random failures for \( M \) equal to 60, to a scattered straight line for \( M \) equal to 400, till finally forming a band for the two cases of \( M \) equal to 2000 and \( M \) equal to 10,000.

![Classification Agreement as function of Trials M](image)

**Figure 29:** Agreement in prediction as a function of the number of trials.

The results provided by the two different algorithms, CT and LR are satisfactory in the three cases analyzed here: Figure 29 shows that the LR maintains an overall 99%, or better, capability to predict the data, this means that even the case with lowest number of failures (for \( M=400 \) the total number of failures is just 13 and only 4 of them are used in the validation dataset) presents no misclassifications. On the other hand, the CT at \( M=400 \) predicts correctly just half of the failures in the dataset, so, for this case, 2 failures.

| HTC measures | Thresholds | Pr(|t|)     |
|--------------|------------|------------|
|              | 400        | 2000       | 10,000     |
| Pr(|t|)       | 1.00E-03   | 6.06E-08   | 1.17E-32   |

**Table 18:** CT thresholds for HTC Core and Summary p-tests for LC
The thresholds, granted by CT, thus provide precise threshold values for the only significant parameter ‘HTC core’ for the last two cases of Table 18 where a dramatic increase in its p-test value, is shown. This decrease in significance is due to the “poorness” of the first dataset. Note that it is the low numbers of failures, as opposed to the total number of realizations, to determine the inefficacy in predictions. Figure 30 shows an extreme case for this by reporting the number of failures obtained by the first 60 realizations. The number of failures is in this case 2, which means that the classifier would have a training set of, at most, 1 failure.

Note that 2 over 60 yields a probability equal to 2.33E-02 which is quite close to the asymptotic value of 3.0E-02 which has been reported in Figure 25 and obtained from 10,000 simulations. This outcome can be interpreted by saying that, despite the apparent low numerical difference between 2.33E-02 and 3.0E-02, the uncertainty related to the patterns of PF is actually significant as illustrated by comparing the last scatter-plot of Figure 31 with the one of Figure 30.

![Cp gas versus HTC Core for M=60](image)

*Figure 30:* Scatter-plot of the failures in the core for a number of samplings M = 60.
Figure 31: Scatter-plots of response and failures as a function of the Number of Realizations M. GFR’s core Model (Example I) for 400, 2000 and 10,000 realizations.
The same analysis has been carried out using the dataset of Example II. In this case the upper value for the number of possible samplings is now set to 128. Figure 32 shows the scatter-plots next to the tree structure. The interesting finding here is given by the ‘Max Leakage’ threshold line which, in order to capture as many failures as possible, translates to the right of the graph, as $M$ increases. This behavior confirms the capability of supervised methods to confine the dependent variable into clustered regions even for the case of relative low values of samplings (and consequently for low number of failures).

Figure 32: Scatter-plots of response and failures as a function of the Number of Realizations $M$. GFR’s core Model (Example II). Threshold lines are indicated in the graphs.
Findings

Through the third example we presented a purely mathematical model in order to explore the implications of an uncertain PF output, as emerged from Examples I and II. We decomposed this problem by stating that these uncertainties could depend on different factors, and the seven observations are the result of this discretization. What is inferred from the above seven considerations is that:

- The value of the output of a given system is affected by uncertainty, the same then is for PF;
- This uncertainty is nonetheless small if the number of samples is high enough;
- If the number of samples is sufficiently high, the issue is non relevant;
- If the number of simulations is limited by – as is the case for Example II - computational time constraints, there is a chance that this small uncertainty can amplify, as in the case of multiple system failures and multiple failure modes.

Once again the critical parameter, amongst the many determined, is the number of simulations available, M.

Since the most common scenario is that of a small number of available simulations, this exercise also explains the need to monitor the PF uncertainty through a set of simulations with a different number of seeds. By doing so, an estimate of the PF, even when its asymptotical value is unknown, could be obtained.

At the same time, the capability of the model to predict the system’s probability of failure was proven to be greatly influenced by the number of simulations M:

The scatter plots provide for a visual understanding of these considerations, by showing that the points, associated with the failure events, are uncertain in position, and that the patterns followed by these failures remain undetermined by changing their boundaries.

Further studies should explore the connection between these two last factors, the uncertainty of the PF, and the unspecified patterns followed by the failures.
4. FINAL REMARKS AND CONCLUSIONS

The purpose of this work was to define a correct methodology that would determine the reliability of Passive Systems. Specifically, the method proposed here concentrated on the study of those fundamental parameters that are most significant in leading the system to failure.

The issue of correctly assessing these fundamental quantities is not trivial. Concurrently to the appropriate selection of the parameters, this work studied to define a hierarchy among them. For this purpose, it is crucial to properly rank and prioritize the different components, or modules, of the system that fail as a consequence of those parameters. The method aims to discover the possible combinations of the latter that would result in a system failure.

The methodology described in this work was designed to be computationally swift and easy to understand and apply, all the while being mathematically robust. A statistical approach fulfilled these requirements, and the tools selected were that of the Classification Tree and Logistic Regression. An effort was made to validate each step by means of visual verification combined with more classification analysis, including the tools mentioned, p-tests and Wald Statistics. The method was applied to systems undergoing transient conditions so that failures of the system could be simultaneously and immediately detectable.

In order to gauge the efficacy of the method, 3 examples were developed, each one with a specific purpose in mind.

Example I was designed to be purely illustrative of the method: an unlikely failure of the core was conveniently simulated by means of a systems dynamics model.

Example II studied a more likely scenario, by taking into account the model of full plant in order to better study the propagation of failures from one system to the next.

With the purpose of containing the already intense computational time involved in these models, only one failure criterion was used (T cladding for the first Example and T structural collapse of the ECCS Pipe for the second).

Example III was a consequence of the results obtained by these analyses.
It appeared in fact, that while the first Example yielded a 99% agreement between the output predicted by the method and that of the model, the results of second Example were a bit more difficult to interpret, since the number of simulations available was low (128).

Examples I and II differed, aside from in the complexity of their models, also in the number of simulations available, $M$, for each. This difference in $M$ was non-trivial: 10,000 for the first, and 128 for the second. It became significant to understand whether there was a relation between $M$ and the uncertainty of the output $PF$.

Where Example I succeeded in validating the method proposed, Example II proved that the low number of simulations hindered the capacity to correctly predict $PF$.

Example III was developed to analyze this issue, and thus purely mathematical, with no apparent physical connotation. This test was critical since it also considered the possibility, not present in the first two examples, of multiple failures combined with multiple failure modes. The problem was approached by creating two distinct models, one linear and one non-linear, coherent with the topology of the previous examples. The outputs of these 2 models were then analyzed based on the number of simulations to which the two systems were subjected.

The outcome of this analysis was that the output did in fact vary depending on the number of simulations available. Furthermore, it became apparent that for a low $M$, other factors became increasingly significant in determining the uncertainty of the $PF$. These factors are the shape and distribution of the input parameters, their seed number, the sampling criterion used to select values for the input (in this case: LHS), the type of model used to simulate their behavior and a combination of one or more of the above.

This result was achieved, among other means, by that of the visual study of the scatter plots yielded by the analysis. These showed that those points, associated with failure events, were uncertain in position and presented no determinate pattern when $M$ was low.

The work conducted highlights the importance of the number of simulations when predicting that probability of failure of a system. Furthermore, since the most common scenario is that of a low number of trial runs, it calls for further studies into the connection between the uncertainty of the $PF$ and the unspecified patterns followed in the scatter plots.
5. RISK-INFORMED REGULATIONS

The initial objective of this work was to improve the current state of risk-informed regulations. The present study supported this original goal in two ways:

First, in a practical sense, by increasing the body of cases available in the literature of applications of load-capacity studies in the nuclear field. This means that this study succeeded in providing a direct contribution to the need of having a systematic approach to the proper analysis of uncertainties’ propagation.

Secondly, by combining practical and numerical methods in order to provide further insights when applying risk-informed regulations. What is meant by this, is that the statistical method used throughout this study, was aimed at improving the current approach and problem solving algorithms exercised in this type of problems.

The accuracy of the prediction methods was enhanced by the study of the relevance of the available simulations, M. The confidence statement concerning the applicability of a given number of simulations, usually set at 128, was debated through a methodology that was further supported by three fundamental examples. In this case, a systematic way to include uncertainties in the models, or codes, representing the plant’s “model of the world” was approached both, from the performance and from the acceptance criteria points of view. The analysis performed on the numerical implementation of those procedures connecting these models to the simulator, could lead to further significant improvements to the present-day risk-informed regulations.
6. FUTURE DEVELOPMENTS

In the development of future works some of the aspects not covered here could be addressed. The following list summarizes those aspects that were left aside and in the author's opinion would deserve further attention.

1. The specific role played in the prediction models by the results obtained through p-value testing, should be looked into more closely; the p-value technique was availed as a criterion to evaluate the statistical significance of the parameters, yet the use of the 0.05 value, provided by the chi square, has to be substantiated.

2. The effects generated by the presence of non-linearities in the model of the world.

3. Check the validity of other supervised learning techniques, especially when it comes to problems with multiple outputs. This issue was addressed earlier in this work while explaining LHS, and when discussing the approach used in Example III. In other words, while in the first two examples the input dataset was monitored for the case of just one output variable (the probability of failure of a single system), in Example III the problem went from being simply 2-Dimensional to becoming N-Dimensional. This increase in complexity implicates the loss of relevance of the scatter-plots, thus other methods, such as cluster analyses, could be used in order to verify the failures. Another approach to this problem could be to perform a principal component analysis in order to assess the variance of the parameters. These two possible solutions have the advantage being always applicable, even when no prior knowledge of the model is available.

4. Test the methodology with an enhanced model of the whole plant developed with a high-level software (GoldsSim® instead of Vensim®) capable of tolerating the use of deterministic and probabilistic variables for a high number of variables and/or for a high number of simulations, M. GoldsSim® has the distinct advantage of providing specific tools such as containment transport and reliability modules. These in particular, facilitate the development of complex engineering designs. The model is under development and it is named GFRDYN_LUMPED_5N. Its code was conceived to be capable of considering the interface between the Passive System, the containment, and the core. The motivation to present a final model, complete of all the features discussed throughout the 3 examples, is to present more credible results to risk-informed regulations by providing a more realistic and comprehensive scenario.
7. REFERENCES


8. APPENDICES

8.1 Appendix A: Sub-Model of the GFR’s Core used for Example I

The following algorithm was modeled by means of Vensim 5.4 by Ventana® Systems.

SYSTEM:

CORE, ECCS, HXC

~

~ Plant’s Systems

****************************************************

.Overall Heat Transfer Core

****************************************************~

Fuel Resistivity=

\[
((1-(\text{Internal Radius Channel RF1}/\text{Inner Cladding Radius Rci})^2)/4-\ln(\text{Inner Cladding Radius Rci}/\text{Internal Radius Channel RF1}))/2/(2\times3.14\times\text{Fuel Conductivity Kf})\sim\frac{\text{mK}}{\text{J/sec}}
\]

Active Length core=

1.54

~ m/channel

Gas Channel Resistivity=

\[
1/(\text{Area Active Core}\times\text{Htc Core})\sim\frac{\text{K}}{\text{J/s}}
\]

Average Radius Gap=

\[
4.2/1000+70/1e+006/2\sim\text{m}
\]

Cladding Conductivity Kc=

20

~ J/(m*s*K)

Cladding Resistivity=

\[
\ln(\text{Outer Cladding Radius Rco}/\text{Inner Cladding Radius Rci})/(2\times3.14\times\text{Cladding Conductivity Kc})\sim\frac{\text{mK}}{\text{J/sec}}
\]
Equivalent Diameter = 0.006983 m

External Equivalent Channel Diameter = 6.983/1000 m

Fuel Conductivity $K_f$ = 3.7 J/(m·s·K)

Gap Resistivity = $\frac{1}{2 \times 3.14 \times \text{Average Radius Gap} \times \text{Heat transfer Coeff Gap}}$ m·K/(J/sec)

Gap Thickness = 70/1e+006 m

Internal Radius Channel $R_{f1}$ = 4.2/1000 + 70/1e+006 m

Gas Conductivity $K_{gap}$ = 0.08 J/(m·s·K)

Heat transfer Coeff Gap = $\frac{\text{Gas Conductivity } K_{gap}}{\text{Gap Thickness}}$ J/(m·m·K)

Outer Cladding Radius $R_{co}$ = 4.2/1000 m

Inner Cladding Radius $R_{ci}$ = Equivalent Diameter/2 m
Number of Channels=
100170
~
channels
~

Lumped Core's Length=
Active Length core*Number of Channels
~
m
~

Overall Heat Transfer Coefficient=
1/(Cladding Resistivity/Lumped Core's Length+Fuel Resistivity/Lumped Core's Length+Gap Resistivity/Lumped Core's Length+Gas Channel Resistivity)
~
J/(s*K)
~

Htc Core=
100
~
J/s/(K*m*m)
~

Area Active Core=
1260
~
m*m
~

********************************************
*.Validity Test Indexes
********************************************

Heat Pulse=
Amplitude Pulse*pulse(1000,1000)
~
MJ/sec
~

switch Pulse=
0
~
Dmnl
~
1= no pulse ; 0= pulse in heat

Amplitude Pulse=
130
~
~
Failure Model

Failure Counter Cladding =
  IF THEN ELSE(T Cladding >= T Limit Cladding, 1, 0)
  ~   Dmnl
  ~

Time Coefficient Cladding =
  1
  ~   1/s
  ~

Time of residence in a failure state Cladding =
  IF THEN ELSE(Time = FINAL TIME, T Counter Cladding, 0)
  ~   Dmnl
  ~

T Counter Cladding = INTEG (Failure Counter Cladding * Time Coefficient Cladding, 0)
  ~   Dmnl
  ~

T Limit Cladding =
  1200 + 273.16
  ~   K
  ~

Failure Counter =
  IF THEN ELSE(Tmax Fuel >= T Fuel Limit, 1, 0)
  ~   Dmnl
  ~

T core Counter = INTEG (Failure Counter * Time Coefficient, 0)
  ~   Dmnl
  ~

Peaking Factor =
  1.3 * 1.15
  ~   Dmnl
  ~
Tmax Fuel =
  T Core*Peaking Factor
  ~ K
  ~ |

T Fuel Limit =
  2200+273.16
  ~ K
  ~ |

Time of residence in a failure state =
  IF THEN ELSE(Time=FINAL TIME, T core Counter, 0 )
  ~ Dmnl
  ~ |

"J/MJ Converter" =
  1e+006
  ~ J/MJ
  ~ |

*******************************************************************************
  .Initial Values
*******************************************************************************

Gas Mass flow Rate 0 =
  300
  ~ kg/s
  ~ |

Q Gas Core 0 =
  Mass Gas Core*Cp Gas*T gas Core 0 /"J/MJ Converter" 
  ~ MJ
  ~ |

Q Core 0 =
  Mass Core*Cp Core*T core 0 /"J/MJ Converter"
  ~ MJ
  ~ |

T core 0 =
  600
  ~ K
  ~ |

T gas Core 0 =
  +400
  ~ K
  ~ |

Q Extracted from Core 0 =
0
~ ~ MJ
INITIAL TIME = 550
~ Second
~ The initial time for the simulation.

******************************************************************************
.Model Variables
******************************************************************************

Q gas transfer rate=
  IF THEN ELSE(switch Pulse=1, Overall Heat Transfer Coefficient*(T Core-T gas Core)/"J/MJ Converter", Overall Heat Transfer Coefficient*(T Core-T gas Core)/"J/MJ Converter"-Heat Pulse)
  ~ MJ/sec
  ~

Tin Core=
  400
  ~ K
  ~

Tout Core=
  2*T gas Core- Tin Core
  ~ K
  ~

T Cladding=
  T Core+Q gas transfer rate/(1/(Cladding Resistivity/Lumped Core's Length+Fuel Resistivity/Lumped Core's Length+Gap Resistivity/Lumped Core's Length))"J/MJ Converter"
  ~ K
  ~

Q decay rate=
  a*Power Th*(Time+1e-006)^(-b)
  ~ MJ/sec
  ~

Q core extraction rate=
  Gas Mass flow Rate 0*Cp Gas*"deltaTcore=Tout-Tin"/"J/MJ Converter"
  ~ MJ/sec
  ~

"deltaTcore=Tout-Tin"=
  Tout Core-Tin Core
  ~ K
  ~

T Core=
Q Core"/MJ Converter"/(Cp Core*Mass Core)
~ K
~

T gas Core=
Q Gas Core"/MJ Converter"/(Cp Gas*Mass Gas Core)
~ K
~

Parameters

Volume Tot=
70000
~ m*m*m
~
Cp Core=
536
~ J/(K*kg)
~

a=
0.173
~ Dmnl
~

b=
0.28
~ Dmnl
~

Cp Gas=
1180.5
~ J/(K*kg)
~

Mass Core=
115942
~ kg
~

Mass Gas Core=
36000
~ kg
~

Power Th=
2400
MJ/sec

.Level Variables

| Q Extracted from Core = INTEG (  
| Q core extraction rate,  
| Q Extracted from Core 0)  
| ~ MJ  
| ~ |

Q Core = INTEG (  
| Q decay rate-Q gas transfer rate,  
| Q Core 0)  
| ~ MJ  
| ~ |

Q Gas Core = INTEG (  
| +Q gas transfer rate-Q core extraction rate,  
| Q Gas Core 0)  
| ~ MJ  
| ~ |

.Unit Converters

| Time Coefficient =  
| 1  
| ~ 1/s  
| ~ |

.Control: Simulation Control Parameters

| FINAL TIME = 10000  
| ~ Second  
| ~ The final time for the simulation.  

| SAVEPER = TIME STEP  
| ~ Second [0,?]  
| ~ The frequency with which output is stored.  

| TIME STEP = 1  
| ~ Second [0,?]  
| ~ The time step for the simulation.  


8.2 Appendix B: Dataset from RELAP5-3D used for Example II

The following dataset was created by Francisco Mackay and can be found in [F. MacKay, 2007].

<table>
<thead>
<tr>
<th>Roughness In the core</th>
<th>Valve 674 Leakage</th>
<th>Valve 665 Leakage</th>
<th>Heat Transfer Coefficient in Containment</th>
<th>Heat Transfer Coefficient in Core</th>
<th>Shaft Moment of Inertia</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.0266E-05</td>
<td>0.860972906</td>
<td>0.132791381</td>
<td>1.018099269</td>
<td>0.175240146</td>
</tr>
<tr>
<td>2</td>
<td>3.17861E-05</td>
<td>0.464606103</td>
<td>0.009216567</td>
<td>0.619031756</td>
<td>1.49147449</td>
</tr>
<tr>
<td>3</td>
<td>1.41968E-05</td>
<td>0.189936412</td>
<td>0.446353642</td>
<td>0.518304501</td>
<td>4.089563936</td>
</tr>
<tr>
<td>4</td>
<td>0.000165969</td>
<td>0.2247406</td>
<td>1.665210385</td>
<td>3.074084819</td>
<td>1.339039701</td>
</tr>
<tr>
<td>5</td>
<td>2.23325E-05</td>
<td>0.110838306</td>
<td>0.363779265</td>
<td>3.883312254</td>
<td>0.543303791</td>
</tr>
<tr>
<td>6</td>
<td>8.95976E-05</td>
<td>0.082665107</td>
<td>0.273528511</td>
<td>0.94501105</td>
<td>0.936337946</td>
</tr>
<tr>
<td>7</td>
<td>2.53847E-05</td>
<td>0.392584938</td>
<td>0.085119427</td>
<td>1.183833099</td>
<td>0.646817192</td>
</tr>
<tr>
<td>8</td>
<td>4.4716E-05</td>
<td>0.032207799</td>
<td>0.215695098</td>
<td>1.643178094</td>
<td>0.376749072</td>
</tr>
<tr>
<td>9</td>
<td>5.38544E-05</td>
<td>0.607582295</td>
<td>0.156639456</td>
<td>1.772914465</td>
<td>1.793946565</td>
</tr>
<tr>
<td>10</td>
<td>7.85191E-05</td>
<td>0.302996778</td>
<td>0.042048643</td>
<td>0.277030229</td>
<td>2.391928364</td>
</tr>
<tr>
<td>11</td>
<td>3.85214E-05</td>
<td>0.17446045</td>
<td>0.09064391</td>
<td>4.829592916</td>
<td>0.293658031</td>
</tr>
<tr>
<td>12</td>
<td>8.47645E-06</td>
<td>0.140217132</td>
<td>0.220896755</td>
<td>0.422911142</td>
<td>0.864267335</td>
</tr>
<tr>
<td>13</td>
<td>5.91768E-05</td>
<td>0.058180793</td>
<td>0.70785134</td>
<td>0.838450117</td>
<td>0.465594267</td>
</tr>
<tr>
<td>14</td>
<td>2.84935E-05</td>
<td>0.012386872</td>
<td>0.534292341</td>
<td>2.266994934</td>
<td>1.104884296</td>
</tr>
<tr>
<td>15</td>
<td>1.82596E-05</td>
<td>0.337699928</td>
<td>0.348002238</td>
<td>0.745117621</td>
<td>0.761977965</td>
</tr>
<tr>
<td>16</td>
<td>0.000123131</td>
<td>0.890796357</td>
<td>0.023758162</td>
<td>1.333115501</td>
<td>0.429398495</td>
</tr>
<tr>
<td>17</td>
<td>7.08761E-06</td>
<td>0.066021338</td>
<td>0.127291225</td>
<td>11.24755112</td>
<td>1.596157589</td>
</tr>
<tr>
<td>18</td>
<td>1.10777E-05</td>
<td>0.099241677</td>
<td>0.061949683</td>
<td>1.369543295</td>
<td>1.013147351</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>19</td>
<td>0.000105528</td>
<td>0.529594859</td>
<td>0.189806486</td>
<td>0.948642402</td>
<td>0.727023454</td>
</tr>
<tr>
<td>20</td>
<td>4.31121E-05</td>
<td>0.74219072</td>
<td>0.114714985</td>
<td>3.263201796</td>
<td>1.224111492</td>
</tr>
<tr>
<td>21</td>
<td>0.000211663</td>
<td>0.03001568</td>
<td>0.069037775</td>
<td>0.210871326</td>
<td>0.676329216</td>
</tr>
<tr>
<td>22</td>
<td>1.92448E-05</td>
<td>0.202227439</td>
<td>0.985357812</td>
<td>0.905538264</td>
<td>1.170526951</td>
</tr>
<tr>
<td>23</td>
<td>3.40267E-05</td>
<td>0.263548363</td>
<td>0.178068757</td>
<td>0.803267482</td>
<td>3.684155094</td>
</tr>
<tr>
<td>24</td>
<td>4.96232E-05</td>
<td>0.498487678</td>
<td>0.238331675</td>
<td>0.468322129</td>
<td>0.095135605</td>
</tr>
<tr>
<td>25</td>
<td>0.00014868</td>
<td>0.043266545</td>
<td>0.286472444</td>
<td>2.184673244</td>
<td>0.802686297</td>
</tr>
<tr>
<td>26</td>
<td>1.71049E-05</td>
<td>1.190170101</td>
<td>0.512037361</td>
<td>1.487914895</td>
<td>0.242057146</td>
</tr>
<tr>
<td>27</td>
<td>5.68576E-05</td>
<td>0.001891644</td>
<td>0.593915737</td>
<td>1.121569675</td>
<td>0.409672324</td>
</tr>
<tr>
<td>28</td>
<td>2.34316E-05</td>
<td>0.422684746</td>
<td>0.321484425</td>
<td>0.347961093</td>
<td>2.105991987</td>
</tr>
<tr>
<td>29</td>
<td>3.63424E-05</td>
<td>0.127905542</td>
<td>0.852944013</td>
<td>0.689984288</td>
<td>0.511600116</td>
</tr>
<tr>
<td>30</td>
<td>6.49453E-05</td>
<td>0.256755913</td>
<td>0.422403824</td>
<td>2.608512339</td>
<td>0.328936234</td>
</tr>
<tr>
<td>31</td>
<td>8.00589E-05</td>
<td>0.158081862</td>
<td>0.017396258</td>
<td>1.921065859</td>
<td>0.574820163</td>
</tr>
<tr>
<td>32</td>
<td>2.91641E-05</td>
<td>0.330425174</td>
<td>0.039769783</td>
<td>0.542943227</td>
<td>1.919305918</td>
</tr>
<tr>
<td>33</td>
<td>4.68492E-05</td>
<td>0.641145484</td>
<td>0.678139931</td>
<td>0.770016585</td>
<td>0.25968491</td>
</tr>
<tr>
<td>34</td>
<td>9.93625E-05</td>
<td>0.80248295</td>
<td>0.911992663</td>
<td>1.061765697</td>
<td>2.910536583</td>
</tr>
<tr>
<td>35</td>
<td>8.51434E-05</td>
<td>0.154193833</td>
<td>0.30223248</td>
<td>1.397918526</td>
<td>0.889602906</td>
</tr>
<tr>
<td>36</td>
<td>5.12343E-05</td>
<td>0.084045934</td>
<td>0.333844337</td>
<td>0.303841247</td>
<td>0.792966876</td>
</tr>
<tr>
<td>37</td>
<td>7.10865E-05</td>
<td>0.480768573</td>
<td>1.230679167</td>
<td>1.90661706</td>
<td>1.247906264</td>
</tr>
<tr>
<td>38</td>
<td>0.00011937</td>
<td>0.23343891</td>
<td>0.163818188</td>
<td>0.811362806</td>
<td>0.152467084</td>
</tr>
<tr>
<td>39</td>
<td>2.0305E-05</td>
<td>0.136108256</td>
<td>0.490839118</td>
<td>1.721271504</td>
<td>0.397270165</td>
</tr>
<tr>
<td>40</td>
<td>9.97261E-06</td>
<td>0.994760314</td>
<td>0.561209485</td>
<td>2.350701033</td>
<td>0.999122583</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>41</td>
<td>2.17877E-05</td>
<td>0.187909689</td>
<td>0.24745941</td>
<td>3.571845903</td>
<td>0.4874615</td>
</tr>
<tr>
<td>42</td>
<td>1.48581E-05</td>
<td>0.244850439</td>
<td>0.183940052</td>
<td>1.239715178</td>
<td>1.684962153</td>
</tr>
<tr>
<td>43</td>
<td>4.71902E-05</td>
<td>0.282812486</td>
<td>0.202202188</td>
<td>0.358308616</td>
<td>2.057624559</td>
</tr>
<tr>
<td>44</td>
<td>3.09792E-05</td>
<td>0.403472191</td>
<td>0.793651269</td>
<td>6.032402101</td>
<td>0.698321791</td>
</tr>
<tr>
<td>45</td>
<td>7.38153E-05</td>
<td>0.060647096</td>
<td>0.629844302</td>
<td>0.71082589</td>
<td>0.956883392</td>
</tr>
<tr>
<td>46</td>
<td>0.00020055</td>
<td>0.244850439</td>
<td>0.183940052</td>
<td>1.239715178</td>
<td>1.684962153</td>
</tr>
<tr>
<td>47</td>
<td>2.77006E-05</td>
<td>0.104256214</td>
<td>0.153300021</td>
<td>5.406115063</td>
<td>8.141160654</td>
</tr>
<tr>
<td>48</td>
<td>3.98161E-05</td>
<td>0.01308486</td>
<td>0.884653856</td>
<td>1.560774312</td>
<td>568.5798634</td>
</tr>
<tr>
<td>49</td>
<td>4.62055E-05</td>
<td>0.315619703</td>
<td>0.383399114</td>
<td>1.591284501</td>
<td>0.631478997</td>
</tr>
<tr>
<td>50</td>
<td>1.61458E-05</td>
<td>0.00693723</td>
<td>0.46147984</td>
<td>2.286157212</td>
<td>620.3083227</td>
</tr>
<tr>
<td>51</td>
<td>6.56956E-05</td>
<td>0.17005815</td>
<td>0.231037463</td>
<td>1.54223527</td>
<td>1.041974271</td>
</tr>
<tr>
<td>52</td>
<td>6.14829E-05</td>
<td>0.018504666</td>
<td>0.012842897</td>
<td>4.001604887</td>
<td>0.361890083</td>
</tr>
<tr>
<td>53</td>
<td>2.61785E-05</td>
<td>0.022433357</td>
<td>0.316280157</td>
<td>0.860842505</td>
<td>0.344565861</td>
</tr>
<tr>
<td>54</td>
<td>0.000267854</td>
<td>1.442496364</td>
<td>0.054688522</td>
<td>0.394960618</td>
<td>0.276108579</td>
</tr>
<tr>
<td>55</td>
<td>1.24528E-05</td>
<td>0.361764165</td>
<td>0.118231636</td>
<td>0.250464186</td>
<td>0.831776376</td>
</tr>
<tr>
<td>56</td>
<td>0.000133637</td>
<td>0.117707471</td>
<td>0.100023068</td>
<td>0.575748717</td>
<td>0.455933767</td>
</tr>
<tr>
<td>57</td>
<td>2.46989E-05</td>
<td>0.094621349</td>
<td>0.03586993</td>
<td>2.021533534</td>
<td>1.396404502</td>
</tr>
<tr>
<td>58</td>
<td>5.5652E-05</td>
<td>0.049276257</td>
<td>0.258345944</td>
<td>0.605107028</td>
<td>0.212640618</td>
</tr>
<tr>
<td>59</td>
<td>1.74366E-05</td>
<td>0.449463712</td>
<td>0.142381726</td>
<td>1.115250099</td>
<td>0.524400738</td>
</tr>
<tr>
<td>60</td>
<td>3.79521E-05</td>
<td>0.038705296</td>
<td>0.075183345</td>
<td>0.181204116</td>
<td>0.560520061</td>
</tr>
<tr>
<td>61</td>
<td>9.74176E-05</td>
<td>0.272326653</td>
<td>0.106640922</td>
<td>0.484303178</td>
<td>2.635348042</td>
</tr>
<tr>
<td>62</td>
<td>4.16683E-05</td>
<td>0.210614439</td>
<td>0.051236791</td>
<td>0.997587303</td>
<td>0.70547236</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>63</td>
<td>3.56232E-05</td>
<td>0.574585443</td>
<td>0.406903451</td>
<td>2.757981398</td>
<td>0.608032193</td>
</tr>
<tr>
<td>64</td>
<td>3.30958E-05</td>
<td>0.374395138</td>
<td>0.028567636</td>
<td>0.443542793</td>
<td>1.406811368</td>
</tr>
<tr>
<td>65</td>
<td>2.44393E-05</td>
<td>0.03892315</td>
<td>0.037643374</td>
<td>1.826227042</td>
<td>0.743613759</td>
</tr>
<tr>
<td>66</td>
<td>3.26633E-05</td>
<td>0.429736516</td>
<td>0.073220858</td>
<td>2.374057349</td>
<td>0.66304439</td>
</tr>
<tr>
<td>67</td>
<td>0.000119456</td>
<td>0.853064915</td>
<td>0.110295296</td>
<td>2.700746814</td>
<td>0.35031969</td>
</tr>
<tr>
<td>68</td>
<td>4.2467E-05</td>
<td>0.163285009</td>
<td>0.057242128</td>
<td>0.163381672</td>
<td>0.44344617</td>
</tr>
<tr>
<td>69</td>
<td>6.11709E-05</td>
<td>0.182661765</td>
<td>0.194803481</td>
<td>0.323280366</td>
<td>0.289654677</td>
</tr>
<tr>
<td>70</td>
<td>1.44549E-05</td>
<td>0.177547568</td>
<td>0.031614643</td>
<td>1.255017446</td>
<td>0.4822816</td>
</tr>
<tr>
<td>71</td>
<td>3.96156E-05</td>
<td>0.294931273</td>
<td>0.663103609</td>
<td>0.515879692</td>
<td>0.118911905</td>
</tr>
<tr>
<td>72</td>
<td>1.63018E-05</td>
<td>0.292796158</td>
<td>0.396770153</td>
<td>4.520676725</td>
<td>0.75463197</td>
</tr>
<tr>
<td>73</td>
<td>9.44872E-05</td>
<td>0.05127336</td>
<td>0.294894594</td>
<td>1.217860942</td>
<td>3.045780555</td>
</tr>
<tr>
<td>74</td>
<td>1.35938E-05</td>
<td>0.562509747</td>
<td>0.09561235</td>
<td>1.060178579</td>
<td>2.559356379</td>
</tr>
<tr>
<td>75</td>
<td>1.56751E-05</td>
<td>0.069275548</td>
<td>0.327920832</td>
<td>2.203124185</td>
<td>0.33698293</td>
</tr>
<tr>
<td>76</td>
<td>9.2457E-05</td>
<td>0.220260843</td>
<td>0.264218317</td>
<td>0.38688355</td>
<td>0.92479469</td>
</tr>
<tr>
<td>77</td>
<td>3.50952E-05</td>
<td>0.321362252</td>
<td>1.155321411</td>
<td>0.787872797</td>
<td>1.950232297</td>
</tr>
<tr>
<td>78</td>
<td>5.46275E-05</td>
<td>0.089781232</td>
<td>0.160602313</td>
<td>1.744233057</td>
<td>0.903230525</td>
</tr>
<tr>
<td>79</td>
<td>4.80789E-05</td>
<td>0.21184961</td>
<td>0.357060873</td>
<td>1.430553609</td>
<td>0.62257917</td>
</tr>
<tr>
<td>80</td>
<td>1.19785E-05</td>
<td>0.348104379</td>
<td>0.060423816</td>
<td>1.51025574</td>
<td>1.281580227</td>
</tr>
<tr>
<td>81</td>
<td>2.29841E-05</td>
<td>0.078958706</td>
<td>0.066758276</td>
<td>0.921965926</td>
<td>0.549252891</td>
</tr>
<tr>
<td>82</td>
<td>5.76447E-05</td>
<td>0.109910569</td>
<td>0.207762362</td>
<td>0.234382257</td>
<td>0.418804668</td>
</tr>
<tr>
<td>83</td>
<td>4.8909E-05</td>
<td>0.312370023</td>
<td>0.477193119</td>
<td>2.429644306</td>
<td>0.568941377</td>
</tr>
<tr>
<td>84</td>
<td>4.05737E-05</td>
<td>0.387344337</td>
<td>0.828494585</td>
<td>1.564471368</td>
<td>1.629139496</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>85</td>
<td>9.09296E-06</td>
<td>0.015774464</td>
<td>0.138237619</td>
<td>1.312702309</td>
<td>1.306318855</td>
</tr>
<tr>
<td>86</td>
<td>7.22034E-05</td>
<td>0.491041341</td>
<td>0.006652834</td>
<td>1.360475912</td>
<td>0.870487068</td>
</tr>
<tr>
<td>87</td>
<td>7.62102E-05</td>
<td>0.405617924</td>
<td>0.018203018</td>
<td>1.484275245</td>
<td>0.217787688</td>
</tr>
<tr>
<td>88</td>
<td>8.1427E-06</td>
<td>0.198004657</td>
<td>0.341567074</td>
<td>0.218745727</td>
<td>1.350890534</td>
</tr>
<tr>
<td>89</td>
<td>6.33097E-05</td>
<td>1.066584672</td>
<td>0.393328526</td>
<td>3.731597047</td>
<td>1.756550384</td>
</tr>
<tr>
<td>90</td>
<td>1.07823E-05</td>
<td>0.063723513</td>
<td>1.027471598</td>
<td>1.087912879</td>
<td>0.600179238</td>
</tr>
<tr>
<td>91</td>
<td>3.42523E-05</td>
<td>0.026950605</td>
<td>0.281080473</td>
<td>2.933790801</td>
<td>1.542390549</td>
</tr>
<tr>
<td>92</td>
<td>3.00186E-05</td>
<td>0.238012053</td>
<td>0.143020405</td>
<td>0.639826242</td>
<td>1.131315686</td>
</tr>
<tr>
<td>93</td>
<td>0.000129671</td>
<td>0.124871384</td>
<td>1.416088606</td>
<td>0.328154632</td>
<td>0.252036602</td>
</tr>
<tr>
<td>94</td>
<td>0.000154807</td>
<td>1.640795395</td>
<td>0.174491436</td>
<td>0.698446807</td>
<td>1.458026427</td>
</tr>
<tr>
<td>95</td>
<td>5.83034E-05</td>
<td>0.114459764</td>
<td>0.548945478</td>
<td>4.244041273</td>
<td>0.30499423</td>
</tr>
<tr>
<td>96</td>
<td>2.08798E-05</td>
<td>0.054269715</td>
<td>0.049615326</td>
<td>0.969358208</td>
<td>0.158179792</td>
</tr>
<tr>
<td>97</td>
<td>4.6118E-05</td>
<td>0.046090401</td>
<td>0.72803977</td>
<td>0.431048625</td>
<td>0.228051476</td>
</tr>
<tr>
<td>98</td>
<td>3.19744E-05</td>
<td>0.123509661</td>
<td>0.378526192</td>
<td>5.068997272</td>
<td>1.194993691</td>
</tr>
<tr>
<td>99</td>
<td>0.000109476</td>
<td>0.091293959</td>
<td>0.202069973</td>
<td>0.28116126</td>
<td>2.718226393</td>
</tr>
<tr>
<td>100</td>
<td>0.000101622</td>
<td>0.162009876</td>
<td>0.021353242</td>
<td>1.178274754</td>
<td>2.191621088</td>
</tr>
<tr>
<td>101</td>
<td>0.000117398</td>
<td>0.197024004</td>
<td>0.089317371</td>
<td>0.632726739</td>
<td>1.862677253</td>
</tr>
<tr>
<td>102</td>
<td>1.81666E-05</td>
<td>1.116133243</td>
<td>0.631878624</td>
<td>0.50362288</td>
<td>0.476034599</td>
</tr>
<tr>
<td>103</td>
<td>0.000140123</td>
<td>0.3694106</td>
<td>0.123210426</td>
<td>2.116812651</td>
<td>3.350699786</td>
</tr>
<tr>
<td>104</td>
<td>2.35482E-05</td>
<td>0.231032974</td>
<td>0.045327924</td>
<td>3.093001052</td>
<td>0.437834921</td>
</tr>
<tr>
<td>105</td>
<td>0.000184661</td>
<td>0.003414411</td>
<td>0.011420538</td>
<td>1.036389138</td>
<td>0.951400214</td>
</tr>
<tr>
<td>106</td>
<td>0.000350829</td>
<td>0.007558315</td>
<td>0.752383237</td>
<td>2.849447745</td>
<td>1.031748519</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>107</td>
<td>3.66486E-05</td>
<td>0.034792339</td>
<td>0.148018768</td>
<td>1.838354829</td>
<td>1.068334272</td>
</tr>
<tr>
<td>108</td>
<td>5.27417E-05</td>
<td>0.267794393</td>
<td>0.166937809</td>
<td>0.722632293</td>
<td>2.239221847</td>
</tr>
<tr>
<td>109</td>
<td>2.72498E-05</td>
<td>0.438173699</td>
<td>0.310950321</td>
<td>1.687224992</td>
<td>0.528280225</td>
</tr>
<tr>
<td>110</td>
<td>8.64598E-05</td>
<td>0.689469885</td>
<td>0.07897467</td>
<td>0.414067652</td>
<td>0.846720021</td>
</tr>
<tr>
<td>111</td>
<td>5.76494E-06</td>
<td>0.720264048</td>
<td>0.567437747</td>
<td>0.954076081</td>
<td>0.388809242</td>
</tr>
<tr>
<td>112</td>
<td>2.59696E-05</td>
<td>0.45873879</td>
<td>0.500678834</td>
<td>2.066317956</td>
<td>0.777646147</td>
</tr>
<tr>
<td>113</td>
<td>2.69684E-05</td>
<td>0.076801692</td>
<td>0.288628703</td>
<td>3.448935736</td>
<td>0.49654075</td>
</tr>
<tr>
<td>114</td>
<td>1.87046E-05</td>
<td>0.14315049</td>
<td>0.432706689</td>
<td>7.888527292</td>
<td>0.653884589</td>
</tr>
<tr>
<td>115</td>
<td>2.68496E-05</td>
<td>0.589125745</td>
<td>0.003257401</td>
<td>0.567182096</td>
<td>0.689344201</td>
</tr>
<tr>
<td>116</td>
<td>5.1259E-05</td>
<td>0.144203165</td>
<td>0.610120945</td>
<td>0.756656021</td>
<td>0.972753315</td>
</tr>
<tr>
<td>117</td>
<td>8.27308E-05</td>
<td>0.622005761</td>
<td>0.252690176</td>
<td>0.893303892</td>
<td>4.6458555</td>
</tr>
<tr>
<td>118</td>
<td>3.06792E-05</td>
<td>0.544848006</td>
<td>0.105402894</td>
<td>0.371657756</td>
<td>0.823148669</td>
</tr>
<tr>
<td>119</td>
<td>0.000242648</td>
<td>0.100470863</td>
<td>0.232242065</td>
<td>0.463772479</td>
<td>0.372780363</td>
</tr>
<tr>
<td>120</td>
<td>2.12095E-05</td>
<td>0.77753316</td>
<td>0.245206874</td>
<td>0.55595609</td>
<td>0.312550768</td>
</tr>
<tr>
<td>121</td>
<td>3.77208E-05</td>
<td>0.147013488</td>
<td>0.890147715</td>
<td>0.135457797</td>
<td>1.435246354</td>
</tr>
<tr>
<td>122</td>
<td>6.81034E-05</td>
<td>0.248773951</td>
<td>0.027479927</td>
<td>0.86335716</td>
<td>1.078165163</td>
</tr>
<tr>
<td>123</td>
<td>4.45943E-05</td>
<td>0.511529166</td>
<td>0.46634545</td>
<td>1.983037627</td>
<td>5.397763593</td>
</tr>
<tr>
<td>124</td>
<td>1.27816E-05</td>
<td>0.951530839</td>
<td>0.098224977</td>
<td>0.831457732</td>
<td>1.197266703</td>
</tr>
<tr>
<td>125</td>
<td>7.78064E-05</td>
<td>0.132872215</td>
<td>0.22159742</td>
<td>0.675999703</td>
<td>0.596440491</td>
</tr>
<tr>
<td>126</td>
<td>6.6777E-05</td>
<td>0.022776215</td>
<td>0.188431304</td>
<td>7.610320216</td>
<td>0.717407639</td>
</tr>
<tr>
<td>127</td>
<td>1.97961E-05</td>
<td>0.34663953</td>
<td>0.433029247</td>
<td>1.158006383</td>
<td>1.743479501</td>
</tr>
<tr>
<td>128</td>
<td>4.36691E-05</td>
<td>0.278621361</td>
<td>0.131480541</td>
<td>0.594252985</td>
<td>0.197257333</td>
</tr>
</tbody>
</table>
6.3 Appendix C: XML script from Insightful Miner used for Example II and III

The following XLM script was written with Insightful Miner by Insightful Corporation®.

```xml
<?xml version="1.0" encoding="UTF-8" ?>
- <IIML version="7.0">
- <Worksheet lastNodeID="26" linkType="diagonal">
- <FileInfo>
  <FileAuthor>Edo</FileAuthor>
  <FileName>C:\Documents and Settings\Edo\Desktop\3param10000\threeParCoreModel.imw</FileName>
  <FileDateCreated>1179601921171</FileDateCreated>
</FileInfo>
- <WorksheetProperties>
  - <XTProps>
    <Property name="maxStringSize" value="33" />
    <Property name="maxCategoricalLevels" value="500" />
    <Property name="numDigits" value="2" />
    <Property name="stringLeftAlign" value="false" />
    <Property name="dateParseString" value="%m[/]/%d[/]/%y [%H:[/]%M[:%S[.%N]]] [%p]" />
    <Property name="dateFormatString" value="%02m/%02d/%Y %02H:%02M:%02S" />
    <Property name="dateCenturyCutoff" value="1950" />
    <Property name="decimalSymbols" value=".," />
    <Property name="maxBufferMB" value="10" />
    <Property name="rowsToChunk" value="10000" />
    <Property name="useCache" value="yes" />
  </XTProps>
</WorksheetProperties>
- <WorksheetParameters>
- <XTProps />
</WorksheetParameters>
- <ComponentComments>
  - <AuthorComments>
    <AuthorName>Edo</AuthorName>
    <AuthorCommentTitle>Not Provided</AuthorCommentTitle>
    <AuthorCommentCreateDate>May 19, 2007</AuthorCommentCreateDate>
    <AuthorCommentText>Not Provided</AuthorCommentText>
  </AuthorComments>
  - <DiscussionComments>
    - <Comment>
      <CommentAuthor>Not Provided</CommentAuthor>
      <CommentCreateDate>Not Provided</CommentCreateDate>
      <CommentTitle>Not Provided</CommentTitle>
      <CommentText>Not Provided</CommentText>
    </Comment>
  </DiscussionComments>
</ComponentComments>
- <ActivityNodeList>
```

113 | Page
<ActivityNode engineClass="com.insightful.miner.ReadExcelFileEngineNode" guiClass="com.insightful.miner.ReadExcelFileNodeModel" numInputs="0" numOutputs="1" id="0">
  <DisplayInfo labelText="Read Dataset" defaultLabelText="Read Excel File"
smallIcon="read_excel_small.gif" largeIcon="read_excel_large.gif" x="43" y="175" />
- <ArgumentList>
- <XTProps>
  <Property name="useGlobal" value="true" />
  <Property name="useCache" value="global" />
  <Property name="filePath" value="C:\Documents and Settings\Edo\Desktop\3param10000\Calc_Class.xls" />
  <Property name="fileType" value="EXCEL" />
  <Property name="rowsToPreview" value="20" />
  <Property name="startRow" value="" />
  <Property name="endRow" value="" />
  <Property name="defaultStringColumnType" value="string" />
  <Property name="specifiedRowsToChunk" value="0" />
  <Property name="rowsToChunk" value="10000" />
  <Property name="randomSeed" value="newSeedEveryTime" />
  <Property name="randomSeedValue" value="5" />
  <Property name="isSampleFraction" value="false" />
  <Property name="isSampleEachN" value="false" />
  <Property name="sampleFraction" value="0.5" />
  <Property name="sampleEachN" value="2" />
  <Property name="worksheet" value="dataset60" />
- <Property name="columns">
- <Property name="Number" isSampled="false" />
- <Property name="Failure Fuel" isSampled="false" />
- <Property name="Failure Clad" isSampled="false" />
- <Property name="Time Out Fuel" isSampled="false" />
- <Property name="Time Out Clad" isSampled="false" />
- <Property name="Col9" isSampled="false" />
- <Property name="T CladEp Lim" isSampled="false" />
- <Property name="T Clad" isSampled="false" />
</Property>
</ArgumentList>
</ActivityNode>
- <Property name="SM">
  <Property name="isSelected" value="false" />
</Property>
- <Property name="Failure Cladding">
  <Property name="datatype" value="categorical" />
</Property>
- <Property name="test">
  <Property name="isSelected" value="false" />
</Property>
</Property>
</XTProps>
</ArgumentList>
- <ComponentComments>
  - <AuthorComments>
    <AuthorName>Edo</AuthorName>
    <AuthorCommentTitle>Not Provided</AuthorCommentTitle>
    <AuthorCommentDateCreated>May 19, 2007</AuthorCommentDateCreated>
    <AuthorCommentText>Not Provided</AuthorCommentText>
  </AuthorComments>
  - <DiscussionComments>
    <Comment>
      <CommentAuthor>Not Provided</CommentAuthor>
      <CommentDateCreated>Not Provided</CommentDateCreated>
      <CommentTitle>Not Provided</CommentTitle>
      <CommentText>Not Provided</CommentText>
    </Comment>
  </DiscussionComments>
</ComponentComments>
</ActivityNode>
- <ActivityNode engineClass="com.insightful.miner.ClassificationRpartEngineNode"
  guiClass="com.insightful.miner.ClassificationTreeNodeModel" numInputs="1" numOutputs="1"
  id="1">
  <DisplayInfo labelText="Classification Tree (1)" defaultLabelText="Classification Tree"
    smallIcon="classification_tree_small.gif" largeIcon="classification_tree_large.gif" x="260"
    y="365" />
  - <ArgumentList>
    - <XTProps>
      - <Property name="copyInputColumns">
        <Property name="independent" value="false" />
        <Property name="dependent" value="true" />
        <Property name="other" value="false" />
      </Property>
      - <Property name="newColumns">
        <Property name="probability" value="true" />
        <Property name="classification" value="true" />
        <Property name="agreement" value="false" />
        <Property name="probabilityFor" value="forSpecifiedCategory" />
        <Property name="specifiedCategory" value="1" />
      </Property>
      - <Property name="tree.ensemble" value="true" />
      <Property name="rows.per.tree" value="300" />
    </ XTProps>
</ActivityNode>
- <ActivityNode engineClass="com.insightful.miner.PartitionEngineNode"
guiClass="com.insightful.miner.PartitionNodeModel" numInputs="1" numOutputs="3" id="6">
  <DisplayInfo labelText="Partition (6)" defaultLabelText="Partition"
smallIcon="partition_small.gif" largeIcon="partition_large.gif" x="160" y="175" />
  <OutputPortToolTips port1="Train" port2="Test" port3="Validate" />
  - <ArgumentList>
    - <XTProps>
      <Property name="trainSize" value="0.5" />
      <Property name="testSize" value="0.5" />
      <Property name="validSize" value="0.0" />
      <Property name="useGlobal" value="true" />
      <Property name="useCache" value="global" />
      <Property name="randomSeed" value="newSeedEveryTime" />
      <Property name="randomSeedValue" value="5" />
      <Property name="specifiedRowsToChunk" value="0" />
      <Property name="rowsToChunk" value="10000" />
    </XTProps>
    </ArgumentList>
    - <ComponentComments>
      - <AuthorComments>
        <AuthorName>Edo</AuthorName>
        <AuthorCommentTitle>Not Provided</AuthorCommentTitle>
        <AuthorCommentDateCreated>May 19, 2007</AuthorCommentDateCreated>
        <AuthorCommentText>Not Provided</AuthorCommentText>
      </AuthorComments>
      - <DiscussionComments>
        <Comment>
          <CommentAuthor>Not Provided</CommentAuthor>
          <CommentDateCreated>Not Provided</CommentDateCreated>
          <CommentTitle>Not Provided</CommentTitle>
          <CommentText>Not Provided</CommentText>
        </Comment>
      </DiscussionComments>
      </ComponentComments>
  </ActivityNode>
  - <ActivityNode engineClass="com.insightful.miner.PredictEngineNode"
guiClass="com.insightful.miner.PredictNodeModel" numInputs="1" numOutputs="1" id="8">
    <DisplayInfo labelText="Logistic Prediction" defaultLabelText="Predict: Logistic Regression"
smallIcon="predict_small.gif" largeIcon="predict_large.gif" x="408" y="285" />
    <ArgumentList>
      - <XTProps>
        <Property name="useGlobal" value="true" />
        <Property name="useCache" value="global" />
        <Property name="propertiesSet" value="true" />
        <Property name="newColumns">
          <Property name="probability" value="true" />
          <Property name="classification" value="true" />
          <Property name="agreement" value="false" />
          <Property name="probabilityFor" value="forLastCategory" />
          <Property name="specifiedCategory" value="0" />
        </Property>
      </XTProps>
      </ArgumentList>
      - <ComponentComments>
        - <AuthorComments>
          <AuthorName>Edo</AuthorName>
          <AuthorCommentTitle>Not Provided</AuthorCommentTitle>
          <AuthorCommentDateCreated>May 19, 2007</AuthorCommentDateCreated>
          <AuthorCommentText>Not Provided</AuthorCommentText>
        </AuthorComments>
        - <DiscussionComments>
          <Comment>
            <CommentAuthor>Not Provided</CommentAuthor>
            <CommentDateCreated>Not Provided</CommentDateCreated>
            <CommentTitle>Not Provided</CommentTitle>
            <CommentText>Not Provided</CommentText>
          </Comment>
        </DiscussionComments>
        </ComponentComments>
  </ActivityNode>
</ActivityNode>
<Property name="setPropertyColumns">
    <Property name="independent" value="false" />
    <Property name="dependent" value="true" />
    <Property name="other" value="false" />
</Property>

<Property name="modellndependentVar">
    <Property name="Ccore" value="true" />
    <Property name="friction Corr Exp[ECCS]" value="true"/>
    <Property name="htc Core" value="true"/>
    <Property name="htc HXC" value="true"/>
    <Property name="K form System[ECCS]" value="true"/>
    <Property name="Mass Gass" value="true"/>
    <Property name="T core 0" value="true"/>
    <Property name="T gas 0" value="true"/>
    <Property name="Roughness" value="true"/>
    <Property name="max leakage" value="true"/>
    <Property name="HTCS" value="true"/>
    <Property name="HTCC" value="true"/>
    <Property name="Rupture Form Factor" value="true"/>
    <Property name="Shaft Moment Inertia" value="true"/>
    <Property name="C1" value="true"/>
    <Property name="C2" value="true"/>
    <Property name="C3" value="true"/>
    <Property name="HTC[Core]" value="true"/>
    <Property name="T fuel 0" value="true"/>
    <Property name="Cp" value="true"/>
</Property>

<Property name="modelOutputType" value="classification"/>
<Property name="modelAlgorithm" value="glm"/>
<Property name="specifiedRowsToChunk" value="0"/>
<Property name="rowsToChunk" value="10000"/>
<Property name="modelDependentVar" value="Failure Cladding"/>
</XTProps>
</ExecuteAfter id="2" />
<ModelParent id="2" />

- <AuthorComments>
  <AuthorName>Edo</AuthorName>
  <AuthorCommentTitle>Not Provided</AuthorCommentTitle>
  <AuthorCommentDateCreated>May 19, 2007</AuthorCommentDateCreated>
  <AuthorCommentText>Not Provided</AuthorCommentText>
</AuthorComments>

- <DiscussionComments>
  <Comment>
    <CommentAuthor>Not Provided</CommentAuthor>
    <CommentDateCreated>Not Provided</CommentDateCreated>
    <CommentTitle>Not Provided</CommentTitle>
    <CommentText>Not Provided</CommentText>
  </Comment>
</DiscussionComments>
<ActivityNode engineClass="com.insightful.miner.PredictEngineNode"
guiClass="com.insightful.miner.PredictNodeModel" numInputs="1" numOutputs="1" id="9">
<DisplayInfo labelText="Classification Tree Prediction" defaultLabelText="Predict: Classification Tree" smallIcon="predict_small.gif" largeIcon="predict_large.gif" x="405" y="435" />
- <ArgumentList>
- <XTProps>
  <Property name="useGlobal" value="true" />
  <Property name="useCache" value="global" />
  <Property name="propertiesSet" value="true" />
  - <Property name="modelIndependentVar">
    <Property name="Ccore" value="true" />
    <Property name="friction Corr Exp[ECCS]" value="true" />
    <Property name="htc Core" value="true" />
    <Property name="htc HXC" value="true" />
    <Property name="K form System[ECCS]" value="true" />
    <Property name="Mass Gass" value="true" />
    <Property name="T core 0" value="true" />
    <Property name="T gas 0" value="true" />
    <Property name="Roughness" value="true" />
    <Property name="max leakage" value="true" />
    <Property name="HTCS" value="true" />
    <Property name="Rupture Form Factor" value="true" />
    <Property name="HTCC" value="true" />
    <Property name="Shaft Moment Inertia" value="true" />
    <Property name="C 1" value="true" />
    <Property name="C2" value="true" />
    <Property name="HTC[Core]" value="true" />
    <Property name="T fuel 0" value="true" />
    <Property name="Cp" value="true" />
  </Property>
  <Property name="modelOutputType" value="classification" />
  <Property name="modelAlgorithm" value="tree" />
  - <Property name="newColumns">
    <Property name="probability" value="true" />
    <Property name="classification" value="true" />
    <Property name="agreement" value="false" />
    <Property name="probabilityFor" value="forSpecifiedCategory" />
    <Property name="specifiedCategory" value="1" />
  </Property>
  - <Property name="copyInputColumns">
    <Property name="independent" value="false" />
    <Property name="dependent" value="true" />
    <Property name="other" value="false" />
  </Property>
  <Property name="modelDependentVar" value="Failure Cladding" />
</XTProps>
</ArgumentList>
<ActivityNode engineClass="com.insightful.miner.ClassificationAssessmentEngineNode" guiClass="com.insightful.miner.ClassificationAssessmentNodeModel" numInputs="-1" numOutputs="0">
  <DisplayInfo labelText="Lift Chart" defaultLabelText="Lift Chart"
    smallIcon="lift_chart_small.gif" largeIcon="lift_chart_large.gif" x="661" y="375"/>
  <InputPortToolTips port1="Classifications"/>
  <ArgumentList>
    - <XTProps>
      <Property name="useGlobal" value="true"/>
      <Property name="useCache" value="global"/>
      <Property name="specifiedRowsToChunk" value="0"/>
      <Property name="rowsToChunk" value="10000"/>
      - <Property name="#0">
        <Property name="useRole" value="true"/>
        <Property name="fitted" value="true"/>
        <Property name="depColumn" value=""/>
        <Property name="evalColumn" value=""/>
      </Property>
      - <Property name="#1">
        <Property name="useRole" value="true"/>
        <Property name="fitted" value="true"/>
        <Property name="depColumn" value=""/>
        <Property name="evalColumn" value=""/>
      </Property>
    </XTProps>
  </ArgumentList>
</ActivityNode>
8.4 Appendix D: Linear and Non Linear Models used for Example III

The code reported below is created by means of the Vensim Software by Ventana Systems.

Number of Realizations =
100
~

Inputs[input number] =
10, 10, 10, 10, 10
~

input number:
input1, input2, input3, input4, input5
~

Linear Weighted Sum Inputs[input number] =
Inputs[input number] * weights[input number]
~

"Non-Linear Weighted Sum Inputs"[input number] =
(Inputs[input number] * weights[input number])
~

Output Linear =
SUM(Linear Weighted Sum Inputs[input number!])
~

Output Non Linear =
(SUM("Non-Linear Weighted Sum Inputs"[input number!]))^2
~

weights[input number] =
0.5, 0.5, 0.5, 0.5, 0.5
~

Failure Criterion 1 =
14
~
Failure Criterion2 =
  250
  ~
  ~

SM1 =
  -Output Linear + Failure Criterion1
  ~
  ~

SM2 =
  -Output Non Linear + Failure Criterion2
  ~
  ~

Failure Counter 1 =
  IF THEN ELSE(SM1 > 0, 0, 1)
  ~
  ~

Failure Counter 2 =
  IF THEN ELSE(SM2 > 0, 0, 1)
  ~
  ~

Cumulative Number of Failures 1 = INTEG(
  Failure Counter 1, 0)
  ~
  ~

Cumulative Number of Failures 2 = INTEG(
  Failure Counter 2, 0)
  ~
  ~

*******************************************************************************
.Control
*******************************************************************************

FINAL TIME = 100
  ~ Second
  ~ The final time for the simulation.

INITIAL TIME = 0
  ~ Second
  ~ The initial time for the simulation.

TIME STEP = SAVEPER = 1 Second [0,?] The time step for the simulation.