2016

Predicting Alarm And Safety System Performance Using Simulation

Ian Hunter Moskowitz
University of Pennsylvania, ianmosk@seas.upenn.edu

Predicting Alarm And Safety System Performance Using Simulation

Abstract

Safety is paramount to the chemical process industries. Because many processes operate at high temperatures and/or pressures, involving hazardous chemicals at high concentrations, the potential for accidents involving adverse human health and/or environmental impacts is significant. Thanks to research and operational efforts, both academically and industrially, the occurrences of such incidents are rare. However, disastrous events in the chemical manufacturing industry are still of relevant concern and garner further attention – the Deepwater Horizon incident (2010) and the Texas City refinery explosion (2005) being two recent examples.

Many techniques have been developed to understand, quantify, and predict alarm and safety system failures. In practice, hazards are identified using Hazard and Operability (HAZOP) analysis, and a network of independently-acting safety systems works to maintain the probabilities of such events below a Safety Integrity Level (SIL). The network of safety systems is studied with Layer of Protection Analysis (LOPA), which uses failure probability estimates for individual subsystems to project the failures of entire safety system networks.

With few alarm and safety system activations over the lifetime of a chemical process, particularly the critical last-line-of-defense systems, the failure probabilities of these systems are difficult to estimate. Statistical techniques have been developed, attempting to decrease the variances of such predictions despite few supporting data. This thesis develops methods to estimate the failure probabilities of rarely activated alarm and safety systems using process and operator models, enhanced by process, alarm, and operator data. Two repeated simulation techniques are explored involving informed prior distributions and transition path sampling. Both use dynamic process models, based upon first-principles, along with process, alarm, and operator data, to better understand and quantify the probability of alarm and safety system failures and the special-cause events leading to those failures.

In the informed prior distribution technique, process and alarm data are analyzed to extract information regarding operator behavior, which is used to develop models for repeated simulation. With alarm and safety system failure probabilities estimated for specific special-cause events, near-miss alarm data are used, in real-time, to enhance the predictions.

The transition path sampling method was originally developed by the molecular simulation community to understand better rare molecular events. Herein, important modifications are introduced for application to understand better how rare safety incidents evolve from rare special-cause events. This method uses random perturbations to identify likely trajectories leading to system failures – providing a basis for potential alarm and safety system design.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Chemical and Biomolecular Engineering

First Advisor
Warren D. Seider

Keywords
Bayesian Analysis, Process Reliability, Process Safety, Transition Path Sampling

Subject Categories
Chemical Engineering

PREDICTING ALARM AND SAFETY SYSTEM PERFORMANCE USING SIMULATION

Ian H. Moskowitz

A DISSERTATION

in

Chemical and Biomolecular Engineering

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

____

Warren D. Seider, Professor, Chemical and Biomolecular Engineering

Graduate Group Chairperson

____

John C. Crocker, Professor, Chemical and Biomolecular Engineering

Dissertation Committee

Raymond J. Gorte, Professor, Chemical and Biomolecular Engineering

Amish J. Patel, Assistant Professor, Chemical and Biomolecular Engineering

Ulku G. Oktem, Professor, Risk Management and Decision Process Center

Masoud Soroush, Professor, Drexel Chemical and Biological Engineering

Jeffrey E. Arbogast, Air Liquide International Expert, American Air Liquide

PREDICTING ALARM AND SAFETY SYSTEM PERFORMANCE USING SIMULATION

DEDICATION

To my family and Julie for their love and support.

ACKNOWLEDGEMENT

The ideas and methods presented herein are a result of the efforts of many, and would not be possible without the love and support of my friends and family.

This thesis represents the work of a wonderful collaboration between Drexel University, Near-Miss Management, Air Liquide, and of course, the University of Pennsylvania. Throughout the entirety of my graduate work, this group met every two months - I presented my research and we discussed the direction of the project. Many of the ideas and methods presented in this thesis were postulated and refined in these meetings – I am quite confident that the quality of this work would be severely less had I not had the unique opportunity to work with so many varied and bright individuals. Masoud Soroush and Taha Mohseni of Drexel University provided invaluable insight to the mathematical formulation of this work. Ulku Oktem and Ankur Pariyani of Near-Miss Management were crucial to keeping this thesis within the framework of industrial applications. Jeffery Arbogast, Darrin Feathers, Brian Besancon, and Benjamin Jurcik of Air Liquide provided crucial support to this project with data and frequent help in learning the techniques to analyze them. At the University of Pennsylvania, Amish Patel provided a key breakthrough when during my fourth year talk it struck him that the techniques being studied in his lab and field could be applied to ours. This realization, along with many insightful discussions and meetings, was critical to the development of this thesis. Ray Gorte, also of UPenn, brought key insight to the applicability of this work, and also provided me the theoretical framework to understand the complex reaction kinetics through his course.

My classmates, labmates, and friends kept me grounded through the swings of graduate research. Our weekly frisbee games, daily lunches, and evenings at Mad Mex were vital in keeping me focused and in high spirits. I had the help of several undergraduate researchers, noteably Eiman Soliman, Tony Barberio, Evans Molel, and Nicholas Baylis – each of whom had key contributions and helped me develop my own understanding of this material through a teaching role. I also need to thank Ankur again, along with Anjana Meel and James Philimister, as the students in our lab group that preceeded me on this project – themselves demonstrating the power of dynamic risk analysis (but still leaving areas to me to work on!). Cory Silva was my labmate for four of the five years I spent at school. When I joined the lab group, he brought me up to speed on many of the technical concepts in this thesis. I am grateful for his support in the lab, and even more so for his friendship outside of it.

My family has always emphasized school and without them I would have never been in a position to achieve this degree. Long before I learned how to perform numerical integration or write effective technical papers, you taught me how to count and you read me 'Goodnight Moon'. My parents, brother, and extended family supported me throughout my school career, and reminded me to harness my energy and competitive nature.

I especially need to thank Julie. I can imagine that dating a grad student for five years has a lot more drawbacks than it does upside, but you were always supportive of me pursuing my degree. You understood when I had to stay late at the lab, and when I had to work on weekends. When I'd come home feeling defeated, your incredible amount of energy would quickly make me forget about my school difficulties, and this allowed me to go into work each morning feeling refreshed and ready to go.

I am confident I will never be able to properly thank Warren Seider for the hours upon hours of advising, help, support, insight, and direction that he provided me during my graduate school career. Warren knew when to be patient, when to ask questions, when to push me, and when to give me space to be creative. Warren truly is a giant in the field of chemical process engineering. There wasn't a topic I stumbled across that he wasn't intimately familiar with, quick to provide the history of the field, the major contributors, key papers, and actionable steps I could take. Warren's passion for teaching and advising is infectious, in the times where I was struggling and felt like I could never get my research to work, I knew I could always rely on a conversation with Warren that would leave me with new ideas as well as new energy. Warren far exceeded his duties as an adviser – often chatting with me about politics, sports, relationships, and of course, our frequent tennis matches. Even though I will not be seeing Warren on a daily basis anymore, I am confident that our work and our friendship will continue for years to come.

Ian H. Moskowitz

Philadelphia

July 29, 2016

ABSTRACT

PREDICTING ALARM AND SAFETY SYSTEM PERFORMANCE USING SIMULATION

Ian H. Moskowitz

Warren D. Seider

Safety is paramount to the chemical process industries. Because many processes operate at high temperatures and/or pressures, involving hazardous chemicals at high concentrations, the potential for accidents involving adverse human health and/or environmental impacts is significant. Thanks to research and operational efforts, both academically and industrially, the occurrences of such incidents are rare. However, disastrous events in the chemical manufacturing industry are still of relevant concern and garner further attention – the Deepwater Horizon incident (2010) and the Texas City refinery explosion (2005) being two recent examples.

Many techniques have been developed to understand, quantify, and predict alarm and safety system failures. In practice, hazards are identified using Hazard and Operability (HAZOP) analysis, and a network of independently-acting safety systems works to maintain the probabilities of such events below a Safety Integrity Level (SIL). The network of safety systems is studied with Layer of Protection Analysis (LOPA), which uses failure probability estimates for individual subsystems to project the failures of entire safety system networks.

With few alarm and safety system activations over the lifetime of a chemical process, particularly the critical last-line-of-defense systems, the failure probabilities of these systems are difficult to estimate. Statistical techniques have been developed, attempting to decrease the variances of such predictions despite few supporting data. This thesis develops methods to estimate the failure probabilities of rarely activated alarm and safety systems using process and operator models, enhanced by process, alarm, and operator data. Two repeated simulation techniques are explored involving *informed prior distributions* and *transition path sampling*. Both use dynamic process models, based upon first-principles, along with process, alarm, and operator data, to better understand and quantify the probability of alarm and safety system failures and the *special-cause events* leading to those failures.

In the informed prior distribution technique, process and alarm data are analyzed to extract information regarding operator behavior, which is used to develop models for repeated simulation. With alarm and safety system failure probabilities estimated for specific special-cause events, near-miss alarm data are used, in real-time, to enhance the predictions.

The transition path sampling method was originally developed by the molecular simulation community to understand better rare molecular events. Herein, important modifications are introduced for application to understand better how rare safety incidents evolve from rare special-cause events. This method uses random perturbations to identify likely trajectories leading to system failures – providing a basis for potential alarm and safety system design.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Despite much attention and many efforts, accidents in the chemical manufacturing industries are relevant, costly, and occasionally fatal. In the past four years, over 100 fatalities have occurred in the United States due to a wide variety of accidents ("Worker Fatalities to Federal and State OSHA"). There have been incidents in the past decade that have drawn much attention due to their severe nature – BP's Deepwater Horizon oil spill ("U.S. Chemical Safety Board Report: BP Deepwater Horizon"), Texas City refinery explosion ("U.S. Chemical Safety Board Report: BP America Refinery Explosion"), and the Kleen Energy Systems explosion ("U.S. Chemical Safety Board Report: Kleen Energy Natural Gas Explosion"), to name a few. Each of these accident scenarios involves two critical similarities – an unexpected event occurred, and the event was not handled properly by operators and plant managers (Kletz, 2009). Because many chemical plants involve dangerous chemicals, high temperatures, high pressures, or are in environmentally fragile areas (e.g., the Gulf Coast), the impacts of accidents can be quite large. The Texas City refinery explosion claimed the lives of 14 workers and injured over 100 more. The BP Deepwater Horizon oil spill devastated the environment along much of the Gulf Coast, and was one of the most costly accidents ever, having damage estimates as high as 42 billion dollars. While these events are rare, their impact is sufficiently high to warrant further research aimed at predicting, mitigating, and preventing these accidents.

The typical approach to preventing accidents in a chemical manufacturing process involves process design coupled with design of operating strategies, process controlsystems, and safety systems. Processes can be designed such that they are inherently less sensitive to disturbances in process units and feedstock fluctuations. This approach, known as inherently safer design (ISD), often varies process-to-process, with specific process units or features installed to handle potential accidents before they develop (Hendershot, 2006). On the inlet of sensitive reactors, it is common for designers to introduce buffer tanks to dampen deviations in feed flow rates, compositions, temperatures, and pressures. Separation units commonly involve extra trays, bed depth, or membrane areas – permitting continued operation in the face of large disturbances. Some units are designed to be used only when a problem arises in a plant. In many cases involving pipes designed for gas flow, a pressure-release line is installed. When the pressure exceeds an upper bound, gas can be redirected to the release line and flared so that it doesn’t rupture a pipe. Stop valves are typically installed on the inlet and outlet of sensitive processing units – allowing operators to isolate problems that occur upstream of the unit or within the unit. Various indices and statistical approaches for quantifiably evaluating and rationalizing ISD have been developed (Srinivasan et al., 2012).

Disturbances in a plant occur on a frequent basis, often minute-to-minute, and need to be handled in an efficient manner. While process design features can help to dissipate disturbances, they are often not responsible for arresting them. This is the role of the process control system. PID controllers are the most basic – a variable is measured, and based on its deviation from its setpoint, the controller typically opens or closes a valve in part or in full (Luyben, 1989; Stephanopoulos, 1984). Here, the controller must be tuned properly, and the measuring device and actuator must be functioning properly. If not, there is potential for the disturbance to propagate further. Control configurations involving PID controllers have been developed, such as cascade or feedforward controllers. These provide tighter and/or more robust process control, assuming that the measuring devices and actuators are working properly. Model-predictive controllers use first-principle or empirical models to yield actuator responses that minimize deviations from set points over the predictive horizon (Garcia et al., 1989). They often improve controllability, but process-model mismatch may keep controllers from adequately arresting disturbances.

When the process design features and control systems are insufficientto regulate a disturbance, the operator, often in response to alarms, is responsible for any corrective actions to move the process back to typical operating conditions with a safety interlock system shutting down the process when it deviates sufficiently far from these conditions (Crowl et al., 2001). Operators typically have the ability to make adjustments to decision variables in a process, open and close valves, and switch control systems on and off, and are aided by a network of alarms. When alarms activate to notify operators that process variables have crossed thresholds, the operators are expected to: (1) diagnose the root cause of the problem, and (2) make appropriate corrective actions to mitigate the consequences (Hollifield et al., 2010). This can be a difficult task, particularly when the root cause problem is shrouded; i.e., the process is undergoing inverse response or there is an undetected rare event occurring.

In addition to the operator, there is an automated safety interlock system. Interlocks work to shut down the plant automatically when specific process variables, called *primary variables*, cross defined thresholds. The automatic safety interlock system is important because it shuts down the process before safety systems, such as quench tanks or relief valves, are activated as a last line of defense in preventing the process from entering a runaway reaction mode where human health and environmental catastrophes are possible. Plant operator actions are important in the continued operation of a process, and **safer operation is realized when plant operators are effective in preventing processes from undergoing shutdown (and associated restart) and activating crucial safety systems.**

Alarms are placed on process variables to alert operators that the process is deviating from its expected regime(s) of operation. A typical alarm has a low-threshold (for L alarms) and a high-threshold (for H alarms) that bound the range of typical operation. When the measured variable moves outside these thresholds, an alarm is activated and a *special-cause event* has occurred. The L and H-alarm thresholds, along with more severe alarm thresholds, are established during the commissioning phase of a process, typically the first one to three years of operation. During the design phase, several measured variables are chosen as primary variables. Strong candidates for primary variables are those that best describe the safety of the process – often, the measurements associated with the most potentially dangerous operations (i.e., process units at high temperature or pressure, or containing hazardous chemicals). Ideally, safety interlocks are activated only when these variables move into unsafe regimes (Rothenberg, 2009). The choice of alarm thresholds and primary variables has a major impact on the effectiveness of the operator response to alarms to reliably maintain safe operation.

Areas of unsafe operation are commonly determined using hazard and operability, HAZOP, analyses (Kletz, 1999). This common and systematic approach is intended to determine all potential hazards to process units. All potential material inlets (through designed inlet ports and backflow through outlet ports, as well as leaks through the vessel walls) are considered, and the potential chemical reactions are postulated. Mechanical failures to piping and valves and electrical failures to compressors, motors, and control systems are also considered. HAZOP has long been performed as a qualitative approach, but computer-based HAZOP approaches and algorithms have been developed, in an effort to reduce the amount of human error that arises during the hazard identification procedure (Venkatasubramanian et al., 1994; Palmer et al., 2008). Human error and "safety culture" has been incorporated into HAZOP approaches, with operator mistakes and failures studied as potential causes of hazards to process operation (Kennedy et al., 1998). The qualitative analysis is then enhanced using quantitative statistics – the failure rates of similar process units are used to gain an understanding of the most severe process risks. This analysis is often the basis for determining the primary variables in the process. Process variables associated with the greatest potential hazards or risks are chosen as primary variables, ensuring that an automatic shutdown is attempted when these variables are far outside their typical operating regions.

With potential hazards to process operation identified, independently-acting safety systems are installed to maintain the probability of failure below a pre-specified *Safety Integrity Level* (SIL) (Dowell, 1998). The independently-acting safety systems are commonly evaluated using event-trees, where the probability of the network of safety systems failing is the product of the failure probability of each activated safety system (Andrews et al., 2000; Phimister et al., 2003). As illustrated using the "Swiss Cheese Model", an accident occurs when the various levels of safety systems fail or are insufficient (Reason, 1990).

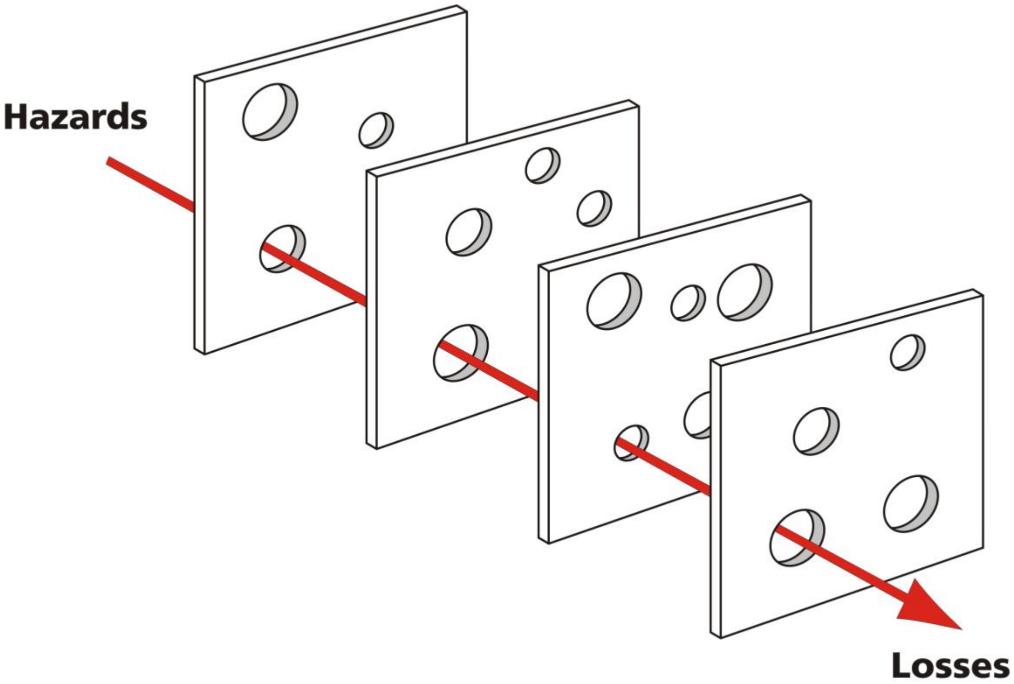


Figure 1.1. Swiss cheese model

Layer of Protection Analysis (LOPA), is the industry standard to quantify the accident probability for specific special-cause events, typically indentified during HAZOP (Summers, 2003). This quantitative procedure is valuable in characterizing the safety of a process during a special-cause event. More recently, techniques to evaluate the process's safety through a period of human error have been developed (Baybutt, 2002; Baybutt, 2003). Various techniques to quantify the failure probability of individual safety systems and the network of safety systems have been developed, all sharing the challenge of few safety system activations over the lifetime of a process. Bayesian networks (Marsh et al., 2008) and neural networks (Ruilin et al., 2010) have been utilized to quantify these failure probabilities.

While LOPA estimates the probability of safety system failure, Fault Tree Analysis (FTA) estimates the probability of special-cause event occurrence. The varying paths leading to a special-cause event are identified and process statistics are used to characterize the probability of such an event occurring (Khakzad et al., 2011; Tanaka et al., 1983). These estimates can be combined with previous event-tree approaches for analyzing the failure probability of the safety system network during a special-cause event. This "bow-tie" approach tracks the special-cause event from its root-cause through the safety system activation (Cockshott, 2005).

In some cases, alarms are officially considered a layer-of-protection and contribute to the SIL rating of the overall safety system. Therefore, the alarms are included in the safety-systems discussed herein – noting that often the full alarm system is not considered part of a plant's safety instrumented system (SIS). The failure probabilities of specific safety systems, as well as the network as a whole, are often difficult to estimate – **the activation of most safety systems occur infrequently**, and oftentimes the root-cause of the event is poorly understood. If the failure probabilities of safety systems, could be known with certainty, the probability of accidents at a process could be guaranteed below the SIL with proper safety system design. Various techniques and methods for quantifying the failure probabilities of rarely activated safety systems have been developed, and this thesis explores new techniques in this area.

Dynamic Risk Analysis (DRA) is used to update risk estimates over the lifetime of the plant (Meel et al., 2006; Kalantarnia et al., 2009). As process and alarm data are collected, in real-time, DRA updates the risk estimations that were made during the design and commissioning phases. Typically Bayesian statistics (Bayesian analysis) are used to generate failure probability estimates using alarm data (Pariyani et al., 2012a). The Bayesian approach has the potential to generate failure probabilities having lower variance than those achieved using classical statistics, and is explained in Chapter 2.

DRA performs best in describing the risk of frequently activated safety systems – with more data available, estimates with narrower confidence intervals can be made. For infrequently used systems, copulas have been introduced to make risk estimates with smaller variances (Pariyani et al., 2012b; Yi et al., 1998). Copulas describe the dependence between the more frequently-activated, low-consequence systems with infrequently-activated, high-consequence systems.

While dynamic risk analysis and copulas are effective in making meaningful risk estimates for many infrequently-used systems, data may be too sparse to permit copulas to reduce the variance of risk estimates sufficiently. Many processes, such as the steam-methane reformer studied herein, are well-understood, and special-cause events are generally handled reliably by plant operators. This thesis explores model-based approaches for better understanding the failure probabilities of operator responses to alarms that rarely lead to safety interlock activations and associated plant shutdowns. Process models, while not a perfect representation of the process, can be simulated many times, generating a large pool of simulated alarm and safety interlock activations. These statistics can then be enhanced with process and alarm data, when available, to improve the failure probability predictions. Various sampling techniques are developed and applied to safety systems. In particular, this thesis explores *informed prior distributions* and *transition path sampling*. These sampling techniques utilize both process and operator models, enhanced by process and alarm data collected at the plant. Pathways, or trajectories, to safety interlock activations are explored. While the safety interlock activations investigated are inherently rare, the failures have the potential to be catastrophic in the unlikely event that safety interlock systems fail. At best, the safety interlock system activations are expensive due to lost product and process shutdowns. The three chapters describing these techniques are briefly introduced in the next three sections.

1.2 Chemical Process Simulation for Dynamic Risk Analysis: Developing Informed Prior Distributions

Chapter 2 describes how dynamic simulations of a manufacturing process can be used to construct *informed prior distributions* for the failure probabilities of alarm and safety interlock systems. Bayesian analysis is used starting with prior distributions and enhancing them with likelihood distributions, constructed from real-time alarm data, to form posterior distributions, which are used to estimate failure probabilities. The use of alarm data to build likelihood distributions has previously been investigated. Rare-event historical data are typically sparse and have high-variance likelihood distributions. When high-variance likelihood distributions are combined with typical high-variance prior distributions, the resulting posterior distributions naturally have high variances yielding unreliable failure predictions. In contrast with prior distributions obtained by maximizing entropy and those that are based on expert knowledge, this chapter introduces a repeated-simulation method to construct informed prior distributions having smaller variances, which in turn yield posterior distributions with lower variances and a more reliable prediction of the failure probabilities of alarm and safety interlock systems. The application of the proposed method is demonstrated for the offline dynamic risk analysis of a steam-methane reformer (SMR) process.

1.3 Improved Predictions of Alarm and Safety System Performance Using Process and Operator Response-Time Modeling

In Chapter 2, a repeated-simulation process-model-based technique for constructing informed prior distributions is introduced. The models used in simulation are crucial to the low-variance risk predictions generated by the sampling technique. This chapter investigates the effect modeling has on the risk predictions, and how both process and operator models can be systematically improved to generate more accurate risk predictions. This chapter presents a method of quantifying process model quality, which impacts prior and posterior distributions used in Bayesian Analysis. The method uses higher-frequency alarm and process data to select the most relevant constitutive equations and assumptions. New data-based probabilistic models that describe important special-cause event occurrences and operators’ response-times are proposed and validated with industrial plant data. These models can be used to improve the estimates of failure probabilities for alarm and safety interlock systems.

1.4 Understanding Rare Safety and Reliability Events Using Transition Path Sampling

There is strong motivation to understand how rare reliability and safety-events develop and propagate. Effective operator training, safety system design, and safety analysis, all benefit from a full understanding of such events. A major challenge in the study of events that propagate to process shutdowns or safety incidents is their sparsity – typically these events occur so rarely that statistical techniques alone are incapable of describing and characterizing them – especially when they have not yet occurred. Simulation of these events could be useful to understand them, however, a daunting computational challenge exists. Typical rare events occur on the order of years or decades apart, while the events occur within minutes or hours. Thus, the bulk of the computational effort in simulating rare events is allocated to normal operation, making the events computationally infeasible to simulate with meaningful frequencies.

A rare-event sampling technique, Transition Path Sampling (TPS), has been developed by the molecular dynamics community. While the time and length scales between molecular dynamics and process dynamics differ greatly, the ratios of the times of the rare events and the waiting times between them are similar. This Monte-Carlo based technique relies on the simulation of perturbed rare-event trajectories – an initial rare-event trajectory is randomly modified such that large numbers of trajectories are generated. Clusters of rare safety-event trajectories are the basis for alarm and safety-system design, assuring that TPS-generated clusters are preventable. Important modifications to the TPS technique are needed to apply it to process dynamics. The backwards integration, a key attribute of TPS, is not possible for most process simulations – instead a boundary-value optimization technique is used. Furthermore, process models use vast amounts of process data for model verification and to estimate the relative likelihood of one trajectory to another. The application of TPS is demonstrated using a simple jacketed exothermic CSTR, as well as a more complex air separation process. This innovative approach allows for a quantitative rationalization of alarm and safety systems to reduce the occurrence of rare, yet serious, safety events.

Chapter 2

Chemical Process Simulation for Dynamic Risk Analysis

2.1. Introduction

The design of accurate process models and optimal flowsheets have challenged process systems engineering researchers for decades – often involving optimizations with decision variables (such as feed-stock or operation variables) adjusted to increase revenue, decrease cost, or increase profit (Seider et al., 2009). From a controls perspective, controller parameters are tuned to improve performance measures (Seborg et al., 2010). Furthermore, superstructures are used to determine which process units and controllers should be included for optimal functionality (Yeomans et al., 1999). But, process models and flowsheets have been under-investigated in the process safety area, where process engineers are challenged to reduce the risk of incidents, the most serious of which may be classified as accidents. Process incidents, resulting in human-health losses, environmental losses, and capital losses, are expensive and occasionally tragic (when safety systems are insufficient to prevent process incidents from becoming process accidents) (U.S. Chemical Safety and Hazard Investigation Board; Process Safety Incident Database).

To design and operate a process with reduced incident and accident risk, it is crucial to quantify the probabilities of incidents. This can be a difficult task, as it involves: (1) determining the probability of each special-cause event, (2) determining the probability of each consequence arising from each special cause, and (3) evaluating the severity of each consequence (Pariyani et al., 2010; Mannan et al., 1999). To quantify accurately the overall risk of an incident, these three tasks are required for every special cause, consequence, and loss, providing quite a daunting challenge! The success or failure of an alarm system depends upon the success or failure of operator actions taken in response to an activated alarm. In contrast, the Safety Instrumented System (SIS) takes automatic actions such as a shutdown initiated by an interlock. In this paper, the focus is on simulating the effects of special cause events to inform and improve design and operation decisions to mitigate incidents. In this manner, process engineers and operators can make more informed decisions to reduce plant risk (Phimister et al., 2003; Jones et al., 1999).

Emphasis is placed on constructing sufficiently accurate process simulations to evaluate plant safety, given measured process and alarm data. Clearly, special attention is needed: (i) in the most risky plant areas, and (ii) when special-cause events are likely to be amplified or masked (Rosenthal et al., 2006). The former typically involve high temperatures, pressures, and hazardous chemicals. The latter are more difficult to identify, especially when their responses occur in rapid transients. Masked responses include inverse responses and delays (dead-times) which may lead operators to take incorrect action in response to alarms. Here, dynamic, first-principles, process models, built with knowledge from historical process and alarm data (Chen et al., 1998), can help operators respond better to these special-causes. While first-principles models have long been used in the chemical process industries to enhance an understanding of processes

(Soroush et al., 1992), this paper provides a new method to estimate the failure probabilities of alarm and safety interlock systems.

The rest of this chapter begins with a discussion of typical alarm and safety interlock systems and their associated event trees and failure probabilities. Next, Bayesian analysis is reviewed, followed by the presentation of a new method that uses dynamic simulations to create informed prior distributions for Bayesian analysis. Then, a detailed steam-methane reforming (SMR) model integrated with a pressure-swing adsorption (PSA) model is presented and the proposed method is demonstrated by simulating the combined model. To our knowledge, no published integrated SMR-PSA model exists including recycle of the PSA-offgas to the SMR fuel system. Finally, conclusions are drawn with recommendations for future work.

2.2. Safety Systems and Event Trees

An abnormal event occurs when a process variable leaves its normal operating range (green-belt zone in Figure 2.1), which triggers an alarm indicating transition into the yellow-belt zone. If the variable continues to move away from its normal range, the variable may transition into its red-belt zone, indicated by a second-level alarm (e.g., LL, HH) activation. Once a variable remains in its red-belt zone for a pre-specified length of time (typically on the order of seconds), an interlock activates and an automatic shutdown occurs (Hosseini et al., 2007).

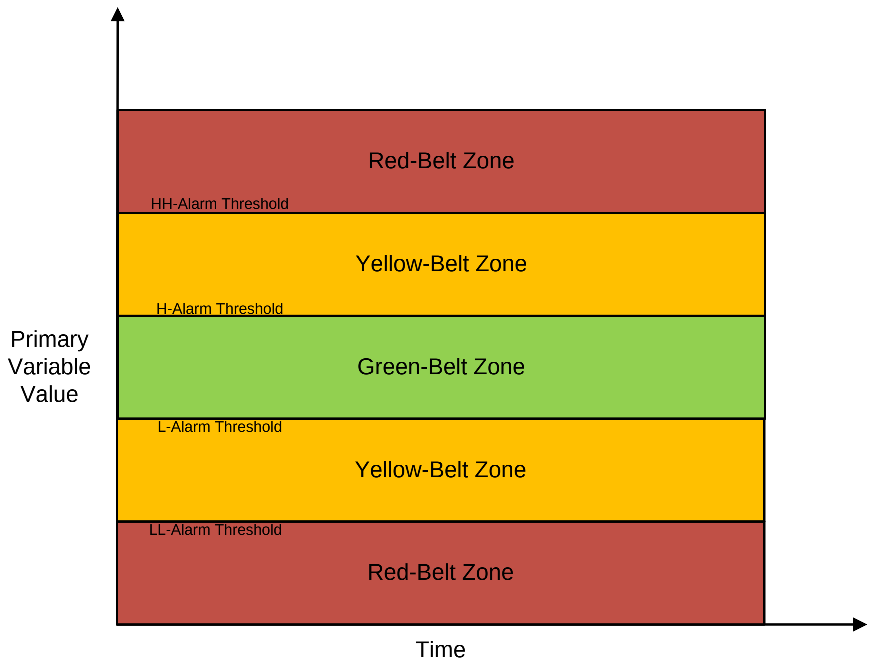


Figure 2.1. Belt-zone map for primary variables.

An event-tree corresponding to a primary variable's transition between belt-zones is shown in Figure 2.2. The first-level (e.g., L, H) alarm system activates safety-system 1 (SS_1), which is typically an operator action. When SS_1 is successful, with probability 1-x_1, continued operation, consequence $C_{1,}$ is achieved. The second-level (e.g., LL, HH) alarm system activates SS_2, which is typically a more aggressive operator action. When successful, with probability 1-x_2, near-miss continued operation, consequence C_2, is achieved. If the primary variable occupies the red-belt zone for a pre-determined length of time (on the order of seconds), SS_3, the automatic interlock plant shutdown, will become activated. The interlock system is designed to be independent of alarm systems, and the activation of SS_3 is determined by an independent set of sensors. It should be noted that if the interlock system is designed to have no delay time, the probability of SS_2

success is equal to zero (x_2 = 1). If SS_3 succeeds, with probability 1-x_3, the interlock shutdown is successful and an accident is avoided, represented by consequence C_3. If the interlock shutdown is unsuccessful, an accident occurs at the plant, represented by C_4. With proper design, x_3 should be very small consistent with the specified Safety Integrity Level (SIL) (Stavrianidis et al., 1998; Stavrianidis et al., 2000). Since the interlock system is independent of the alarm system, the success of SS_3 will not depend on factors such as operator skill or alarm sensor fault. However, it can be concluded that if either SS_1 or SS_2 are successful in arresting the special-cause event, the activation of the interlock system can be avoided altogether. In some cases, alarms are officially considered a layer-of-protection and contribute to the SIL rating of the overall safety system, composed of SS_1, SS_2, and SS_3. Therefore, the alarms are included in the safety-systems herein – noting that often the full alarm system is not considered part of a plant's SIS.

In this way, event trees represent the actions of various alarm and safety interlock systems and their end consequences after abnormal events (Meel et al., 2006). For dynamic risk analyses, alarm and interlock actions must be chronologically tracked and recorded (using the plant alarm historian). Using data compaction techniques and Bayesian analyses, failure probabilities of the alarm and safety interlock systems and the probabilities of plant incidents (Pariyani et al., 2012a; Pariyani et al., 2012b) have been estimated.

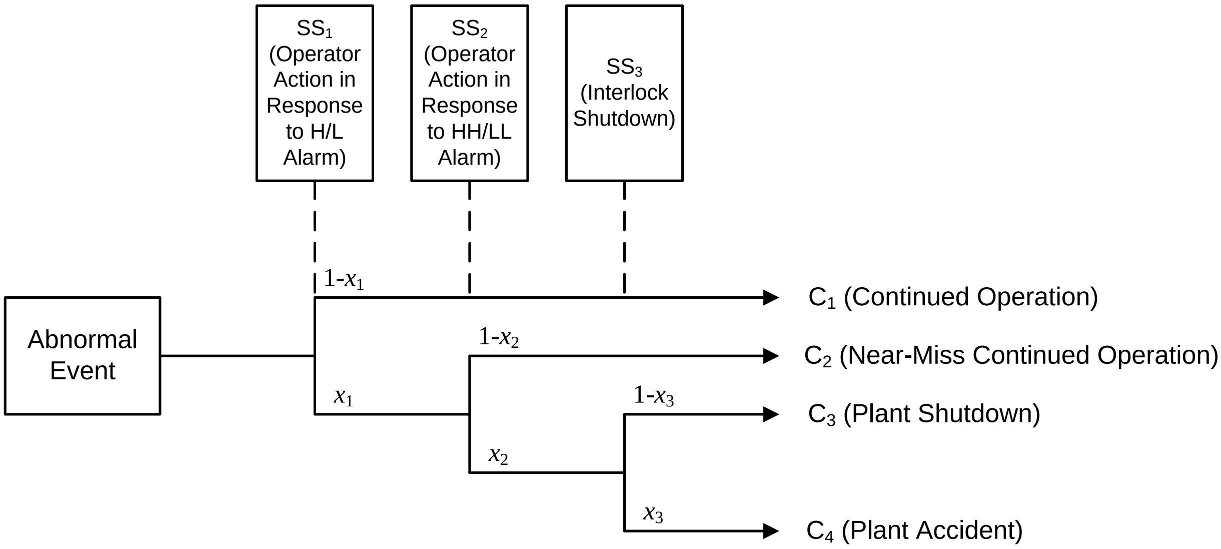

SS_1 (Operator Action in Response to H/L Alarm)
SS_2 (Operator Action in Response to HH/LL Alarm)
SS_3 (Interlock Shutdown)
Abnormal Event
$1-x_1$
x_1
$1-x_2$
x_2
$1-x_3$
x_3
C_1 (Continued Operation)
C_2 (Near-Miss Continued Operation)
C_3 (Plant Shutdown)
C_4 (Plant Accident)

Figure 2.2. Event tree involving three safety systems.

2.3. Bayesian Analysis

Bayesian analysis is often used to determine the failure probabilities of alarm and safety interlock systems. The central dogma of Bayesian analysis is that random-variable distribution parameters (e.g., mean and variance) are themselves distributions. Unlike classical statistics that seeks to capture the true moments of a distribution, Bayesian statistics acknowledges that the moments of a distribution may not be fixed, and seeks to estimate the probability distributions of the moments. This analysis often requires significantly fewer data to make meaningful predictions (Gelman et al., 2014; Berger, 2013). Additionally, as the process dynamics and operators' behavior change with time (because of factors such as process unit degradation and operators' improved skills), real-time data can be collected and used to estimate more accurate failure probabilities in real time.

In Bayesian analysis, the *posterior distribution*, represented as $f(x_i|D)$, the probability distribution of $x_i \in [0,1]$ given the collected data, D, is calculated using Bayes' rule:

$$f(x_i|D) = \frac{f(x_i)f(D|x_i)}{\int_0^1 f(x_i)f(D|x_i)dx_i} \tag{2.1}$$

where $f(x_i)$ is the *prior distribution* of x_i, estimated before data are collected, and $f(D|x_i)$ is the *likelihood distribution* of the data given x_i. The prior distribution is normally estimated using expert knowledge or maximum entropy techniques (Ahooyi et al., 2014). The likelihood distribution captures the probability that the data could have been generated, if the failure probability was equal to x_i, as discussed next.

Herein. a beta distribution is used to represent an informed prior distribution, which is constructed using process simulations:

$$f(x_i) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1}(1-x_i)^{\beta-1} \tag{2.2}$$

where α and β are parameters obtained through simulation, and $\Gamma(z)$ is the gamma function:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \tag{2.3}$$

The beta distribution is well suited to represent a safety-system failure-probability distribution because its domain is [0, 1], and its two parameters can be estimated from only two moments (e.g., the mean and variance) of simulated data.

The alarm data provides a record of each safety system activation, which can be tracked to its failure or success. The binary performance lends itself to being described using a binomial likelihood distribution:

$$f(D|x_i) = \frac{n!}{k!(n-k)!} x_i^k (1 - x_i)^{n-k} \tag{2.4}$$

where D represents the alarm data, n is the number of safety system activations, and k is the number of safety system failures. When Eqs. (2.2) and (2.4) are substituted into Eq. (2.1), the posterior distribution for x_i, given n and k is:

$$f(x_i|n,k) = \frac{\frac{\Gamma(\alpha+\beta)n!}{\Gamma(\alpha)\Gamma(\beta)k!(n-k)!} x_i^{k+\alpha-1}(1-x_i)^{n-k+\beta-1}}{\int_0^1 \frac{\Gamma(\alpha+\beta)n!}{\Gamma(\alpha)\Gamma(\beta)k!(n-k)!} x_i^{k+\alpha-1}(1-x_i)^{n-k+\beta-1} dx_i} \tag{2.5}$$

This is a beta distribution with parameters $a = (k + \alpha)$ and $b = (n - k + \beta)$, recognizing that f is a function of x_i only. Note that for the beta distribution in Eq. (2.2), a = α and b = β. Consequently, using the identity:

$$\int_0^1 x_i^{k+\alpha-1}(1-x_i)^{n-k+\beta-1}dx_i = \frac{\Gamma(k+\alpha)\Gamma(n-k+\beta)}{\Gamma(n+\alpha+\beta)} \tag{2.6}$$

the posterior distribution in Eq. (2.5) simplifies to the beta distribution:

$$f(x_i|n,k) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(k+\alpha)\Gamma(n-k+\beta)} x_i^{k+\alpha-1}(1-x_i)^{n-k+\beta-1} \tag{2.7}$$

As alarm data are collected in real time, the alarm statistics can be updated in real time (Meel et al., 2006; Khakzad et al., 2012; Kalantarnia et al., 2009). In so doing, process engineers gain a better understanding of how the process is performing (Pariyani et al., 2012b).

2.4. Constructing Informed Prior Distributions

The proposed method of construction of informed prior distributions has the eight steps listed in Table 1. In Steps 1-3, a robust, dynamic, first-principles model of the process incorporating the control, alarm and safety interlock systems, is built. The model can then be simulated using a simulator such as gPROMS (gPROMS v.3.6.1; Oh, et. al., 1996), which is used herein. The control system in the model mimics the actual plant control system, with consistent control logic and tuning. Likewise, the alarm and safety interlock systems in the model mimic those in the plant. For operator actions, this can be difficult, as operators often react differently to alarms. In particular, expert operators may take into account the state of the entire process when responding to alarms. When creating a model, the likelihood of operator actions must be considered. Either the modeler can use the action most commonly taken by operators, or a stochastic simulation can be set up in which the different actions are assigned probabilities.

With these models, special-cause events are postulated in Step 4. The list of special-cause events can be developed from various sources: HAZOP or LOPA analysis, observed accidents in the plant (or a similar plant), near-miss events at the plant (or in a similar plant), or from risks suggested in first-principles models of the plant. For each special-cause event, an event magnitude distribution is created in Step 5. A distribution for operator response time, τ, is created in Step 6. These three distributions are used along with the dynamic simulation in Step 7 to obtain simulation data. Lastly, in Step 8, the simulated data is used to regress parameters for the informed prior distribution. The algorithm used to generate simulation data (Step 7) and regress informed prior distribution parameters (Step 8) is described in the paragraph below, and represented pictorially in Figure 2.3.

The script that manages the dynamic simulations starts by sampling A_1 from the event magnitude distribution created in Step 5. Note that Figure 2.3 shows a Normal distribution centered at μ_{SC} with variance σ^2_{SC}, however any distribution can be used. Assign the number of safety system failures, i, to i = 0. With this A_1, the user script samples τ_1 from the distributions created in Step 6. Although Figure 2.3 shows Uniform distributions (with the maximum operator response time at τ_{max}), any distribution can be used. With A_1 and τ_1, a dynamic simulation is run. If the safety system fails to avoid a plantwide shutdown, then $i = i + 1$; if the safety system is successful, i is not incremented. When $n < N$, $n = n + 1$; i.e., for sampled A_i and τ_i, a dynamic simulation is run, and i is adjusted when necessary. After N iterations, $j_1 = i/N$ is calculated, in the range [0,1]. Then m is incremented and A_m sampled, the inner loop is re-executed, and j_m is calculated. When the outer loop has been completed ($m = M$), a vector of M elements (j_1, ..., j_M) has been accumulated. The average and variance of this vector is used to calculate α and β of the Beta distribution. Note that because the Beta distribution is the conjugate prior of the binomial likelihood distribution, it is the recommended choice. The number of simulations, $M \times N$, is chosen, recognizing that more simulations yield a smaller prior-distribution variance.

Table 2.1. Steps to Construct an Informed Prior Distribution

1. Develop a dynamic first-principles process model
2. Incorporate control system into the dynamic process model
3. Incorporate the alarm and safety interlock system into the dynamic process model
4. Postulate potential special-cause events to be studied
5. For each special-cause event, construct a distribution for the event magnitudes, A_{SC} (i.e., for a postulated pressure decrease, construct a probability distribution for a decreasing magnitude)
6. For each special-cause event, construct a distribution for operator response time, τ.
7. For each special-cause event, conduct the simulation study according to the algorithm described in Figure 2.3 to simulate the range of possible event magnitudes, A_{SC}, and operator response time, τ.

8. Estimate parameters of a distribution model (e.g., Beta distribution) representing the data generated in Step 7 – this is used as an informed prior distribution (Gelman et al., 2013).

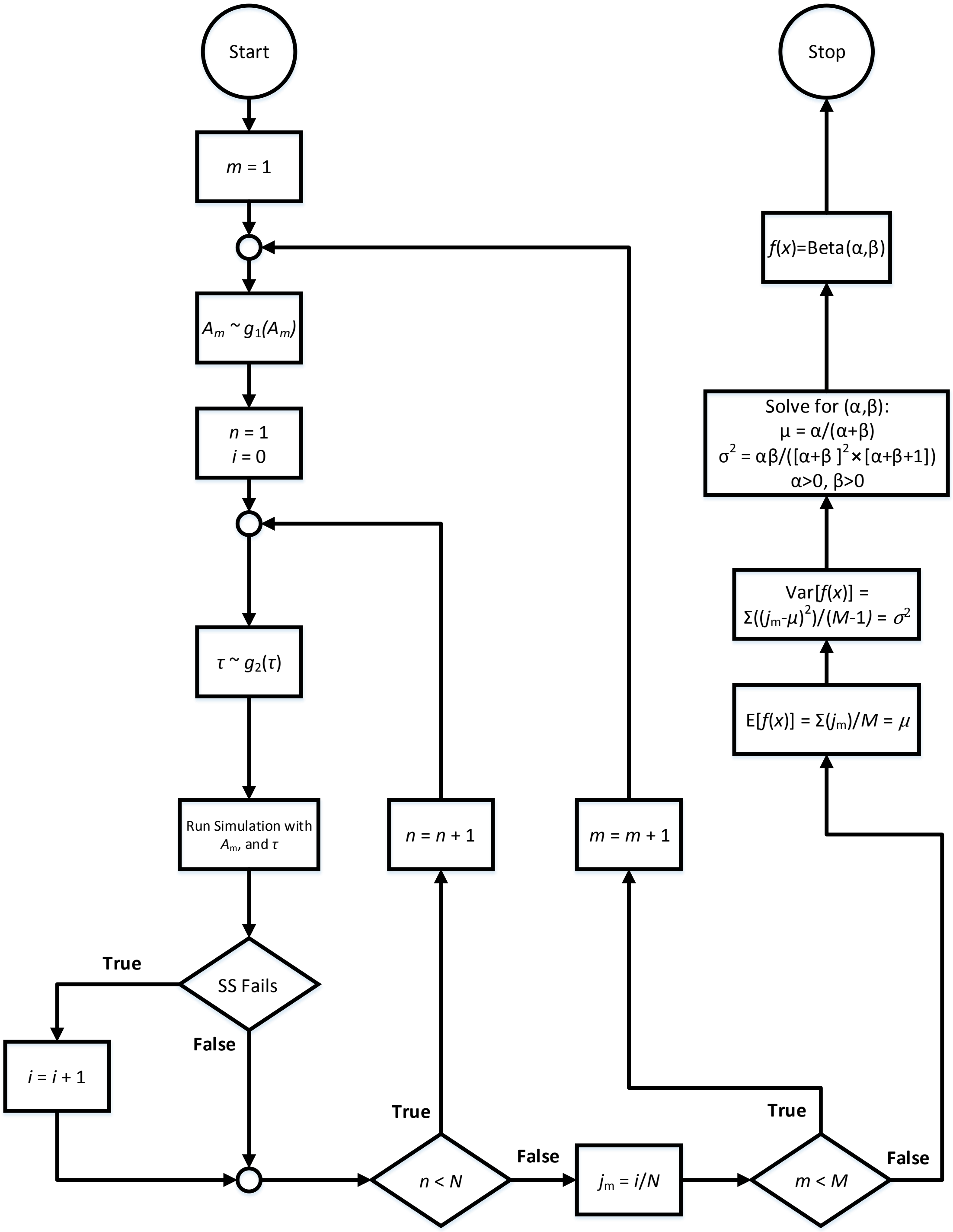


Figure 2.3. Sampling algorithm used in Steps 7 and 8 in Table 2.1.

2.5. Steam-Methane Reforming (SMR) Process

A typical SMR process is shown in Figure 2.4. After pretreatment, natural gas feed (70) and steam (560) are mixed before entering the process tubes of an SMR unit (90), where hydrogen, carbon monoxide, and carbon dioxide are produced. This hot process gas (100) is then cooled and sent to a water-gas shift converter (110), where carbon monoxide and water are converted to hydrogen and carbon dioxide. The process gas effluent (120) is cooled in another heat exchanger, producing stream 170, which is sent to two water extractors. Note that the last section of this heat exchanger is used to transfer heat to a boiler feed water makeup stream in an adjacent process. The gaseous hydrogen, methane, carbon dioxide, and carbon monoxide, in stream 210, are sent to PSA beds. Here, high-purity hydrogen is produced (220), and the PSA-offgas is sent to a surge drum. Stream 800 from the surge drum is mixed with hot air (830) and a small amount of natural gas makeup (815), and sent to the furnace side, where it is combusted to provide heat to the highly-endothermic process-side reactions. Its hot stack gas (840) is sent through an economizer, where it is used to heat steam (520), some of which is used on the process side (560), with the rest available for use or sale as a steam product (570).

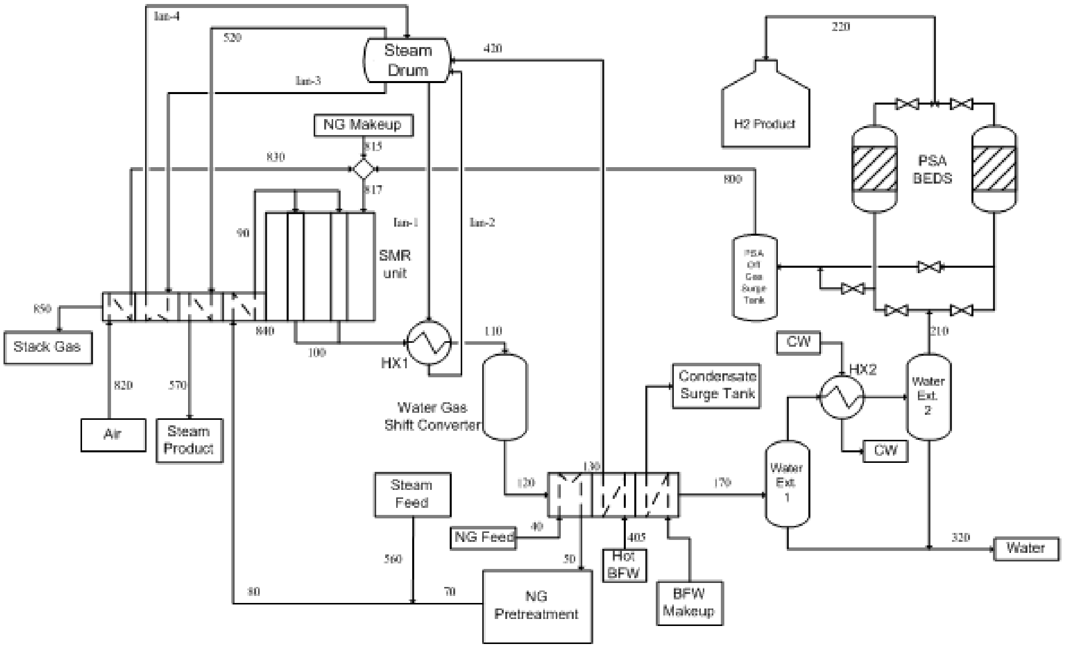


Figure 2.4. SMR process flow diagram

In modeling for process safety, emphasis should be placed on units that present the greatest risk; i.e., have the largest probabilities of incidents multiplied by incident cost (Kalantamia et al., 2009). In an SMR process, temperatures rise above 1,300 K with pressures over 20 atm. Because overheating can lead to process-tube damage and failure, potentially leading to safety concerns, its model received special attention in this work. Partial differential and algebraic equations (PDAE's), that is, momentum, energy and species balances, accounted for variations of pressure, temperature, and composition in the axial direction for both the process- and furnace-side gases. For the reforming tubes, the rigorous kinetic model (Xu et al., 1989) was used, while the furnace-gas combustion reactions were modeled using a parabolic heat-release profile. Convection and radiation were modeled on the furnace-side, where view factors were estimated using Monte-Carlo simulations and gray-gas assumptions. The heat transfer on the process side was modeled by convection only, assuming a pseudo steady-state between the process gas and catalyst. Details of the models are Section 2.6.

The PSA beds represent a cyclic process, with beds switched from adsorption to regeneration on the order of every minute. This type of separation scheme induces oscillatory behavior throughout the SMR process. As the flow rates, compositions, and pressures fluctuate in effluent streams from the PSA beds, variables throughout the entire plant fluctuate as well. In processes with such cyclical units, buffer tanks are often used to dampen fluctuations. However, typical buffer-tank sizes (comparable to SMR-unit sizes) reduce the amplitude of these fluctuations by on the order of 50%. Herein, the SMR process test-bed involves four PSA beds, which operate in a 4-mode scheme, with each bed undergoing adsorption, depressurization, desorption, and repressurization steps. PDAE's are used to model the momentum, energy, and species balances, dynamically tracking pressure, temperature, and composition in the axial direction. Langmuir isotherms are used to model adsorption kinetics. Details of the models are in Section 2.7.

In the full safety process model, the SMR-unit and PSA-bed models are used in conjunction with dynamic models for the water-gas shift reactor, water extractor, surge tank, heat exchanger, and steam drum. Furthermore, the controls used with the dynamic process model are consistent with those used in the real process. The full process is modeled using the software package, gPROMS. A challenging aspect of the full process model involves convergence of the PSA-offgas recycle loop.

To my knowledge, no published SMR model exists with this level of detail. In particular, this process model combines SMR and PSA-bed units within a plant-wide scheme with PSA-offgas recycle. The results computed by gPROMS are consistent with the process data from the industrial plant. This plant-wide model is extremely useful for building leading indicators and prior distributions of alarm and safety interlock system failure probabilities.

With a dynamic process model, process engineers can simulate special cause events and track variable trajectories. Consider an unmeasured 10 percent decrease in the Btu-rating, due to a composition change of the natural gas feed (40), in Figure 2.4. Note that the makeup stream (815) on the furnace side is relatively small and is not changed in the simulation. Initially, because the process stream contains less carbon, less H_2 is produced. Because these reactions are endothermic, less heat from the furnace is consumed by the reactions and the furnace temperature rises, as shown in Figure 2.5. Also, the process-side temperature increases. Eventually, the low-carbon PSA-offgas reaches the SMR furnace. With less methane for combustion, the furnace temperature decreases, as does the temperature of the process gas. This effect is shown in Figure 2.5. Note that the temperatures oscillate due the natural gas oscillation in stream 800 from the PSA-offgas surge tank – due to the cyclic nature of the PSA process.

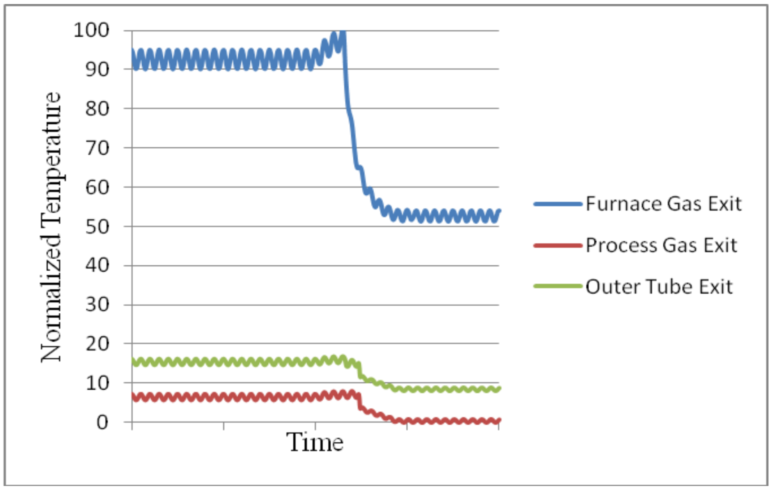


Figure 2.5. SMR effluent temperatures for a 10% decrease in the Btu content of the natural gas feed.

2.5.1. Reformer Model

The SMR herein is a top-fired unit consisting of approximately 400 process tubes. Steam and CH_4 are fed on the process side (Stream 90 in Figure 2.4). In the tubes, H_2 is produced via a set of endothermic reactions. On the furnace side, a fuel source (Stream 817) is combusted to provide heat for the process side. A schematic of the SMR unit is shown in Figure 2.6 (Latham et al., 2011).

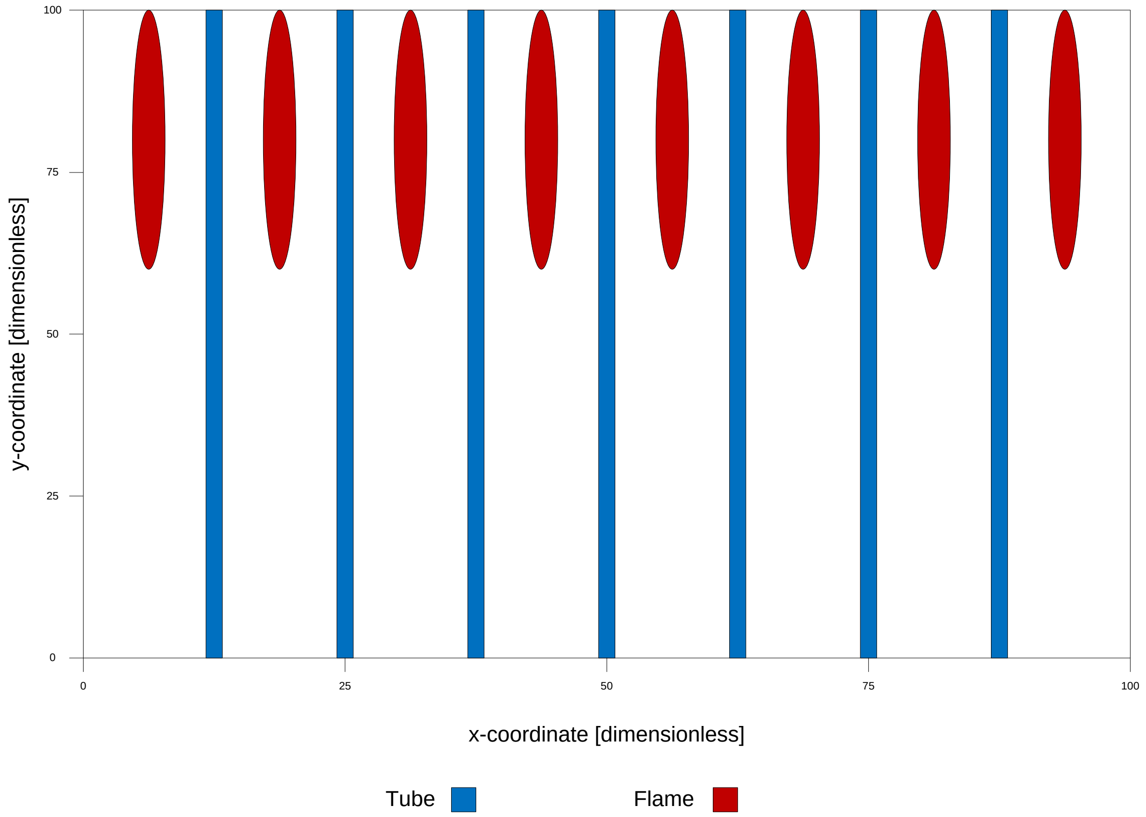


Figure 2.6. Front-view schematic of SMR.

The model proposed by (Latham et al., 2011), which describes the SMR in the steady-state, was converted to a dynamic model. Also, for the furnace-side, radiation view factors replaced the software RADEX used by Latham. In this work, the SMR is modeled as four units: the process gas, the process tubes, the furnace gas, and the reformer brick. The process gas and the furnace gas are modeled as networks of PDAE's, having derivatives with respect to time (t) and axial direction (z). Each model is discretized in the axial direction with central-difference approximations. The resulting ODEs are integrated in time, with the discretized equations solved for the state variables at the end of each time step. The process tubes and reformer brick are modeled as networks of PDAE's, having derivatives with respect to time (t), axial direction (z), and lateral direction (y). These are also discretized in the spatial coordinates with central difference approximations.

In the process and furnace gas models, the state variables are the molar flow rates of each species i ($\dot{n}_i$) and temperature (T). For the process gas, the mass balances for species i are:

$$\frac{d}{dt}(C_i) = -\frac{1}{A_c}\frac{d}{dz}(\dot{n}_i) + \sum\nolimits_{\mathrm{rxns}\ j} \mathrm{v}_{i,j} r_j \qquad i = \{\mathrm{CH_4, H_2O, H_2, CO, CO_2}\} \qquad (2.8)$$

and the energy balance is:

$$\frac{d}{dt}\left(C_\mathrm{v} T \sum\nolimits_i C_\mathrm{i}\right) = -\frac{1}{A_c}\frac{d}{dz}\left(C_\mathrm{p} T \sum\nolimits_i \dot{n}_i\right) + \frac{2h}{r}(T_{tube} - T) + \sum\nolimits_{\mathrm{rxns}\ j} r_j \Delta H_j \qquad (2.9)$$

where C_i is the concentration of species i, A_C is the cross-sectional area of a process tube, $\nu_{i,j}$ is the stoichiometric coefficient of species i in reaction j, C_v is the molar heat capacity at constant volume, C_p is the molar heat capacity at constant pressure, h is the heat-transfer coefficient, r is the inner tube radius, T_tube is the tube wall temperature, and ΔH_j is the enthalpy of reaction j. The heat capacities are functions of C_i, which are functions of T and $\dot{n}_i$:

$$C_i \ = f_1(\dot{n}_{CH_4}, \dots, \dot{n}_{CO_2}, T), \quad i = \{\mathrm{CH_4, H_2O, H_2, CO, CO_2}\} \qquad (2.10)$$

$$C_\mathrm{v} = f_2(\dot{n}_{CH_4}, \dots, \dot{n}_{CO_2}, T) \tag{2.11}$$

$$C_\mathrm{p} = f_3(\dot{n}_{CH_4}, \dots, \dot{n}_{CO_2}, T) \tag{2.12}$$

These functions are used in Eqs. (2.8) and (2.9).

The reaction rates are calculated using the kinetic model of (Xu et al., 1989), which involves three reforming reactions:

$$\mathrm{H_2O + CO \rightarrow H_2 + \ CO_2} \tag{R2.1}$$

$$\mathrm{H_2O + CH_4 \rightarrow H_2 + \ CO} \tag{R2.2}$$

$$\mathrm{H_2O + CH_4 \rightarrow H_2 + \ CO_2} \tag{R2.3}$$

The reaction rates are:

$$r_1 = \frac{\rho_{cat}\eta k_1}{P_\mathrm{H_2}} \frac{\left(P_\mathrm{CO}P_\mathrm{H_2O} - \frac{P_\mathrm{H_2}P_\mathrm{CO_2}}{K_1}\right)}{(DEN)^2} \tag{2.13}$$

$$r_2 = \frac{\rho_{cat}\eta k_2}{P_\mathrm{H_2}^{2.5}} \frac{\left(P_\mathrm{CH_4}P_\mathrm{H_2O} - \frac{P_\mathrm{H_2}^{3}P_\mathrm{CO}}{K_2}\right)}{(DEN)^2} \tag{2.14}$$

$$r_3 = \frac{\rho_{cat}\eta k_3}{P_\mathrm{H_2}^{3.5}} \frac{\left(P_\mathrm{CH_4}P_\mathrm{H_2O}^{2} - \frac{P_\mathrm{H_2}^{4}P_\mathrm{CO_2}}{K_3}\right)}{(DEN)^2} \tag{2.15}$$

where:

$$k_j = k_{j,0} \exp\left[-\frac{E_j}{R}\left(\frac{1}{T} - \frac{1}{T_{\mathrm{ref}}}\right)\right] \quad , \qquad j = \{1,2,3\}$$

$$K_j = \exp\left[-\frac{\mathit{\Delta} G_j^T}{RT}\right] \quad , \quad j = \{1,2,3\}$$

$$DEN = 1 + K_{\mathrm{CO}}P_{\mathrm{CO}} + \; K_{\mathrm{H_2}}P_{\mathrm{H_2}} + K_{\mathrm{CH_4}}P_{\mathrm{CH_4}} + \frac{K_{\mathrm{H_2O}}P_{\mathrm{H_2O}}}{P_{\mathrm{H_2}}}$$

$$K_i = A_i \exp\left[-\frac{\mathit{\Delta} H_{(\mathrm{adsb},i)}}{RT}\right] \quad , \quad i = \{\mathrm{CO}, \mathrm{H_2}, \mathrm{CH_4}, \mathrm{H_2O}\}$$

and $k_{j,0}$, T_{ref} , and A_i are constants and η is a parameter that describes the diffusion-limitation of the reaction.

Note that R2 and R3 are highly endothermic. The Ergun equation is used to solve for the pressure:

$$\frac{d}{dz}(P) = f_{\mathrm{p}} \frac{\rho v_{\mathrm{s}}^2}{D_{\mathrm{p}}} \times \frac{1-\varepsilon}{\varepsilon^3} \tag{2.16}$$

where:

$$f_{\mathrm{p}} = \frac{150(1-\varepsilon)\mu}{D_{\mathrm{p}} v_{\mathrm{s}} \rho} + 1.75$$

and ρ is the gas density, v_{s} is the superficial gas velocity, D_{p} is the catalyst particle diameter, ε is the void fraction, and μ is the gas viscosity.

The furnace gas model is similar with one major exception: radiation heat transfer is included in the energy balance. The radiation heat loss rate emitted by a volume of gas is characterized as:

$$q_r = 4\kappa V \sigma T^4 \tag{2.17}$$

where V is the gas volume, κ is the gas emissivity (a function of composition, temperature, and pressure), σ is the Stefan-Boltzmann constant, and T is the absolute temperature. When incorporating radiation into the model, it is critical to remember that radiation can travel from any one section of the furnace unit to another without being absorbed first. Said differently, each discretized section of furnace gas undergoes radiation heat transfer with each other discretized section of furnace gas, as well as each discretized section of process tube and reformer brick that is exposed to the furnace gas. It should be noted that the radiation heat-loss rate emitted by the process tubes and the reformer brick are, respectively:

$$q_{\text{tube}} = \epsilon_{\text{tube}} A_{\text{tube}} \sigma T_{\text{tube}}^4 \tag{2.18}$$

$$q = \epsilon_{\text{wall}} A_{\text{wall}} \sigma T_{\text{wall}}^4 \tag{2.19}$$

where A_{tube} is the tube area for heat transfer, A_{wall} is the wall area for heat transfer, and $\varepsilon_{\text{tube}}$ and $\varepsilon_{\text{wall}}$ are the emissivities of the tube and wall. Thus, the energy balance for the discretized furnace gas volume p is:

$$
\begin{aligned}
&\frac{d}{dt}\left(C_{\mathrm{v}}T\sum_{i}C_i\right)\\
&\qquad = -\frac{d}{dz}\left(C_{\mathrm{p}}T\sum_{i}\dot{n}_i\right)+\frac{A_{\mathrm{tube}}h}{V_{\mathrm{fur}}}(T_{\mathrm{tube}}-T)+\frac{A_{\mathrm{wall}}h}{V_{\mathrm{fur}}}(T_{\mathrm{wall}}-T)\\
&\qquad + \sum_{\mathrm{rxns}\ j}\mathrm{r}_j\Delta\mathrm{H}_j - 4\kappa\sigma T^4 + \sum_{\mathrm{fur\ volumes}\ q}P_{p,q}4\kappa\sigma T_q^4\\
&\qquad + \sum_{\mathrm{tube\ surfaces}\ q}P_{p,q}\epsilon_{\mathrm{tube}}\sigma T_q^4\times\frac{A_{\mathrm{tube}}}{V_{\mathrm{fur}}}\\
&\qquad + \sum_{\mathrm{fur\ volumes}\ q}P_{p,q}\epsilon_{\mathrm{wall}}\sigma T_q^4\times\frac{A_{\mathrm{wall}}}{V_{\mathrm{fur}}} \qquad (2.20)
\end{aligned}
$$

The coefficients $P_{p,q}$ represent the probability that a ray of radiation leaving the radiation zone q is absorbed by zone p. These are calculated using Monte-Carlo techniques. At each radiation zone, a large number of points are randomly chosen, each with a random direction (representing a ray of radiation emitted). Each ray's absorption is tracked, permitting the estimation of $P_{p,q}$ for radiation heat transfer between zones p and q. Also, the probability of a furnace zone absorbing a ray of radiation depends upon the zone's κ value, with high κ values characterizing a 'gray' gas, which readily absorbs radiation, and low κ values characterizing a 'clear' gas which lets radiation pass through. Therefore, the temperature, composition, and pressure of the furnace gas affect each $P_{p,q}$ within the furnace (Hottel et al., 1967). Because the Monte-Carlo integration for determining each $P_{p,q}$ is difficult to install in gPROMS, the Monte-Carlo integration was carried out offline for a grid of reasonable temperature, pressure, and composition values. Within gPROMS, values of $P_{p,q}$ are interpolated from this grid.

Figure 2.7 shows solutions for the temperature profiles on the process- and furnace-sides of the SMR for typical operation. On the furnace side, the temperature quickly rises in the first third of the unit, where the combustion reaction takes place. Over the next two thirds of the unit, the temperature on the furnace side decreases as heat is transferred to the process side. On the process side, the temperature increases throughout, however its slope is greatest where the furnace gas is hottest. Species flow rates on the process side are shown in Figure 2.8, with the bulk of H_2 produced in the top section, and more than half of the CH_4 consumed. Also note that the reformer is sufficiently long, and consequently, little reaction takes place near its bottom.

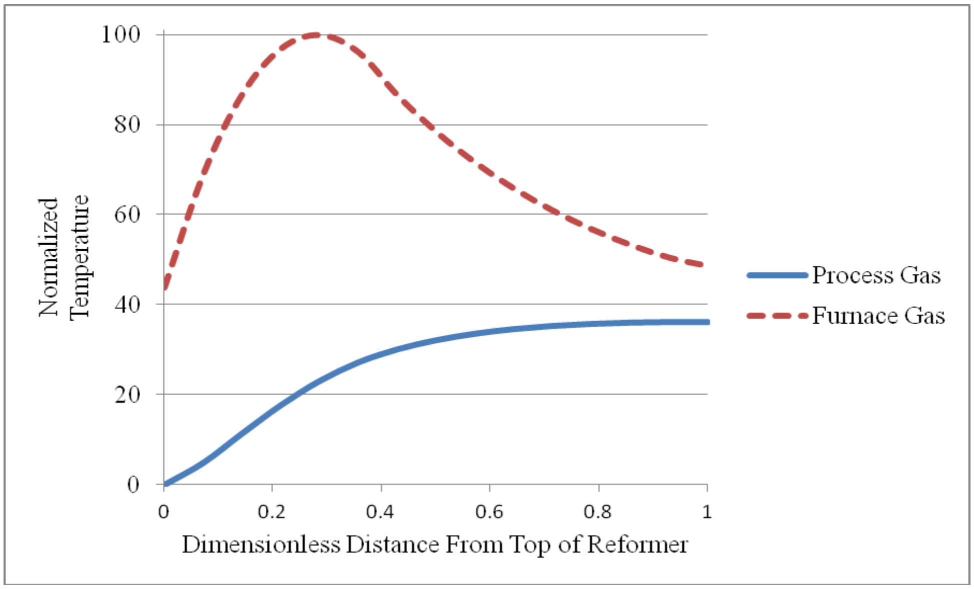


Figure 2.7. Temperature profile in the SMR.

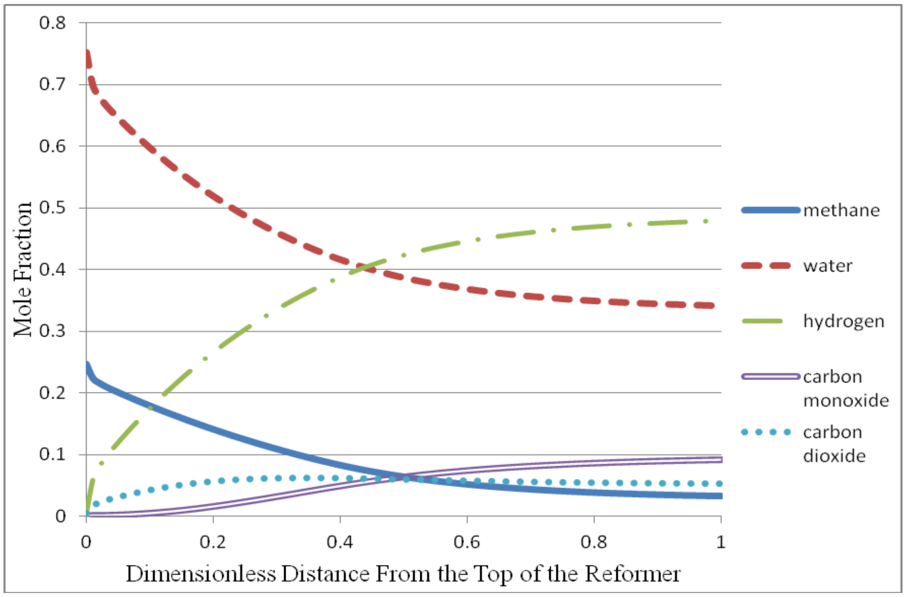


Figure 2.8. Mole fraction profile on the process-side of the SMR.

2.5.2. Pressure Swing Adsorption (PSA) Model

The pressure-swing adsorption model consists of four beds, each of which is described by a set of PDE's (derivatives are taken with respect to time and the axial direction) and associated boundary conditions. The PSA cycle consists of four steps: adsorption, depressurization, desorption, pressurization. At any given time, one of the beds is in each step, cycling to the next step every minute. To model the PSA process, just a single 4-bed unit was used, with each bed triple its size in the plant. These beds are sufficiently large to adsorb all of the carbon-compounds during the 1-minute adsorption step. Also, a larger surge tank is used to dampen the bed swings, because three 4-bed groups cycle out of phase, creating destructive interference among the limit cycles.

Using the method of lines, the PDEs were discretized in the axial direction with central-difference approximations (and forward/backward differences at the boundaries) to represent the derivatives with respect to distance. Then, backward-difference formulae were used to integrate the resulting ODEs in time. The following state variables were used in modeling the beds: temperature, pressure, gas density, gas velocity, mole fraction of each species, loading of each species on the adsorbent, and the equilibrium loading of each species on the adsorbent. The Langmuir Isotherm was used to calculate the equilibrium loading of each species.

Below is the schedule used for a 4-bed PSA cycle, similar to that implemented by Agarwal et al. (Agarwal et al., 2008), and Khajuria and Pistokopolous (Khajuria et al., 2011) (for a 2-bed cycle) and used herein. The four beds (A, B, C, and D), shown schematically in Figure 2.9a, cycle through four steps, 1-4, during modes 1-4. Initially, just prior to the first mode, bed A is at high pressure (HP) and unoccupied (U) by species to be adsorbed; bed B is also at HP, but occupied (O); bed C is at low pressure (LP) and O; and bed D is at LP and U. The valves are open or closed as shown. During the first mode, bed A implements step 1; bed B, step 2; bed C, step 3; and bed D, step 4. In mode 2, A moves to step 2, B to step 3, C to step 4, and D to step 1. Similar moves are made for the third and fourth modes. For the PSA in the SMR process, the beds are occupied with CH_4, CO, and CO_2, and unoccupied with H_2, which is used to desorb the adsorbents – HP > 20atm, and LP = 1.5 atm – each mode (and step) is of duration, 1 min.

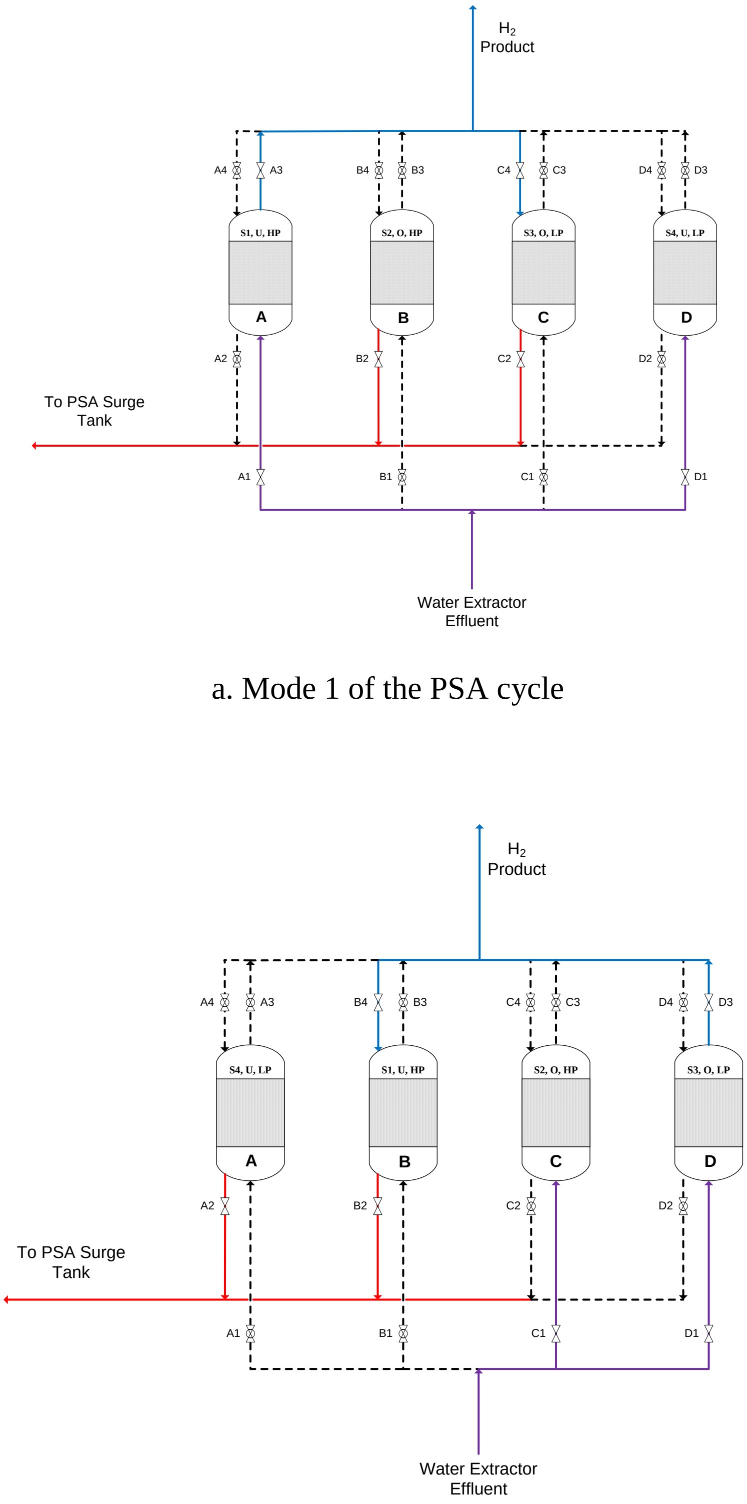
Next, the four steps are described, focusing on bed A. Note that Figures 2.9a-d, show the states of each bed and the valve positions at the start of each mode. In this analysis, it is assumed that the valves are adjusted instantaneously.

- Step 1: Adsorption in bed A (see Figure 2.9a at outset of mode 1)
  - Step 1 begins with bed A at HP and U
  - The water extractor effluent (210) is fed to the base of bed A (valve A1 open)
  - A high-purity stream of H_2 leaves the bed (valve A3 open), as the adsorbent accumulates CH_4, CO, and CO_2 in time, with most of the H_2 product sent in stream 220 to the H_2 Product tank, and the remainder sent to bed C (valve C4 open).

- Step 2: Depressurization of bed A (see Figure 2.9b at outset of mode 2)
  - Step 2 begins with the valve A1 and A3 closed.
  - Gas exits the pressurized bed through the base of bed A (valve A2 open), and in time the pressure in the bed equilibrates to a lower pressure (pressure of the surge tank)

- Step 3: Desorption in bed A (see Figure 2.9c at outset of mode 3)
  - Step 3 begins with valve A4 open, with H_2 product from bed C entering the top of bed A
  - Reminder: bed A is at LP and O
  - In time, the H_2 product adsorbs on the adsorbent and the CH_4, CO, and CO_2 is released into the PSA-offgas

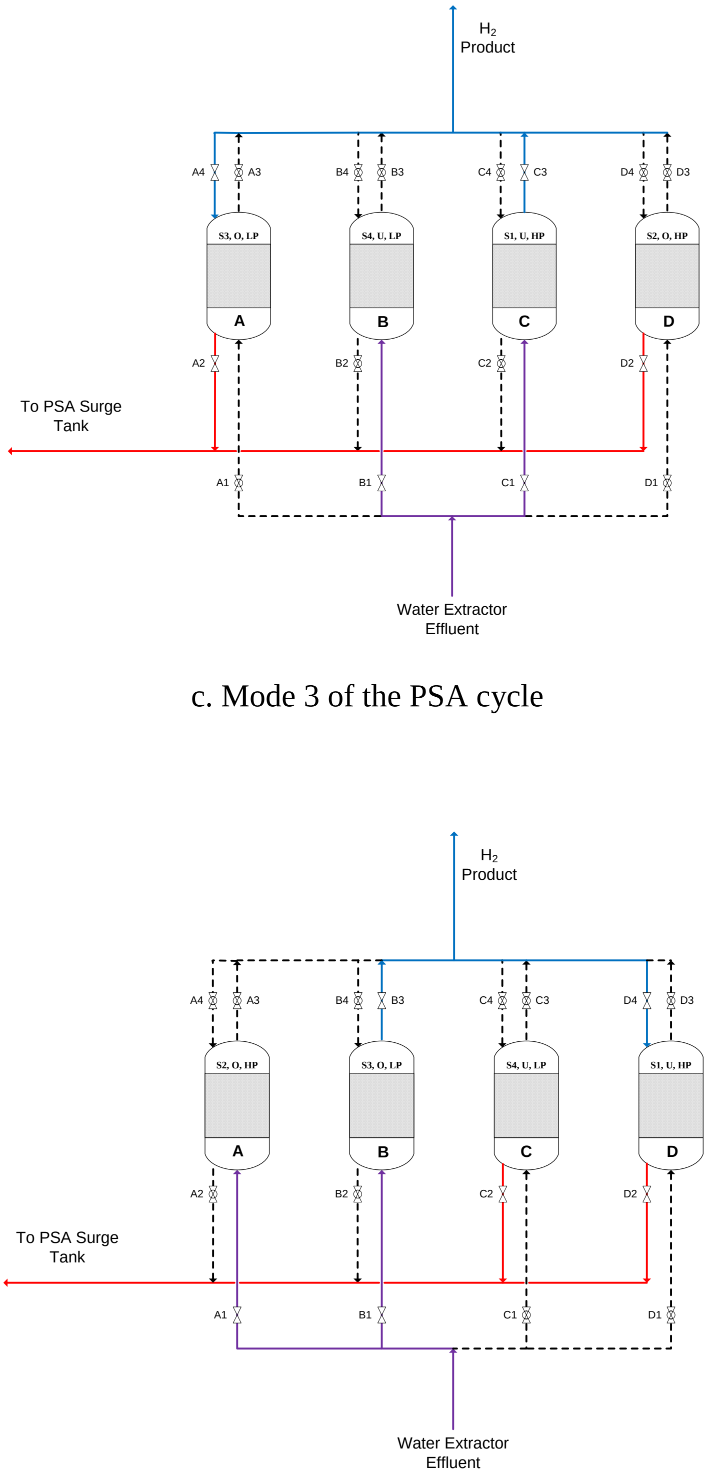
- Step 4: Pressurization of bed A (see Figure 2.9d at outset of mode 4)
  - Step 4 begins with valve A2 and A4 closed
  - The water extractor effluent (210) is fed to the base of bed A (valve A1 open)
  - In time, the pressure of bed A equilibrates to the pressure of stream 210



a. Mode 1 of the PSA cycle

b. Mode 2 of the PSA cycle

Figure 2.9. Schematic of PSA process.



c. Mode 3 of the PSA cycle

d. Mode 4 of the PSA cycle

Figure 2.9. Schematic of PSA process (Cont'd.)

To evaluate process safety using dynamic risk analysis, it's important to monitor the breakthrough of CH_4, CO, and CO_2 in the product H_2 stream. Previous models did not consider simulating a carbon breakthrough into the H_2 product stream, since they use a semi-infinite boundary condition with respect to molar composition at the end of the bed (Agarwal et al., 2008; Khajuria et al., 2011). The model developed herein uses the same boundary condition, but at the end of an elongated bed. In this way, the semi-infinite B.C. (zero derivative) is maintained at the end of the elongated bed, but a methane or CO/CO_2 breakthrough at the end of the bed can be observed. Shown in Figure 2.10 is the mole fraction of H_2 in the PSA bed during Step 1 of the PSA cycle. The mole fraction of H_2 along the bed is shown minute-by-minute. In time, the purity of the H_2 product stream, taken at the exit of the 1 meter long bed, drops below an acceptable level (set point). Clearly, this model is capable of simulating a carbon breakthrough into the H_2 product stream.

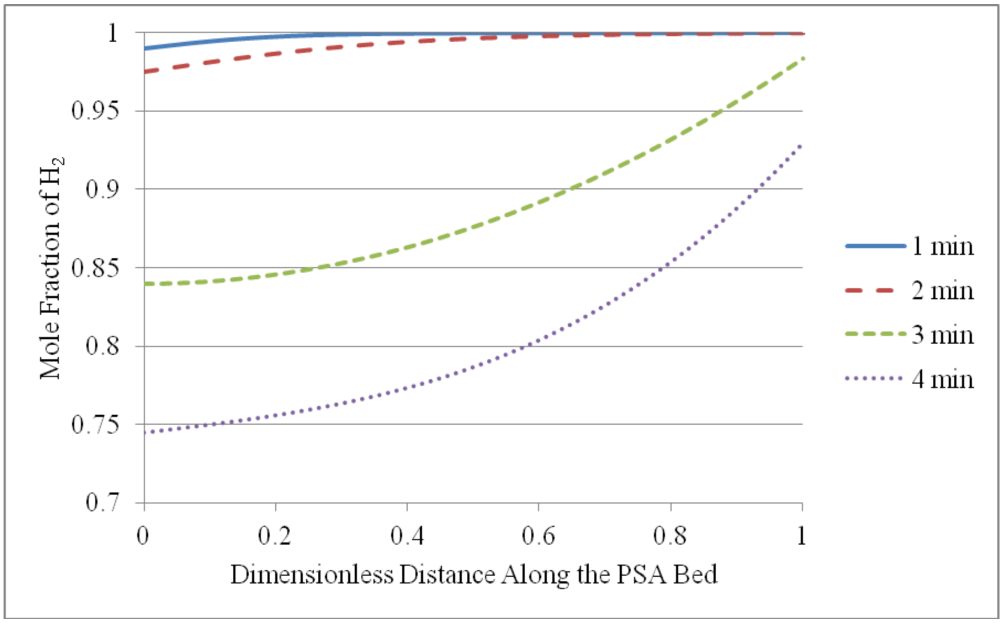


Figure 2.10. Simulated mole fraction of H_2 along the PSA bed during Step 1.

The most important aspect of the PSA modeling as it relates to dynamic risk analysis of the SMR process is the oscillations in the PSA-offgas. PSA-offgas is fed to the surge tank from the PSA bed undergoing Step 3 (desorption.) When a PSA bed switches to Step 3, its concentration of carbon compounds is the highest. This is because the adsorbent near the bottom of the bed, z = 0, contains the most $CH_4/CO_2/CO$. In Step 3, the carbon compounds are continuously desorbed, and consequently, the concentration of H_2 in the PSA-offgas increases. Because the PSA-offgas provides fuel for the furnace side of the SMR, the oscillations in its Btu-rating are important. Most significant, for the 4-bed model, are the oscillations in the effluent for the surge drum, which are shown in Figure 2.11. This is consistent with documentation for the industrial process studied.

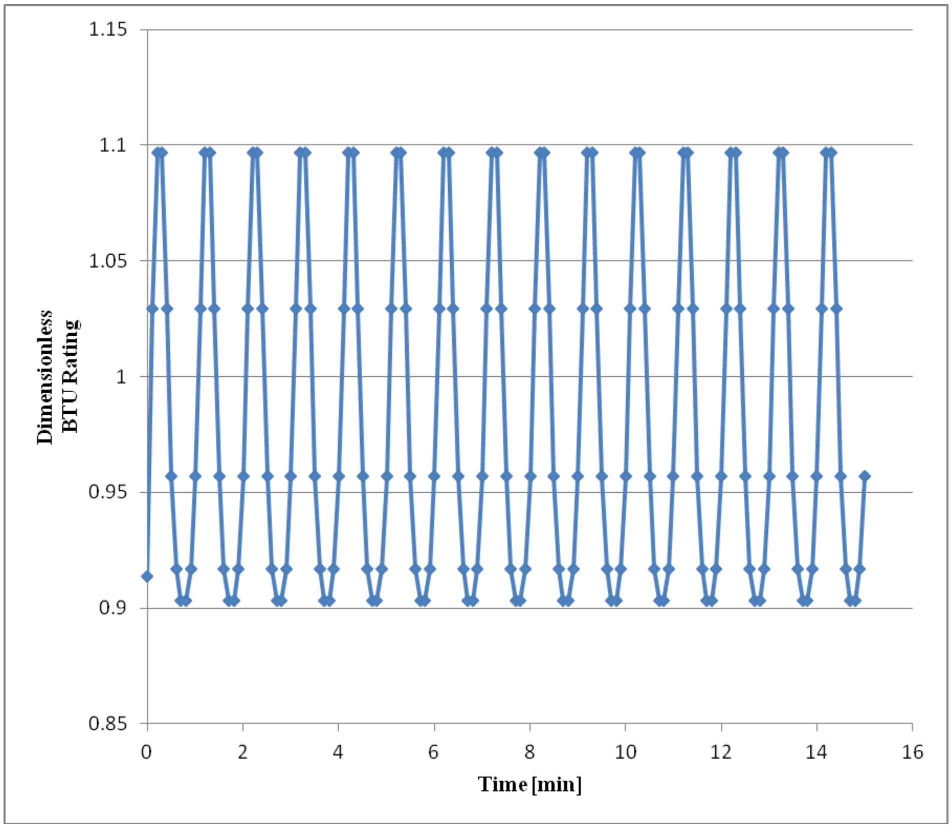


Figure 2.11. Simulated PSA-offgas Btu-rating

2.6. SMR Informed Prior Distributions

For the SMR process in Figure 2.4, a loss in steam pressure to the reformer-side (stream 560), was simulated with the responses of the safety systems monitored. For small disturbances, the process control system handled the effect of steam pressure decreases. There is a flow controller on the steam line, whose set point is generated using a linear equation involving the flow of natural gas into the SMR process-side, seeking to achieve a constant steam-to-carbon ratio in the process tubes. This control system normally arrests typical fluctuations in steam pressure and flow rate, but for large steam-pressure decreases, feedback control alone is insufficient. In this case, the control valve is wide-open, with a flow rate insufficient to accompany natural gas fed to the SMR-unit. For this reason, an investigation was undertaken to assess the effectiveness of the alarm systems associated with the SMR steam line. When the steam flow rate is below its L-alarm threshold, the steam-to-carbon ratio drops, accompanied by an increase in the process-side temperature, and potential tube failure. Because of these operating limits, an interlock was placed at the HH-alarm threshold with a time delay. This time delay, of several seconds, reflects that the temperature threshold may be exceeded in this case for a short period of time and permits the operator to respond rapidly in an attempt to bring the furnace temperature below its HH-alarm threshold.

Three operator responses to the alarm are simulated: (1) the valve on the steam line is opened, (2) the valve on the makeup fuel line is pinched, and (3) the dampers associated with air flow in the furnace are opened (effectively increasing air flow rate). When the operator is able to bring the furnace temperature below the HH-threshold before the interlock delay times out, an automatic shutdown is avoided. If, however, the operator is unable to do so, the interlock is activated and a plant shutdown occurs. The simulated abnormal event leads to either a success of SS_2 (interlock is avoided), or a failure of SS_2 (interlock is activated). It is desirable to have a reliable estimate of x_2, the probability that the operator is not successful, despite only a few activations of this HH-alarm during the recorded history over several years.

Herein, a pressure decrease in the steam line to the process-side of the SMR-unit was simulated. The magnitude of the pressure decrease was a random variable, sampled from a normal distribution centered at 50% of stream pressure. The response time of the operator was taken as a random variable, sampled from a uniform distribution ranging from 0 to 15 seconds. The operators three responses were all incorporated into the simulation as step-changes in valve settings. One thousand simulations were run, and the effectiveness of the operator's response in each simulation was tracked. In some simulations, the operator successfully reduced the furnace temperature below the interlock threshold in the allotted time before the automatic shut-down. In others, the operator failed and the plant was shut-down. In Figure 2.12, a temperature trajectories for events resulting in an interlock activation and in a near-miss are shown. In the scenario where SS_2 succeeds, the temperature is brought below the interlock threshold within the interlock delay time. Note that the action of the control system was observed early in the trajectory, but it was insufficient to avoid the abnormal event and eventual plant shut-down. The average number of safety-system failures was recorded for the simulations, as well as the failure variance, which were used to generate a beta-distribution to describe the failure probabilities. The beta-distribution, which has just two parameters, was created easily and is supported only in the range [0,1], which bounds the failure probability.

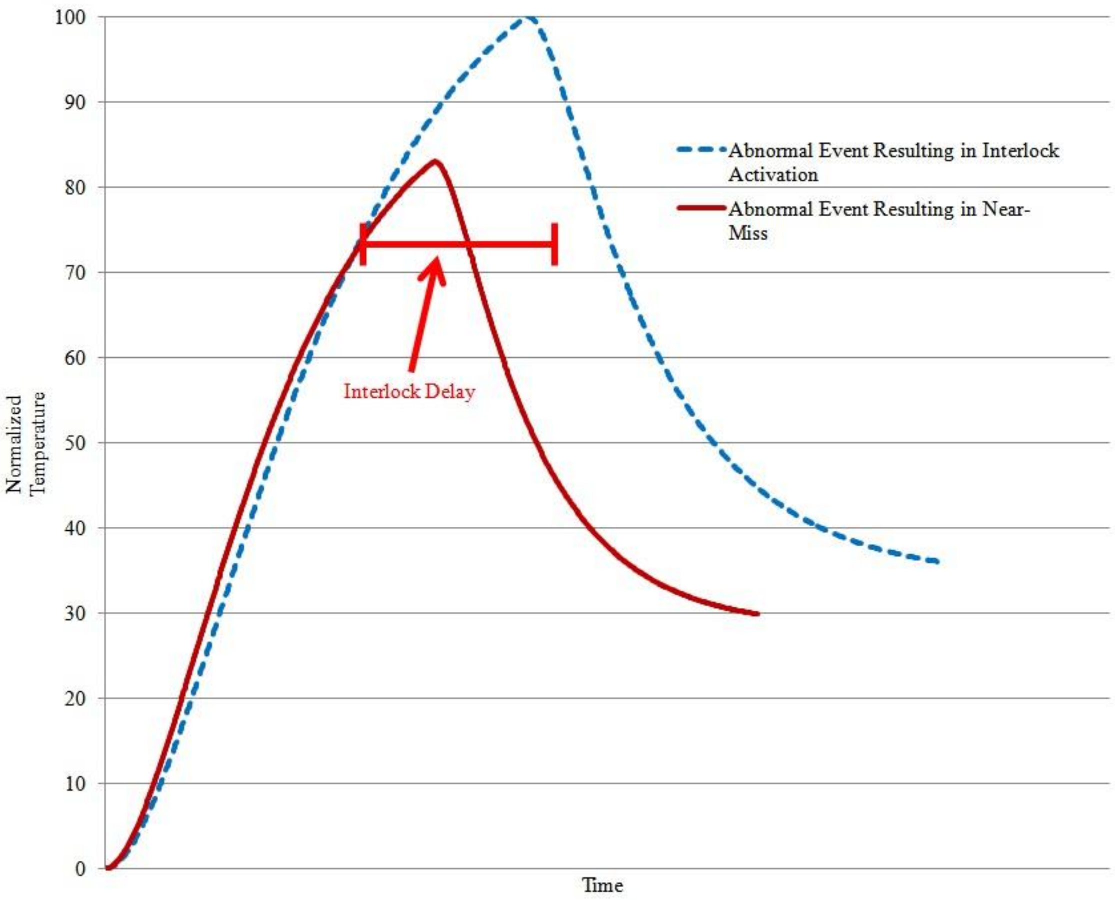


Figure 2.12. Furnace outlet temperature for a decrease in steam pressure.

This informed prior distribution was built using dynamic simulations with first-principle models. Even with no data available to update the distribution, process engineers and plant operators can make improved risk predictions (Levenson et al., 2014). The alarm data are used to build a likelihood distribution, in this case a binomial likelihood distribution of a few trials, all of which are successes. In Figure 2.13, the prior and posterior distributions are shown. The informed posterior is shifted to the left of the prior distribution by the 0 percent failure rate observed in the data. Unlike the commonly used uninformed prior distributions, its posterior distribution has a similar shape to its informed prior distribution. The posterior distribution generated using the informed prior distribution can alert process engineers that a significant decrease in steam pressure has the high probability (>20%) of causing a plant shutdown. This may lead operators to pay special attention to the steam pressure measurements, and may lead process engineers to install a more robust controller on the steam line.

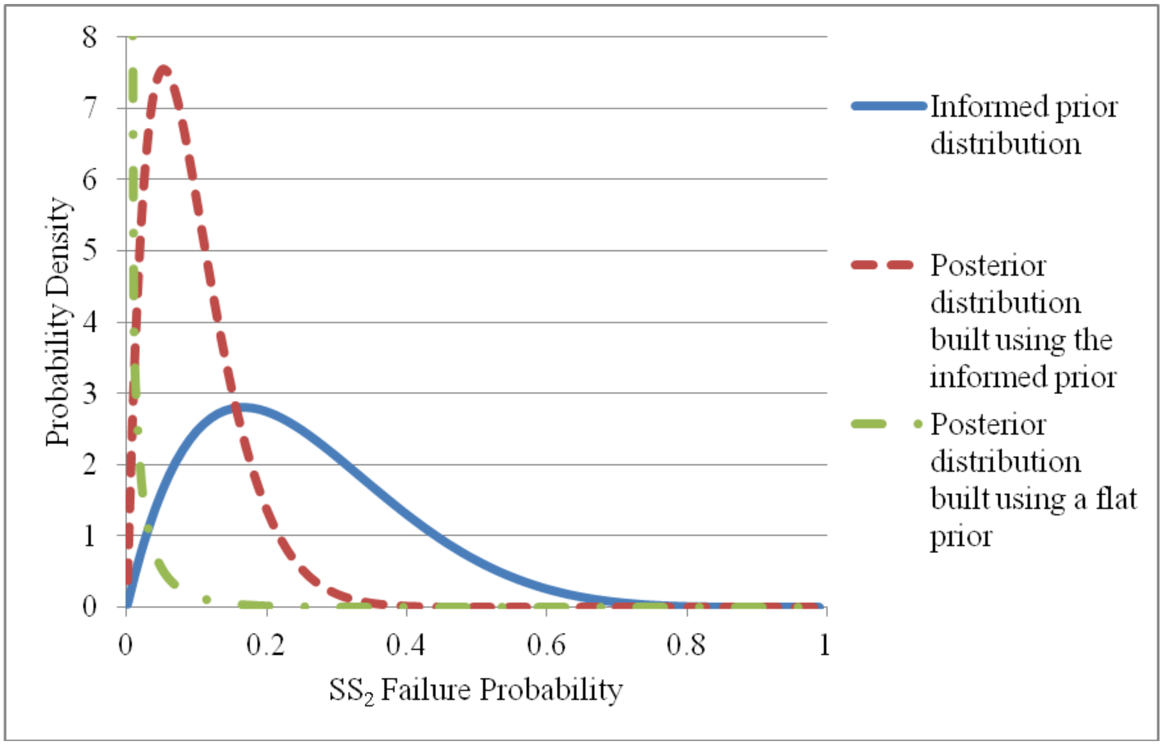


Figure 2.13. Prior and posterior distributions generated by dynamic simulations

2.7. Conclusions

A method, involving repeated dynamic process simulations, for constructing informed prior distributions was presented. The method was used in estimating the failure probabilities of alarm and safety interlock systems that are rarely called into action in chemical processes. The method requires combining a dynamic, first-principles, process model with the control, and alarm and safety interlock systems. Its application was demonstrated for offline dynamic risk analysis of a steam-methane reformer (SMR) process. The high probability of a plant shutdown calculated by the method can alert operators to pay special attention to the steam pressure measurements, and may lead process engineers to improve the controller on the steam line, avoiding or reducing the cost of plant shutdowns. Key aspects of the reformer/furnace and PSA models used to demonstrate the proposed methodology were presented. The modeling of probable operator responses with respect to operator skill, shift (day or night shift), severity of alarm (H/L or HH/LL), and the difficulty of diagnosing the special cause, among other factors, is considered in Chapter 3.

2.8. NOMENCLATURE

A	area
A_m	special-cause magnitude m
D	observed alarm data
i	failure counter; number of failures for a sampled special-cause magnitude
j_m	observed failure probability for a sampled special-cause magnitude, m
m	special-cause magnitude counter, m = 1, ..., M
M	number of sampled special-cause magnitudes
n	sample operator variables counter, n = 1, ..., N
N	number of sampled operator variables per special-cause sample
p	number of possible operator response orders (sequences)
SS_i	safety system i
x_i	failure probability of a safety system i
Greek	
α	first parameter of the Beta distribution
β	second parameter of the Beta distribution
ε	emissivity
ΔP	sampled pressure decrease
μ	average value of j_m
μ_{SC}	average of sampled special-cause magnitude
σ^2	variance of j_m
σ^2_{SC}	variance of sampled special-cause magnitude
τ	operator response time
τ_{max}	maximum operator response time

Chapter 3

Improved Predictions of Alarm and Safety System Performance Through Process And Operator Response-Time Modeling

3.1. Introduction

In the chemical process industries, there are many incentives to mitigate the frequency and consequences of incidents and accidents ("U.S. Chemical Safety and Hazard Investigation Board"). To evaluate the effectiveness of alarm and safety interlock systems reliably, the probabilities of alarm and safety interlock failures and the failure consequences must be quantified. Said differently, to compare two safety systems, quantitative estimates for their effectiveness in mitigating special causes are needed, where a special cause is a disturbance the basic process control system is unable to arrest. This work proposes a method of improving process models and introduces new probabilistic models that describe special-cause event occurrences and operator response-times, allowing for estimating alarm and safety-system failure probabilities more accurately. In industrial practice, methods such as HAZOP and HAZAN are commonly utilized to make safety and reliability estimates of processes on a unit-operation basis. The estimated failure probabilities (from statistical data and manufacturer estimates) of specific components are used to estimate failure probabilities in a process. But, more recently, dynamic risk analysis has been employed to update these probabilities as real-time data are measured. The focus herein is on events that are inherently rare, where failure predictions remain uncertain. A process model is utilized to generate simulation data that enhance sparse measured data. The effects of decisions involving process models, special-cause events, and operator behavior, on risk predictions are investigated.

An informed prior distribution was constructed in Chapter 2, shown in Figure 2.13, to estimate the distribution of failure probabilities (x_2) of the safety system associated with the HH-Temperature alarm (SS_2) in the SMR furnace. Special-cause events were repeatedly simulated involving a substantial decrease of steam pressure at the inlet of the SMR reactor – sufficient to cause the furnace temperature to rise out of its green-belt zone through its yellow-belt zone, and into its red-belt zone. It was shown that without operator actions, the short interlock time delay, Δt_{int}, would elapse and the process would undergo an automatic shutdown. When the operators responded sufficiently quickly to the special-cause event, simulations showed the process often returning to its green-belt zone, with the interlock shutdown avoided. The distribution of the simulation results estimates that SS_2 fails at about a 20% rate during these dramatic pressure decreases; i.e., the process is estimated to undergo an automatic shutdown in 20% of these rare cases - this is not a projection of the process accident rate or the interlock shutdown rate.

In process operation, SS_2 responded successfully to the rare HH-alarm activations, resulting in sparse alarm data. A binomial likelihood distribution was used to calculate the posterior distribution of x_2 given the alarm data D. Since all of the collected data were successful SS_2 actions, the posterior distribution is shifted to the left of the prior distribution. With just a handful of SS_2 activations, the resulting likelihood distribution has a very large variance, and the accuracy of the informed prior distribution to the posterior distribution is uncertain. If the process had undergone hundreds of SS_2 activations, the accuracy could be assessed. This poses a major challenge – how can the user be confident that the process and operator behavior models are sufficiently accurate given few data to assess the accuracy of the informed prior distribution? If the models used to generate the informed prior distribution do not predict the special-cause event well, the results obtained from the posterior distribution may be unreliable.

3.2. Development and Refinement of Models to Construct Informed Prior Distributions

Over years of process operation, SS_1 activations are infrequent, but still provide sufficient data for studying the propagation of special-cause events. The activation of SS_2 is rarer, occurring $\frac{1}{x_1}$ times less frequently – often resulting in very sparse data. While the data associated with SS_2 are insufficient alone to analyze the performance of the safety system, the similarities between the safety systems can be utilized. The activations of both safety systems originate from a control system failure, which, for example, can be due the large magnitude of the disturbance, the inability of the control system to handle the disturbance, and/or the occurrence of an electrical or mechanical failure. It is assumed that the same group of operators are involved. If highly skilled, they should arrest the special-causes at a high rate (Meel et al., 2007; Meel et al., 2008; Chang et al., 2007).

Clearly, the need for urgent responses of SS_2 are greater. Also, when operators take action (e.g., as furnace temperatures become elevated), the need to respond within the interlock delay times dominate their concerns and actions. This would normally stimulate a strong reaction to avoid automatic shutdown.

As the fundamental basis for the proposed method, a sufficiently-accurate first-principles process model is needed. Also, the automated safety system models should be sufficiently accurate. The second model, represented by $g_1(A_m)$ in Figure 2.3, is the distribution of special-cause event magnitudes to be simulated. The operator behavior model, $g_2(\tau)$, unlike automatic safety systems, must reflect human behaviors. Here, the speed and effectiveness of operator responses often depend on the state of the process, the number of competitive active alarms, distractions, personal health and conflicts, and the like.

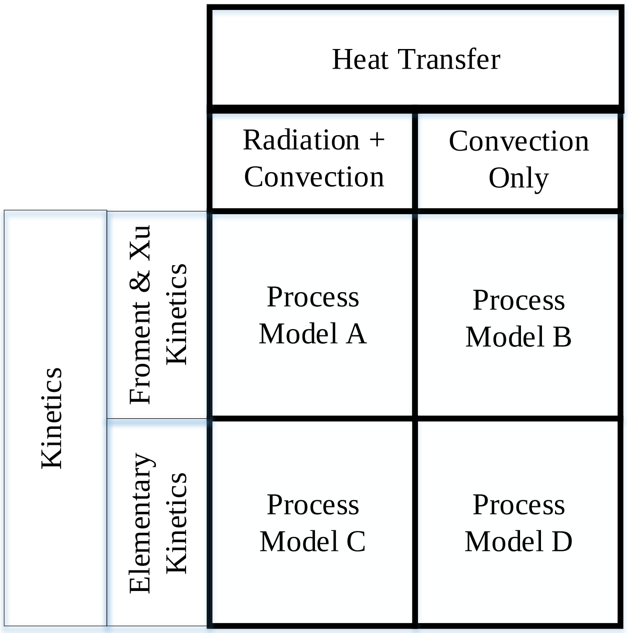
In the method introduced herein, first, SS_1 models are constructed and validated with plentiful data. After constructing these models, and validating them with measured SS_1 data, they are modified to handle SS_2 activations (recognizing that their rare occurrences do not allow for reliable model validation). In the next three sections, all three models are described with respect to SS_1. Their modifications to handle SS_2 activations are then described. Lastly, the failure probability estimates generated by using the SS_2 informed prior distributions are presented.

3.2.1. Dynamic Process Models

Because dynamic first-principles process models are widely used, approaches to model development are not considered here. Instead, this section focuses on model evaluation and improvement for constructing informed prior distributions.

Often, process engineers have developed dynamic process models for control scheme testing during the design phase. These are commonly used initially for carrying out dynamic risk analysis. However, process models used for process design and control are normally developed to track responses in their typical operating regimes (green-belt zones) – but may not respond to special-cause events with sufficient accuracy; i.e., their predictions far from set points may be poor for risk analysis. Consequently, dynamic process models should often be improved to construct informed prior distributions.

For the SMR process shown in Figure 2.4, four dynamic process models are constructed, as summarized in Figure 3.1. The first, Process Model A, is the same as the one described in Chapter 2. This model includes constitutive equations to model the endothermic reformer reactions, the furnace that provides their heat, the exothermic water-gas shift reaction, the separation of hydrogen product from offgas in adsorption beds, and the production of steam (for process heating or sale), as well as models of associated PID controllers.



		Heat Transfer	
		Radiation + Convection	Convection Only
Kinetics	Froment & Xu Kinetics	Process Model A	Process Model B
	Elementary Kinetics	Process Model C	Process Model D

Figure 3.1. Steam-methane reforming process models.

In Process Model A, to model the radiative heat transfer (~90% of the total heat transfer to the tubes), view factors are estimated from each surface or volume zone to each other zone, with the dynamics of radiative heat transfer modeled between all discretized zones (Hottel et al., 1967). The remaining convective heat transfer is simpler, because heat transfer only occurs between physically adjacent zones (Latham et al., 2011).

In Process Models B and D, convection heat transfer is modeled only, with radiative heat transfer accounted for by overstating the heat-transfer coefficients between the furnace gases and tube surfaces. Herein, to estimate the overstated heat-transfer coefficients, 50 steady-state windows were identified in the historical process data. Each window corresponds to a duration of operation, on the order of a day, where process variables are at steady state. Many different steady-state windows exist in the process data due to different demand rates of hydrogen and steam, different feed ratios of steam to natural gas, and natural aging of the catalyst (in the reformer as well as the water-gas shift reactor). The heat-transfer coefficient was estimated from the temperature and flow rate of the process and furnace gas inlet and outlet measurements.

In Process Models A and B, the reforming reaction kinetics proposed by (Xu et al., 1989), which have been shown to be quite accurate over a broad range of temperatures and reactant concentrations, are used. Note that, due to the presence of a complex denominator in the kinetic equations, the spatially-distributed SMR model can be difficult to converge. Accurate guess values for the concentration of the reactants and products along the axial direction of the reformer tubes must be available, or generated using homotopy-continuation techniques, to converge the steady-state model.

However, in Process Models C and D, elementary reaction kinetic equations are simpler to converge. The rate constants of the elementary reactions are estimated, similar to the convection heat-transfer coefficients. Using the data in the 50 steady-state windows, along with measured hydrogen product flow rates and offgas concentrations, the rate constants are estimated. .

Initially, the four process models are compared in the 50 steady-state windows. Beginning with the measured inlet temperatures and flow rates for each mode, predicted and measured outlet temperatures are compared for each model. The root-mean square outlet temperature differences are shown in Figure 3.2. For this steady-state evaluation, Process Model A provided the best agreement with the data, whereas Process Model D was least accurate.

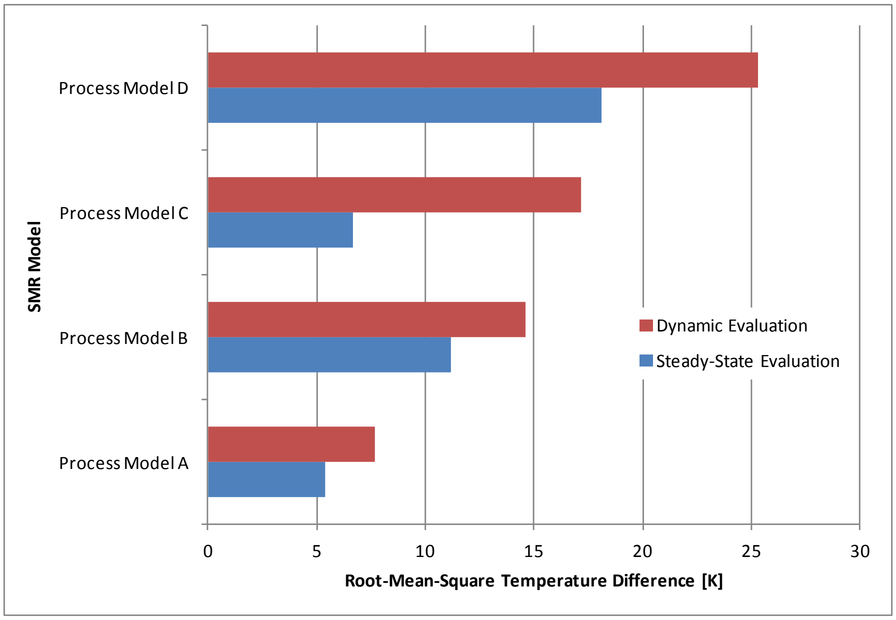


Figure 3.2. Process model goodness-of-fit using steady-state and dynamic evaluations.

However, because the models are used to estimate the responses to special-cause events, agreement with dynamics data is more important. Fifty dynamic windows were identified in the historical process data – periods of time where the process variables describing the operation of the steam-methane reforming reactor are transient. These windows are on the order of minutes to hours, and typically arise when hydrogen or steam demand rates change, feed ratios of steam to natural gas change, or operational changes occur in another process unit (such as a pair of pressure-swing adsorbent beds are taken offline). For each of the 50 dynamic windows, inlet temperature and flow rate trajectories are input to each model, with model-predicted outlet temperature trajectories compared to measured outlet temperature trajectories. Dynamic predictions are typically less accurate than the steady-state ones. Here, Process Model B outperforms

Process Model C, but when comparing just steady-state outlet temperature differences, Process Model C provides a closer fit to the data. Clearly, Process Model B should be selected, rather than Process Model C, when constructing informed prior distributions.

Next, the four process models are used to construct informed prior distributions for the failure of SS_1 – using the uniform distributions for special-cause magnitude and operator behavior used in Chapter 2. The 300 measured SS_1 failures/successes are then used to construct a low-variance binomial likelihood distribution (see Eq. (2.4)). The four informed prior distributions for the failure of SS_1 and the binomial distribution are shown in Figure 3.3. To compare the informed four prior distributions with this likelihood distribution, the $\xi_{i,\mathrm{m}}$ index is defined:

$$\xi_{i,m} = 1 - \frac{1}{2}\int_0^1 |f_m(x) - f_i(x)|dx \tag{3.1}$$

where i represents the model (i = A, B, C, D) and m represents the data-based likelihood distribution. This index ranges from [0, 1], where unity corresponds to perfect matching between the informed prior distribution of model i and the measured likelihood distribution m. As shown in Table 3.1, this index is consistent with the dynamic model accuracies in Figure 3.2, but low levels of agreement are obtained. Note that using more detailed operator response-time models, in the next subsection, the performance indices are improved significantly.

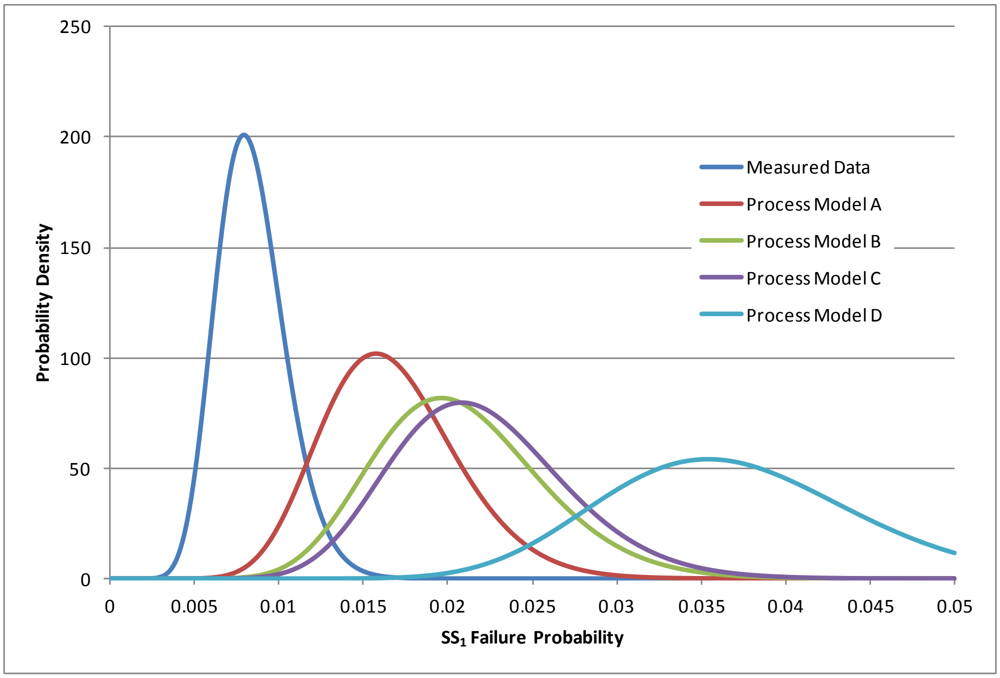


Figure 3.3. Informed prior distributions created using the four process models, as well as the binomial likelihood distribution created using the measured alarm data.

Table 3.1. Performance Index for Process Models A-D.

	Process Model A	Process Model B	Process Model C	Process Model D
$\xi_{i,\mathrm{m}}$	0.159	0.063	0.045	0.026

3.2.2. Special-Cause Event Occurrence Model

When constructing informed prior distributions to estimate the failure probabilities of safety systems that act infrequently, it is important to assess the special-cause events that can activate specific safety systems. Given that process units fail in many ways (e.g., as inlet stream compositions, temperatures, flow rates, and pressures vary; controllers experience measurement bias; valves malfunction; controller electronic mechanisms fail), special-cause modeling deserves attention. Clearly, for specific special-cause events, known to trigger safety systems, it is crucial to account for them when creating informed prior distributions. Some are known to have a high likelihood of occurrence over the lifetime of a plant, while others may be *un-observed* locally, having occurred at other plant sites or even related plants. For all potential special-cause events, a probability distribution should be constructed, even when likelihood data are unavailable.

When developing occurrence models to estimate safety system failure probabilities, special-cause events must be selected and their magnitudes must be investigated, as special-cause events are likely to have devastating consequences (e.g., propagation of runaway reactions, leading to explosions). To identify these events, HAZOP and LOPA analyses, especially, are particularly helpful. HAZOP is the industry standard for postulating all possible special-cause events.

The effect of a special-cause event (SCE) depends on its magnitude. SCE models (e.g., $g_1(A_m)$ in Figure 2.3) are needed to estimate failure probability distributions. SCE models having low expected values are most representative and rarely activate second-level alarms, while SCE models having high expected values represent extreme cases, allowing for the study of second-level alarms. Various steam-pressure-decrease magnitudes are shown in Figure 3.4 – each being a delta function centered at its corresponding point on the abscissa. The expected value of the SS_1 failure probability, j_{m}, is graphed accordingly.

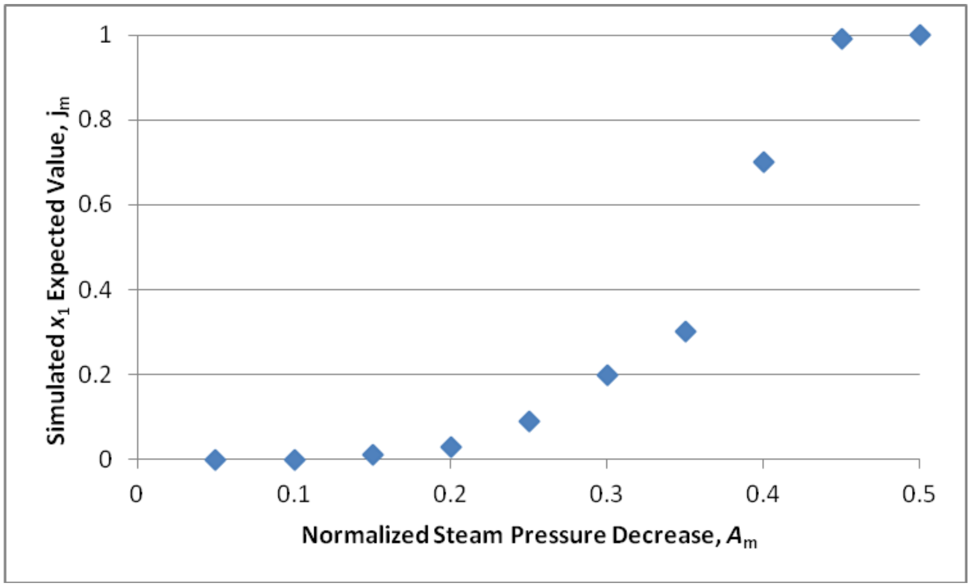


Figure 3.4. SS_1 failure probability as a function of a steam pressure decrease.

3.2.3. Operator Response-Time Models

In the simulation of a special-cause event, the behavior of the operator must be well understood. In Figure 2.3, $g_2(\tau)$ represents operator response times in taking action following activated alarms associated with SS_1. An initial construction of $g_2(\tau)$ can be made using the histogram of operator response times to high-frequency alarms. This provides valuable information about how the operators tend to act when a variable is under alarm (Macwan et al. 1994; Bendoyl et al., 2006; Stylios et al., 1999). In some cases, when operators anticipate that the alarm thresholds have been set conservatively, they are slower to respond to expected nuisance alarms. On the other hand, when operators recognize that alarms tend to trigger a flood of alarms elsewhere in the process, they view these alarms as critical – even though they just signal entry into the yellow-belt zone. To the extent possible, it is important to make quantifiable justifications when modeling operator effectiveness (Hollnagel, 1998; Reason, 2000).

A histogram of approximately 300 observed operator response times to the H-alarm associated with the SMR furnace effluent temperature (Moskowitz et al., 2015) is shown in Figure 3.5. The operator response times are collected using the alarm data log, which records the time of each alarm activation and the time of each operator manipulation. The time between an alarm activation and the initial operator manipulation is the operator response time. This calculation method provides the most accurate data on operator response time. Alarm data are convenient to work with, but without alarm data, process data can be sampled to obtain operator response times. A script can be written to record the times process variables cross their thresholds, as well as the times controlled-variable set points or actuators undergo step-changes (considered to be operator actions) (Pariyani et al., 2012a). In either case, the data sampling interval is important to consider. If data are sampled or recorded infrequently (such as by a composition analyzer), operator response times may be inaccurate. Depending upon the frequency of process data points, reasonable estimates of operator response times can be obtained. The wide range of operator response times, nearly all well represented, suggest many kinds of operator actions. The highest number of responses are associated with the shortest response times – with operators taking action in less than one minute. Nearly all of the operator response times lie between zero and six minutes, with far fewer of longer duration. Past a six minute response time, the number of responses decreases rapidly, with just three response times beyond eight minutes.

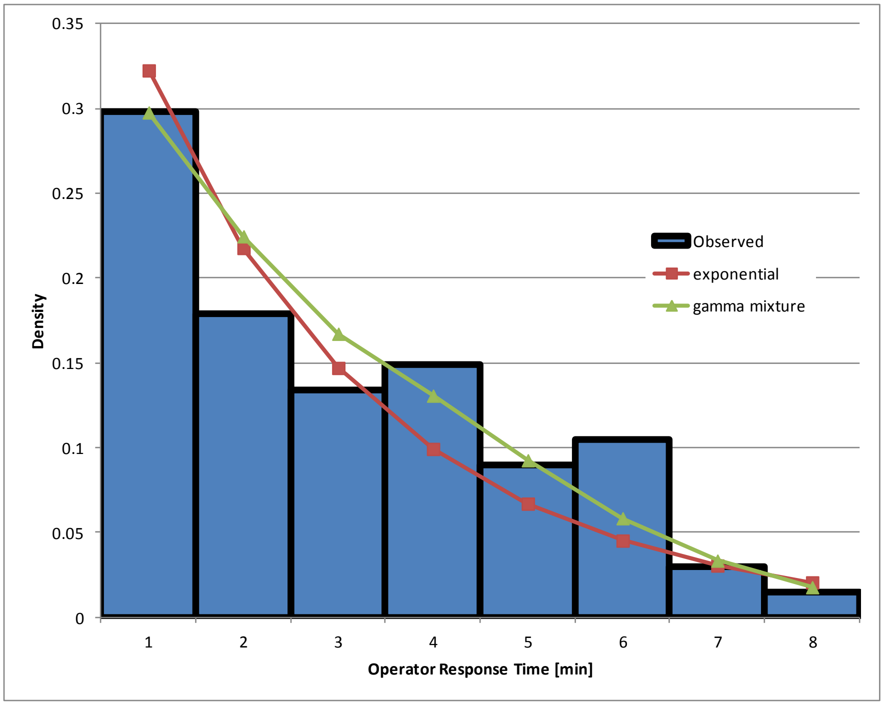


Figure 3.5. Operator response time histogram.

Two parametric distributions are proposed to model the operator response-time distribution. The first distribution is an exponential distribution, which is called Operator Response-Time Model A:

$$g_{2_\mathrm{A}}(\tau) = \lambda e^{-\lambda\tau} \tag{3.2}$$

where λ is a parameter to be estimated by maximizing the likelihood function:

$$L(\lambda|\tau) = \prod_{i=1}^{n} \lambda e^{-\lambda\tau_i} \tag{3.3}$$

where i is the response counter, *n* is the total number of response times measured (on the order of 300), and τ_i is response time *i*. It is often convenient to maximize the log-likelihood function instead of the likelihood function:

$$LL(\lambda|\tau) = \sum_{i=1}^{n}(\ln[\lambda] - \lambda\tau_i) \qquad (3.4)$$

The maximum of the log-likelihood function has a simple analytical form:

$$\lambda = \frac{1}{\bar{\tau}} \qquad (3.5)$$

where $\bar{\tau} = \frac{\sum_{i=1}^{n}\tau_i}{n}$ is the sample mean of the measured response times.

The second is a weighted-sum of three gamma distributions (Operator Response-Time Model B):

$$g_{2_{\mathrm{B}}}(\tau) = \sum_{j=1}^{3}\theta_j \frac{{b_j}^{a_j}}{\Gamma(a_j)}\tau^{a_j-1}e^{-b_j\tau}, \qquad \sum_{j=1}^{3}\theta_j = 1 \qquad (3.6)$$

where each θ_j is a weighting coefficient for gamma distribution *j*, and where a_j and b_j are the parameters of gamma distribution *j*. While any number of gamma distributions can be used, here the fourth distribution gives a negligible increase in the likelihood function (compared to the impact of the third distribution). The eight parameters in Eq. (3.7) are estimated by maximizing the likelihood function:

$$L(a_1, a_2, a_3, b_1, b_2, b_3, \theta_1, \theta_2|\tau) = \prod_{i=1}^{n}\left(\sum_{j=1}^{3} \theta_j \frac{b_j^{a_j}}{\Gamma(a_j)} \tau_i^{a_j-1} e^{-b_j\tau_i}\right) \tag{3.7}$$

Using Newton's optimization method with an analytical Hessian matrix, the parameter values in Table 3.2 were estimated.

Table 3.2. Parameters for Operator Response-Time Models A and B.

λ	a_1	a_2	a_3	b_1	b_2	b_3	θ_1	θ_2	θ_3
0.39	1.83	3.56	4.51	0.68	1.04	0.88	0.48	0.07	0.45

While Operator Response-Time Models A and B represent operator response times well, they do not account for the rate of change of each variable crossing its alarm threshold, as well as the number of activated alarms being monitored by an operator(s). Clearly, operator responses gain urgency, and often speed, when a variable crosses one of its thresholds rapidly. Also, as the number of active alarms decreases, operators are less distracted and respond more rapidly.

To account for the rate of change of each variable when the variable crosses an alarm threshold, using the SMR plant data, operator response times are displayed in Figure 3.6 as a function of the furnace effluent temperature derivative, $\frac{dx}{dt}$, as the temperature crosses its high-alarm threshold. The dependence of the operator response time, τ, on the rate of change is well-described by:

$$\hat{\tau}_C = \kappa_1 e^{-\nu_1 \frac{dx}{dt}} \tag{3.8}$$

where κ_1 and ν_1 are the model parameters. These parameters are estimated by minimizing the sum of the squared errors:

$$SSE_1 = \sum_{i=1}^{n} \left(\kappa_1 e^{-\nu_1 \frac{dx}{dt}_i} - \tau_i \right)^2 \tag{3.9}$$

The estimated values of κ_1 and ν_1 as well as the corresponding SSE_1 value are given in Table 3.3.

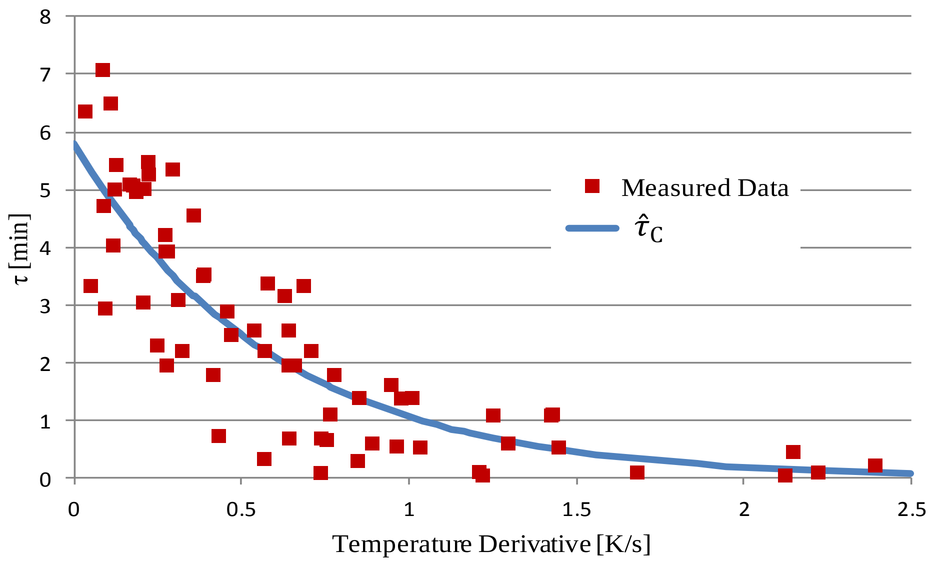


Figure 3.6. Operator response time as a function of temperature rate of change (plant data and model prediction).

Table 3.3. Parameters Used for Operator Response-Time Models C, D and E.

κ_1 [min]	ν_1[s/K]	SSE_1	κ_2[min]	ν_2	SSE_2
4.33	1.54	302	3.8	4.2	413

To construct a $g_2(\tau)$ that accounts for the alarmed variable time-derivative, a stochastic component must be maintained – because if $g_2(\tau)$ were purely deterministic (like $\hat{\tau}_{\mathrm{C}}$), the variance of the safety system failure probability with respect to A_m cannot be calculated. While $\hat{\tau}_{\mathrm{C}}$ is an estimate for the operator response time, it must be incorporated into a random variable distribution for τ. One choice for $g_2(\tau)$ is the exponential distribution having an expected value equal to $\hat{\tau}_{\mathrm{C}}$. The exponential distribution in Eq. (3.8) is known to have an expected value of $\frac{1}{\lambda}$. Therefore, Operator Response-Time Model C is proposed:

$$g_{2_{\mathrm{C}}}(\tau) = \frac{1}{\kappa_1 e^{-\nu_1 \frac{dx}{dt}}} \exp\left[\frac{-\tau}{\kappa_1 e^{-\nu_1 \frac{dx}{dt}}}\right] \qquad (3.10)$$

A similar method can be used to account for the effect of multiple alarm activations in the process. When many alarms are active, competing for operator(s) attention, response times are expected to increase. Here, also, the exponential distribution is appropriate:

$$\hat{\tau}_{\mathrm{D}} = \kappa_2 e^{-\nu_2 \gamma} \qquad (3.11)$$

where γ is the reciprocal of active alarms. The parameters κ_2 and ν_2 are estimated by minimizing the sum of the squared errors:

$$SSE_2 = \sum_{i=1}^{n}(\kappa_2 e^{-\nu_2\gamma_i} - \tau_i)^2 \tag{3.12}$$

The estimated values for κ_2 and ν_2 as well as the corresponding SSE_2 value are given in Table 3.3.

Using this logic, the Operator Response-Time Model D is formulated by setting the expected value of an exponential distribution equal to $\hat{\tau}_{\mathrm{D}}$:

$$g_{2_{\mathrm{D}}}(\tau) = \frac{1}{\kappa_2 e^{-\nu_2\gamma}} \exp\left[\frac{-\tau}{\kappa_2 e^{-\nu_2\gamma}}\right] \tag{3.13}$$

Finally, the Operator Response-Time Model E is formulated that incorporates both $\hat{\tau}_C$ and $\hat{\tau}_D$. The effectiveness of $\hat{\tau}_{\mathrm{C}}$ and $\hat{\tau}_{\mathrm{D}}$ can be compared by their associated SSEs. Herein, the expected value of Operator Response Model E is set equal to:

$$\mu_{3E} = \frac{SSE_1\hat{\tau} + SSE_2\phi}{SSE_1 + SSE_2} \tag{3.14}$$

where the weighting coefficients for each distribution are proportional to their SSEs. This yields Operator Response-Time E model:

$$g_{2_{\mathrm{E}}}(\tau) = \frac{SSE_1 + SSE_2}{SSE_1\left(\kappa_1 e^{-\nu_1 \frac{dx}{dt}}\right) + SSE_2(\kappa_2 e^{-\nu_2\gamma})} \exp\left[-\frac{SSE_1 + SSE_2}{SSE_1\left(\kappa_1 e^{-\nu_1 \frac{dx}{dt}}\right) + SSE_2(\kappa_2 e^{-\nu_2\gamma})}\tau\right] \tag{3.15}$$

Next, Operator Response-Time Models A-E are used to construct informed prior distributions for the failure probability of SS_1, along with dynamic Process Model A. The results are shown in Figure 3.7, along with the binomial likelihood distribution of the measured SS_1 data, Eq. (2.4). The model performance index, $\xi_{i,\mathrm{m}}$, in Eq. (3.1) is used to quantify the model performance of Operator Response-Time Models A-E, with results shown in Table 3.4. Operator Response-Time Models A and B, which are independent of the process state, describe the measured alarm data poorly. As expected, Operator Response-Time Model B, with eight parameters, performs better than Operator Response-Time Model A, with just a single parameter. The incorporation of $\hat{\tau}_\mathrm{C}$ and $\hat{\tau}_\mathrm{D}$ in Operator Response-Time Models C and D, clearly improves the informed prior distributions, with Model C performing better than Model D – expected because SSE_1 < SSE_2. Of the five models, Model E is in the closest agreement with the likelihood data. Given preferred Model E, the choice of process model can be revisited. In Figure 3.8, the four dynamic process models are used to build informed prior distributions with Operator Response-Time Model E. The model performance indices are shown for Process Models A-D in Table 3.5. Once again, Process Model A yields the best agreement with the observed likelihood data, and can be considered the most appropriate process model.

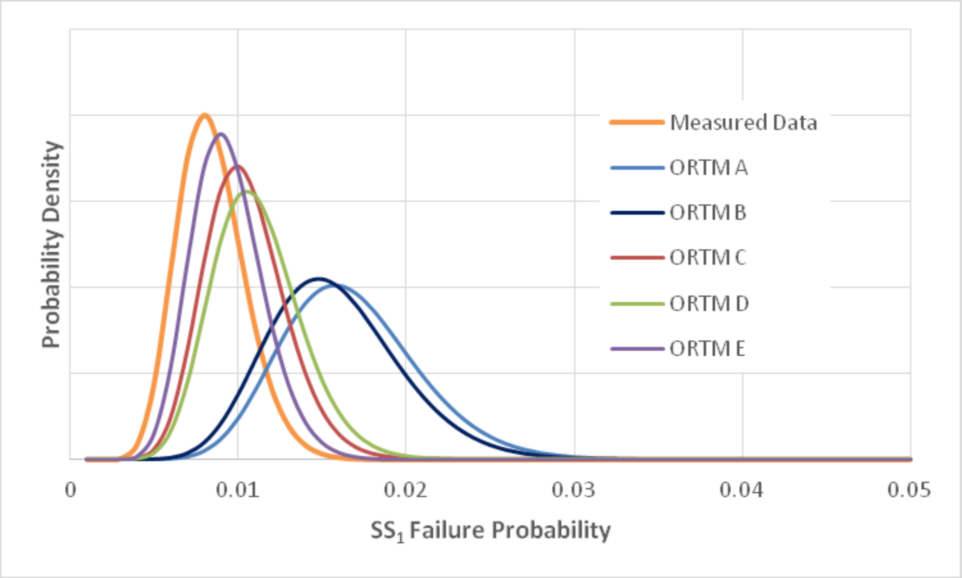


Figure 3.7. *SS_1* informed prior distributions constructed using the five operator response-time models (ORTMs) with dynamic Process Model A.

Table 3.4. Performance Index for Operator Response-Time Models A-E with Process Model A.

	ORTM A	ORTM B	ORTM C	ORTM D	ORTM E
$\xi_{i,\mathrm{m}}$	0.029	0.030	0.633	0.701	0.812

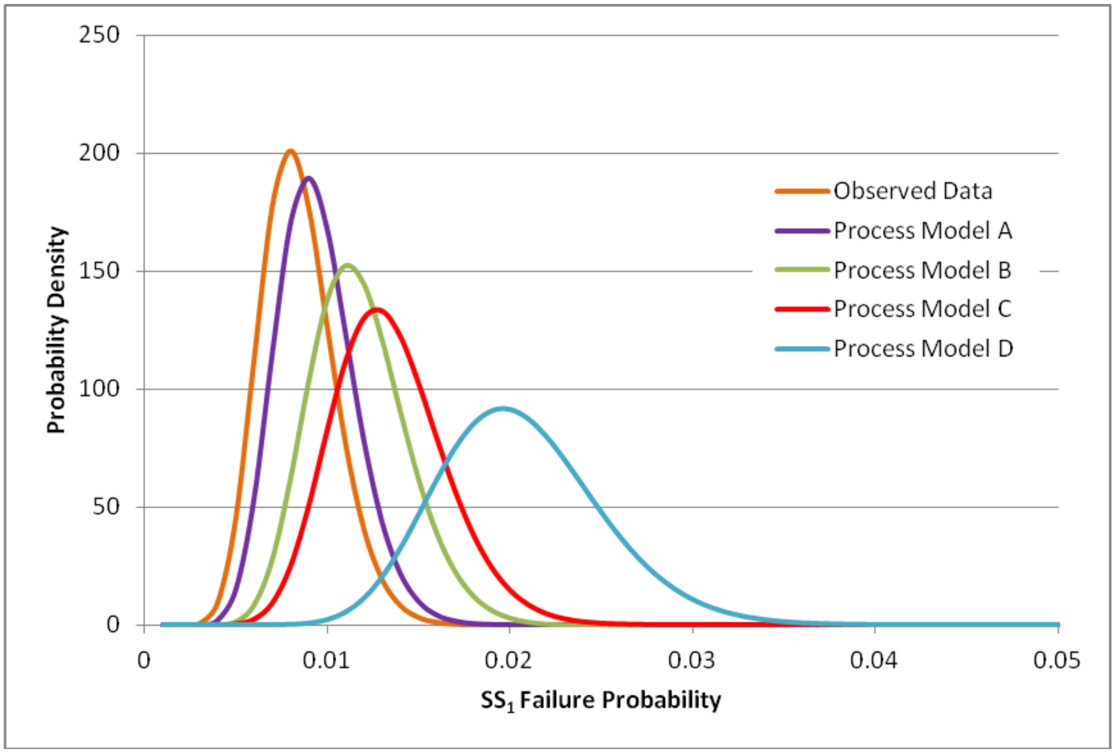


Figure 3.8. SS_1 informed prior distributions constructed using the four process models with Operator Response-Time Model E.

Table 3.5. Performance Index Revisited for Process Models A-D Using Operator Response-Time Model E.

	Process Model A	Process Model B	Process Model C	Process Model D
$\xi_{i,\mathrm{m}}$	0.812	0.481	0.325	0.051

3.3. Modeling SS_2 Failures Using Models with Parameters Estimated from SS_1 Failures

Once the three types of models (process, special-cause event, and operator response-time models) are chosen and their parameters are estimated, they are used to estimate the failure probabilities of SS_1. These models must then be adjusted to handle simulations that involve the activation of SS_2 (i.e., after the failure of SS_1). While it is desirable to keep the models intact, a few adjustments are recommended.

The dynamic process model, Process Model A, should not be altered much to simulate SS_2 activation events – because the simulations from high and high-high alarms, and beyond, are similar, with small changes in physical properties as temperatures rise. In general, Process Model A adjustments would be required when physical and chemical phenomena change abruptly – for example, with shifts from laminar to turbulent flows, or the introduction of two-phase flows.

The Special-Cause Event Occurrence Model needs significant adjustment because the $g_1(A_m)$ distribution used for simulating L/H alarm activations infrequently activates LL/HH alarms. To achieve this, a normal distribution is chosen for $g_1(A_m)$, having mean $\overline{A_m}$ and standard deviation σ. A lower-tail, bounded by θ_L and an upper-tail bounded by θ_U, each two standard deviations from the mean, are defined:

$$\theta_L = \overline{A_m} - 2\sigma \,;\, \theta_U = \overline{A_m} + 2\sigma \tag{3.16}$$

noting that θ_L is a special-cause magnitude closer to zero; that is, closer to typical operation. The normal distribution is described by θ_L and θ_U, rather than the typical mean and standard deviation:

$$g_1(A_m) = \frac{1}{\sqrt{2\left(\frac{\theta_U-\theta_L}{4}\right)^2\pi}} \exp\left(-\frac{\left(A_m-\frac{\theta_U+\theta_L}{2}\right)^2}{2\left(\frac{\theta_U-\theta_L}{4}\right)^2}\right) \tag{3.17}$$

The lower bound of θ_L is set at A_m values for which the L/H alarms failed in simulation – with smaller choices of θ_L yielding many simulations where the LL/HH alarms are not activated. Referring to Figure 3.4, with SS_1 failures frequently observed when $A_m \geq 0.4$, 0.4 is a good lower bound for θ_L. θ_U is set such that the special-cause events are of interest and relevance. Three Special-Cause Event Occurrence models are shown in Figure 3.9, each sharing $\theta_L = 0.4$. SCEM A has θ_U closest to θ_L, and samples special cause events that are most likely (closest to typical operation), yet have the least potential for SS_2 failures. The other extreme is SCEM C, which has θ_U furthest from θ_L. SCEM B, having $\overline{A_m} = 0.56$, is chosen as an interior candidate to analyze the risk of SS_2 failures. The choice of $g_1(A_m)$ has a significant impact on the estimated failure probabilities – the user must keep this in mind when analyzing the simulation results and making statements about the risk of the process. The failure probability estimate attempts to describe the probability of failure while undergoing a special-cause event sampled from $g_1(A_m)$, which is very different than typical day-to-day process fluctuations.

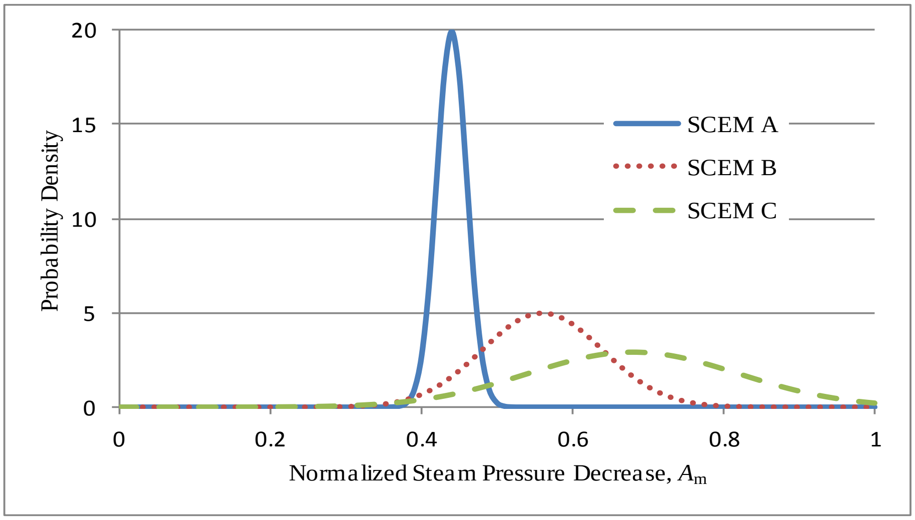


Figure 3.9. Special-Cause Event Models for SS_2 simulation.

The Operator Response-Time Models also need adjustment, it being expected that operators react quicker to alarm activations that are associated with more urgent consequences. Given an immediate threat of an automatic plant shutdown, operator actions should be accelerated. To account for this, when SS_1 is activated to SS_2, the time response to the SS_2 activation is divided by the ratio of the 90% percentile of operator response to a SS_1 activation, $\tau_{90\%}$, over the interlock shutdown time, Δt_{int}. For this case, the Operator Response-Time Model E takes the form:

$$g_{2_{\mathrm{E}}}(\tau) = \frac{SSE_1+SSE_2}{SSE_1\left(\kappa_1 e^{-\nu_1 \frac{dx}{dt}}\right)+SSE_2\left(\kappa_2 e^{-\nu_2 \gamma}\right)} \frac{\Delta t_{\text{int}}}{\tau_{90\%}} \exp\left[-\frac{SSE_1+SSE_2}{SSE_1\left(\kappa_1 e^{-\nu_1 \frac{dx}{dt}}\right)+SSE_2\left(\kappa_2 e^{-\nu_2 \gamma}\right)} \left(\frac{\Delta t_{\text{int}}}{\tau_{90\%}} \tau\right)\right] \tag{3.18}$$

Having constructed, chosen, and regressed the three types of models using the H-alarm data, adjustments are made to model SS_2 activation events. An informed prior distribution is then constructed for the failure probabilities of SS_2. Figure 3.10 shows the resulting informed prior and associated posterior distributions describing the failure probability of SS_2. The sparse alarm data are used to build binomial likelihood distributions that modify the prior distributions to form posterior distributions. It can be seen that the most accurate posterior distribution (formed using Process Models A and Operator Response-Time Model E with an *urgency* adjustment for time responses in SS_2) is not shifted as dramatically as the posterior distribution formed using the simple informed prior distribution. This indicates that the most accurate posterior distribution is more effective at handling the special-cause simulations – leading to more accurate failure probability predictions.

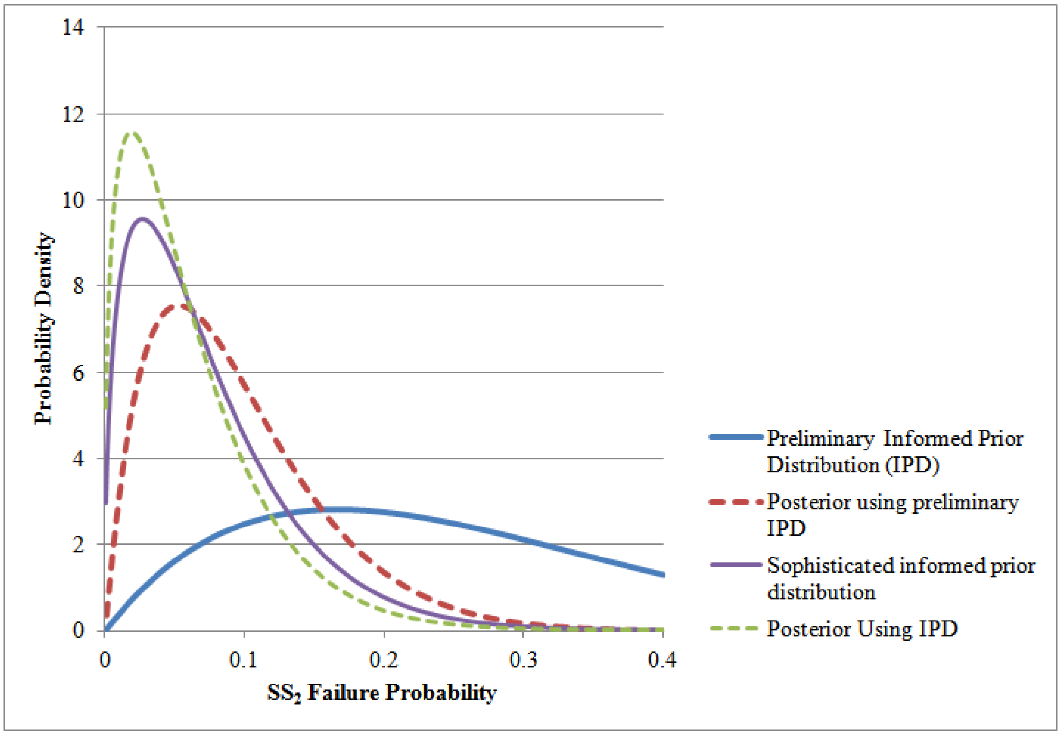


Figure 3.10. Informed prior distributions and associated posterior distributions describing the failure probability of SS_2.

The interpretation of the posterior distribution is that during a severe loss in steam pressure, the probability that the process will undergo an automatic shutdown is on the order of 5%. This allows engineers responsible for setting reliability estimates to have quantifiable justification when they do so. An estimate for the reliability for various special-cause events can provide engineers with a broader understanding of the events that pose the greatest odds of an interlock activation, thus motivating different designs to handle these events. The operator also benefits from these distributions, as he/she becomes aware that, during such a severe pressure drops, his/her reactions have been projected to result in interlock activations after on the order of 5% of occurrences. If this value is perceived to be too high, the operator may be motivated to act more urgently during this type of event.

3.4. Conclusions

Alarm and safety interlock system failure probabilities are difficult to estimate, but warrant careful consideration using the strategies introduced herein. The safety of the operators and employees at a chemical plant, and those in the neighboring community and environment, is crucial to the chemical process industries.   For safety interlock systems and their associated alarms, statistical techniques on the sparse records of activations are alone insufficient to make meaningful evaluations of their failure probabilities. The usage of alarm and process data associated with the relatively frequent alarm activations (e.g., H-alarms) to systematically improve the performance of less frequently activated alarms (e.g., HH-alarms) and safety interlock systems is very promising. As demonstrated for an SMR plant example, the three types of models can be applied to a variety of chemical manufacturing processes. The resulting models provide new insights into the performance of rarely-activated alarm and safety interlock systems, for which historical data are sparse.

Chapter 4

Understanding Rare Safety and Reliability Events Using Transition Path Sampling

4.1. Introduction

Safety and reliability are paramount to the chemical manufacturing industries. Because chemical plants are often operated at high temperatures and pressures, and with hazardous materials, the potential for adverse human health and environmental impacts exists. With proper process design, effective implementation of control and safety instrumented (SIS) systems mitigate such risks. Less severe are product losses which result from poor plant reliability. As chemical manufacturing processes approach dangerous operating conditions, automatic safety interlocks activate, shutting them down before dangerous consequences are realized. When functioning correctly, the dangerous consequences are avoided, but manufacturing processes lose valuable production over the time period encompassing the automatic shutdown, process maintenance, and startup. Furthermore, plant startup is often the most dangerous mode of operation because large transients are often not as well understood compared with steady-state and cyclic operations. There is clear motivation, both financially and ethically, to prevent chemical manufacturing processes from operating in regions where safety interlocks are activated – resulting in automatic plant shutdowns or potentially in safety incidents.

Safety interlocks are often based on HAZOP (hazard and operability analysis) (Kletz, 1999; Venkatasbramanian et al., 1994; Kennedy et al., 1998) and LOPA (layer of protection analysis) (Dowell, 1998; Summers, 2003). With HAZOP, potential hazards to personnel and capital equipment that may occur during process operation are identified through a meticulous (yet qualitative) procedure. It provides “a more complete identification of the hazards, including information on how hazards can develop as a result of operating procedures and operational upsets in the process” (Crowl and Louvar, 1990). With LOPA, the probabilities of identified hazards occurring are maintained under a low, pre-specified value by utilizing a system of high-performing, independently-acting safety systems. Said differently, the hazards identified by HAZOP analysis are mitigated to lower-consequence events (such as plant shutdowns) with high probability by using safety systems identified through LOPA. Through these analyses, safety interlock thresholds are determined. From a reliability perspective, operators seek to avoid costly shutdowns by adjusting valves when control systems are too slow or insufficient in responding to severe disturbances (known as special-cause events). Avoiding shutdowns is also beneficial from a safety perspective, as transient shutdowns and startups are avoided.

Operators are aided by an alarm structure in which process variables pre-specified to be important to the reliability and safety of the process are equipped with alarms. When a variable moves outside of its typical (safe) operating region, the green-belt zone, either a low (L) or a high (H) alarm activates accordingly. Often, process variables have several levels of alarms, possibly a yellow belt-zone (bounded by L and H alarms), an orange belt-zone (bounded by LL and HH alarms), and a red belt-zone (bounded by LLL and HHH alarms). Such an alarm scheme is depicted in Figure 4.1. Here, in Figure 4.1a, a process variable is displayed over months and years, normally residing within its green-belt zone – and, when perturbed into its yellow-belt zone, safety systems/operator actions usually return it to its safe green-belt zone. Rare events result in the automatic shutdown (safety interlock) of the process, followed by a shutdown and restart, which occur over minutes and hours, as shown in Figure 4.1b. The safety-interlock shutdown is activated when the process variable resides in the red belt-zone for a pre-specified length of time, Δt_{int}, typically on the order of seconds to minutes. As a variable moves into each successive belt-zone, the operator becomes aware that interlock activation is impending and takes more severe actions to return the variable to its green-belt zone.

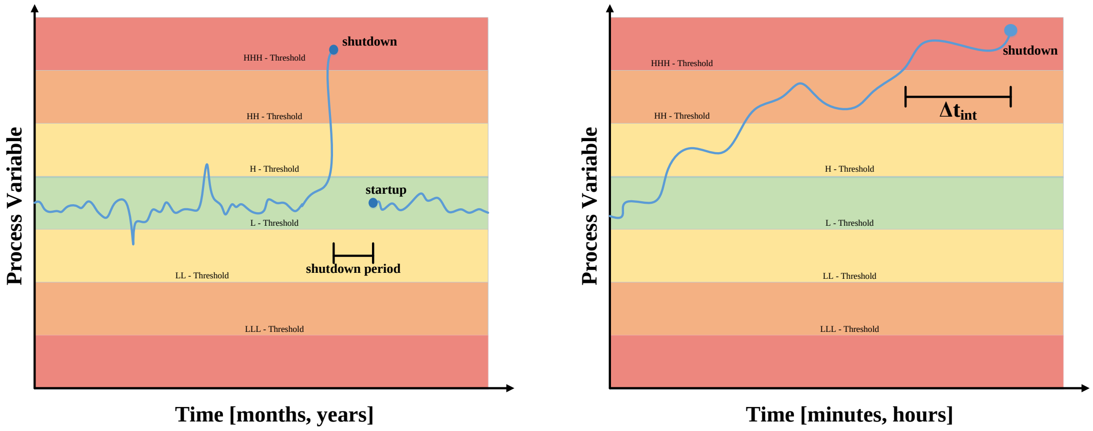


Figure 4.1. Alarm belt-zones and interlock shutdown for a process variable.

The alarm thresholds are set in the process commissioning phase (Hollifield et al., 2010), with competing objectives to: (1) assure that when an alarm is activated operators have sufficient time to act, avoiding subsequent (more severe) alarms or interlock activations, and (2) that the alarm isn’t a nuisance, often activated unnecessarily, and often disregarded by operators. Commissioning is usually performed using expert knowledge of process behavior (based upon the actions of similar processes and upon insights gained in the process design phase), and tests to observe typical transient responses of the variables.

Clearly, alarms are commissioned to alert operators to *postulated*, more common, events that could propagate to interlock activation. But, alarm structures may be insufficient to alert operators to rare or *un-postulated* events. Such unforeseen safety events have the potential to move to the red belt-zone and activate the interlock shutdown faster than the alarms/safety systems are designed to handle. These events may arise early involving variables that are not alarmed, or when some combination of variables leads to such an event. While these events may be easily handled by operators, without proper alarming, operators may not be able to prevent automatic shutdowns.

A quantitative technique to better identify and understand events that lead to process shutdowns would be very useful to engineers responsible for commissioning alarms and operators that respond to those alarms. This paper introduces transition path sampling (TPS) as such a technique for application in the chemical manufacturing industries. TPS is a Monte-Carlo sampling strategy that simulates process models as they propagate toward interlock-activating events. Trajectories of these events are randomly generated, uncovering many un-postulated events, and enabling postulated events to be better understood. With many similar trajectories generated, the probability of a typical trajectory can be estimated, identifying the most likely unsafe events, suggesting more effective alarm thresholds. TPS has been widely investigated by the molecular dynamics community to study rare molecular events (Bolhuis et al., 2002; Dellago et al., 2002), but the application of TPS to process dynamics for studying rare interlock-activating events is novel and presents its own challenges.

4.2. Transition Path Sampling

TPS was invented to study rare molecular dynamic trajectories; for example, the dissociation of a weak acid in an aqueous solution. A weak acid, such as hydrofluoric acid (HF), dissociates in water in approximately a millisecond, but its dissociation event occurs in just nanoseconds (Bolhuis et al., 2002). Hence, its initiation time is on the order of 10^6 times longer than the event itself! Clearly, simulation of the initiation/dissociation sequence involves excess computation time to track the *initiation* phase. In TPS, to circumvent this, just one initiation/dissociation event is simulated. Then, at a random time, t', along the event trajectory (spanning [0, t_f]), state variables are randomly perturbed (such as atom locations and momenta). This new state is simulated forward spanning $[t^{'}, t_f]$ and backwards spanning $[t^{'}, 0]$. If the acid is associated at $t = 0$, and dissociated at $t = t_f$, then a second rare-event trajectory has been generated, which may be accepted. Over many iterations, numerous rare-event trajectories can be generated, with minimal computational effort in simulating the initiation phase (Dellago et al., 2002).

When applied to process dynamics, TPS can identify and explain rare interlock-activating events. The models and time scales in process dynamics are vastly different from those in molecular dynamics, but the challenge of simulating rare-events is similar. A typical rare interlock-activating event may occur over years to decades, while the event itself occurs over minutes to hours. Similarly, TPS can be used to circumvent simulating the initiation phase – the time in between rare safety-events of interest. As shown in Figure 4.2a, a complete trajectory is identified by simulation (or by a rare safety event in a plant or similar facility elsewhere) and then randomly perturbed, as shown in Figure 4.2b, allowing for the generation of many trajectories. These perturbations are applied to state variables, often process unit temperatures, compositions, and pressures. The perturbations are also applied to stochastic variables – either noise to operational and design parameters that effect multiple balance equations (*parametric* noise), or noise introduced as a term to a single balance equation (*non-parametric* noise). The parametric noise can be used to explore rare-events that may arise when operational parameters (such as product demand rate or feed conditions) fluctuate in a specific pattern. Design parameters (such as reaction rate constants or binary interaction coefficients) are fixed over the course of simulations, but perturbations over small ranges can yield rare-event trajectories. Non-parametric noise, introduced in a well-scaled term added to a process unit balance equation, can yield physical interactions not in the first-principles model – such as a side reaction or leak in a vessel. Careful formulation of these terms can improve the effectiveness of non-parametric noise in predicting plant shutdowns and accidents.

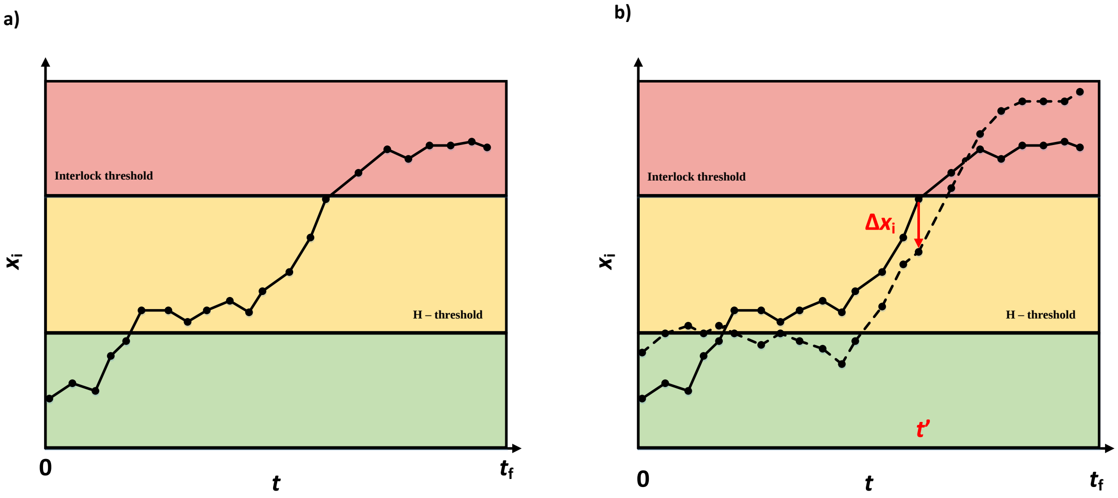


Figure 4.2. TPS used to generate a trial rare-event trajectory from an initial trajectory.

Also, when applying TPS, unlikely rare-event trajectories are often generated easily using process models. But, careful analysis is helpful in generating initial trajectories that enable TPS to discover serious unpostulated plant shutdowns and accidents.

The outline of the TPS algorithm is shown in Table 4.1. The random perturbation of rare-event trajectories may lead to the development of new rare-event trajectories, as shown by Figure 4.2b. When this new trajectory is accepted, further iterations take place from this trajectory. If the new trajectory is not a rare-event trajectory, or the new trajectory is an unaccepted rare-event, further iterations take place from the previous trajectory. The random nature of TPS allows for interesting, possibly unpostulated, rare-event trajectories to be identified. Additionally, by simulating many trajectories of postulated interlock-activating events, engineers and operators can gain a more quantitative understanding of such events.

Table 4.1. TPS Algorithm

1. Identify an initial safety event trajectory
2. Choose a random time, *t*', along the trajectory
3. At *t*', perturb state variables, $\underline{x}$, to $\underline{x}'$ and stochastic variables, $\underline{\eta}$, to $\underline{\eta}'$
4. Integrate forward from *t*' to t_f, and backwards from *t*' to 0
5. Determine if this trial trajectory identifies a safety-event
6. If yes, consider accepting the trial safety event trajectory as the new trajectory, where the trajectory acceptance criteria are defined in Figure 4.5.
7. Return to step 2.

4.2.1. Backward Integration

An important difference between TPS in molecular and process dynamics is the backward integration approach. In molecular dynamics, force balances,

$$\frac{d^2x_i}{dt^2} = F(x) \tag{4.1}$$

are solved, which are second-order ordinary differential equations (ODEs), noting that x_i is the position of atom i, and $F_i(x_i)$ are the forces exhibited on atom i by all other atoms. At t', when backward integration is initiated, the initial conditions are $x_i(t^{'})$ and $\frac{dx_i}{dt}|_{t'}$. For backward integration, the sign of the first derivative is reversed, $-\frac{dx_i}{dt}|_{t'}$, to enable a stable forward integration to $t = 0$. However, in process dynamics, typical systems are first-order ODEs, taking the form:

$$\frac{dx_i}{dt} = f(x, \eta, u) \tag{4.2}$$

where x is a vector of state variables (e.g., moles or energy), η is a vector of stochastic variables, and u is a vector of input variables (e.g., feed conditions). The same approach for backward integration in molecular dynamics is not applicable. If the signs of the time-derivative terms were reversed:

$$-\frac{dx_i}{dt} = f(x, \eta, u)\ , \tag{4.3}$$

the resulting Jacobian matrices of *f*, for typical process systems, would have large positive eigenvalues. Even with linear multi-step integrators (e.g., backward-difference formulae), numerically unstable solutions would be obtained that are often chaotic. Because the resulting trajectories are usually inaccurate, a boundary-value optimization is often formulated, as discussed next.

In one approach, initial conditions in the vector, x_0, are manipulated such that when integration proceeds along $[0, t']$, the state variables at t', $x^*(t^{'})$, approach the desired state variables at t', x'. Said differently, the backward-integration step is performed by solving:

$$\min_{x_0} \sum [x^*(t^{'}) - x']^2 \tag{4.4}$$

$$s.t.$$

$$\frac{dx_i^*}{dt} = f(x, \eta, u)$$

$$x(0) = x_0$$

In this approach, a shooting method (Bock et al., 2000) is used as illustrated in Figure 4.3. Here, $x_{0,1}$ is the initial guess for x_0. After forward integration to t', $x_1^*(t^{'})$ is substantially less than x'. To compensate, a larger guess, $x_{0,2}$, is chosen, yielding $x_2^*(t^{'})$, which is too large. The next guess value, $x_{0,3}$, yields $x_2^*(t^{'})$, which is sufficiently close to x'. This optimization effectively performs the function of backward integration. When the initial value lies within the typical operation region, a rare-event trajectory has been located. If not, this trial trajectory is discarded and a new t' and x' are chosen from the previous trajectory.

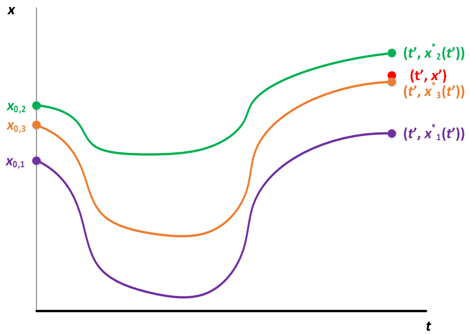


Figure 4.3. Boundary-value optimization to indirectly perform backward integration using initial-value shooting.

In perhaps a more common approach, used herein, the optimization is formulated using orthogonal collocation on finite elements (Cuthrell et al., 1989). This technique involves introducing x_i^K(t), a *K*-order polynomial approximation of *x* over the time range, $t = [t_{i-1}, t_i]$, with *N* finite elements spanning the total time range [0, *t*']. Each x_i^K(t) is constructed using *K* + 1 interpolations of $f(x(\tau), \eta(\tau), u)$, where τ is a normalized time spanning $[t_{i-1}, t_i]$. The minimization problem:

$$\min_{x_0}[x_N^K(t') - x']^2 \tag{4.5}$$

$$s.t.$$

$$x_i^K(t) = \alpha_0^i + \alpha_1^i t + \cdots + \alpha_K^i t^K \; ; \; i = \{1, \dots, N\}$$

$$0 = \frac{dx_i^K(t)}{dt} - f(x(t), \eta(t), u); \; i = \{1, \dots, N\}$$

$$x_0 = x_1^K(0)$$

is solved for x_0. Similar to the shooting method, x_0 is the initial condition that places *x*(*t*) close to *x*'at *t* = *t*'. The orthogonal collocation method is summarized in Figure 4.4.

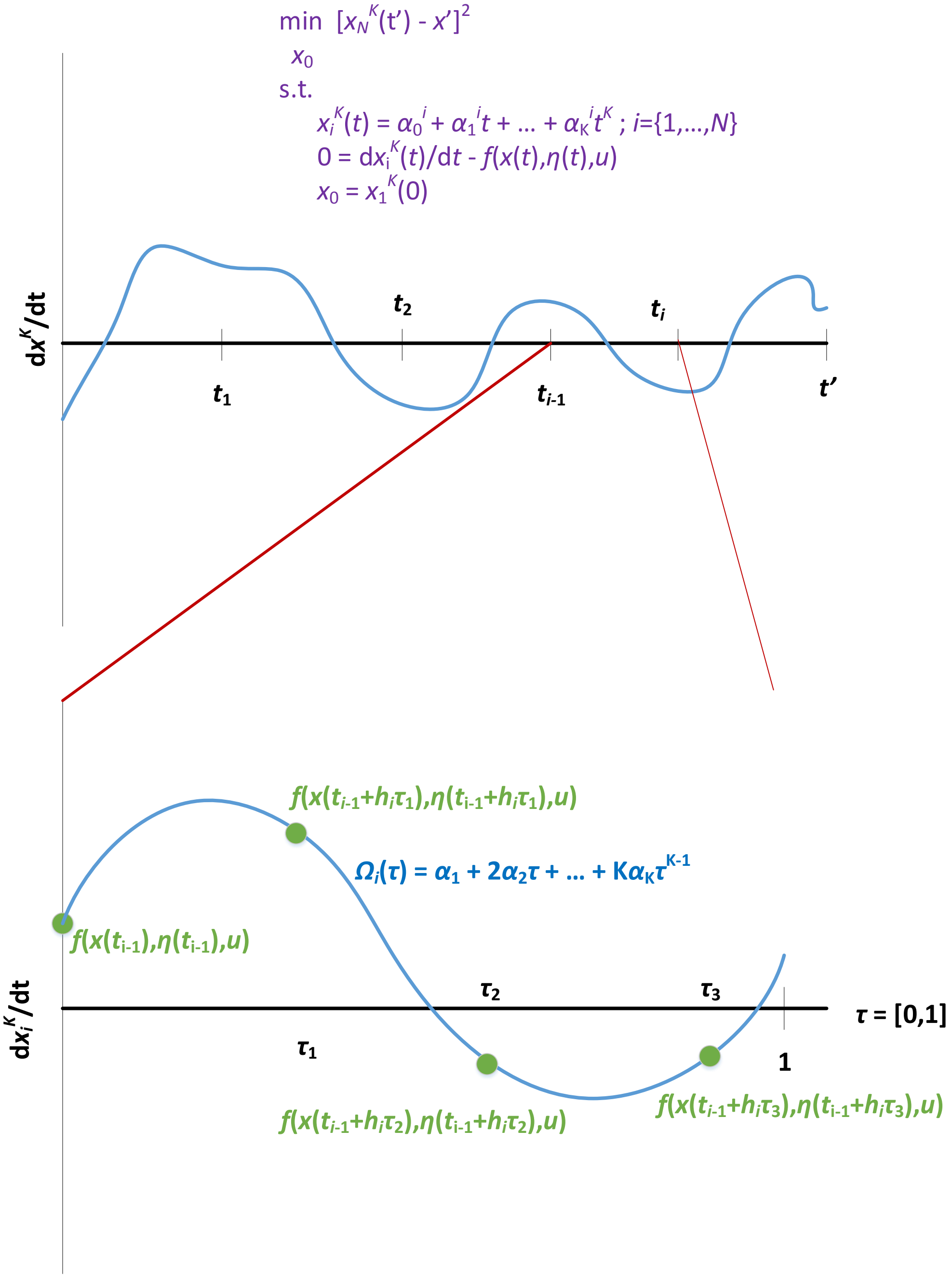

$\min_{x_0} \; [x_N^K(t') - x']^2$
s.t.
$x_i^K(t) = \alpha_0^i + \alpha_1^i t + \ldots + \alpha_K^i t^K$; $i=\{1,\ldots,N\}$
$0 = dx_i^K(t)/dt - f(x(t),\eta(t),u)$
$x_0 = x_1^K(0)$
dx^K/dt
t_1
t_2
t_{i-1}
t_i
t'
$f(x(t_{i-1}+h_i\tau_1),\eta(t_{i-1}+h_i\tau_1),u)$
$\Omega_i(\tau) = \alpha_1 + 2\alpha_2\tau + \ldots + K\alpha_K\tau^{K-1}$
$f(x(t_{i-1}),\eta(t_{i-1}),u)$
dx_i^K/dt
τ_1
τ_2
τ_3
1
$\tau = [0,1]$
$f(x(t_{i-1}+h_i\tau_2),\eta(t_{i-1}+h_i\tau_2),u)$
$f(x(t_{i-1}+h_i\tau_3),\eta(t_{i-1}+h_i\tau_3),u)$

Figure 4.4. Orthogonal collocation over finite-elements.

4.2.2. Trajectory Likelihood Calculation

Another key difference between the two TPS formulations involves the estimation of trajectory likelihood in Step 6, Table 4.1. In molecular dynamics, the likelihood of a trajectory is simply the product of Boltzmann factors corresponding to the atomic configurations at each time step. In process dynamics, likelihood probabilities rely on available or simulated process data. The likelihood probability, p, is often calculated using:

$$p = h_A(0) \times h_B(t_\mathrm{f}) \times f(x_0) \times \prod_{i=1}^{nm} g(\eta(t_i)) \qquad (4.6)$$

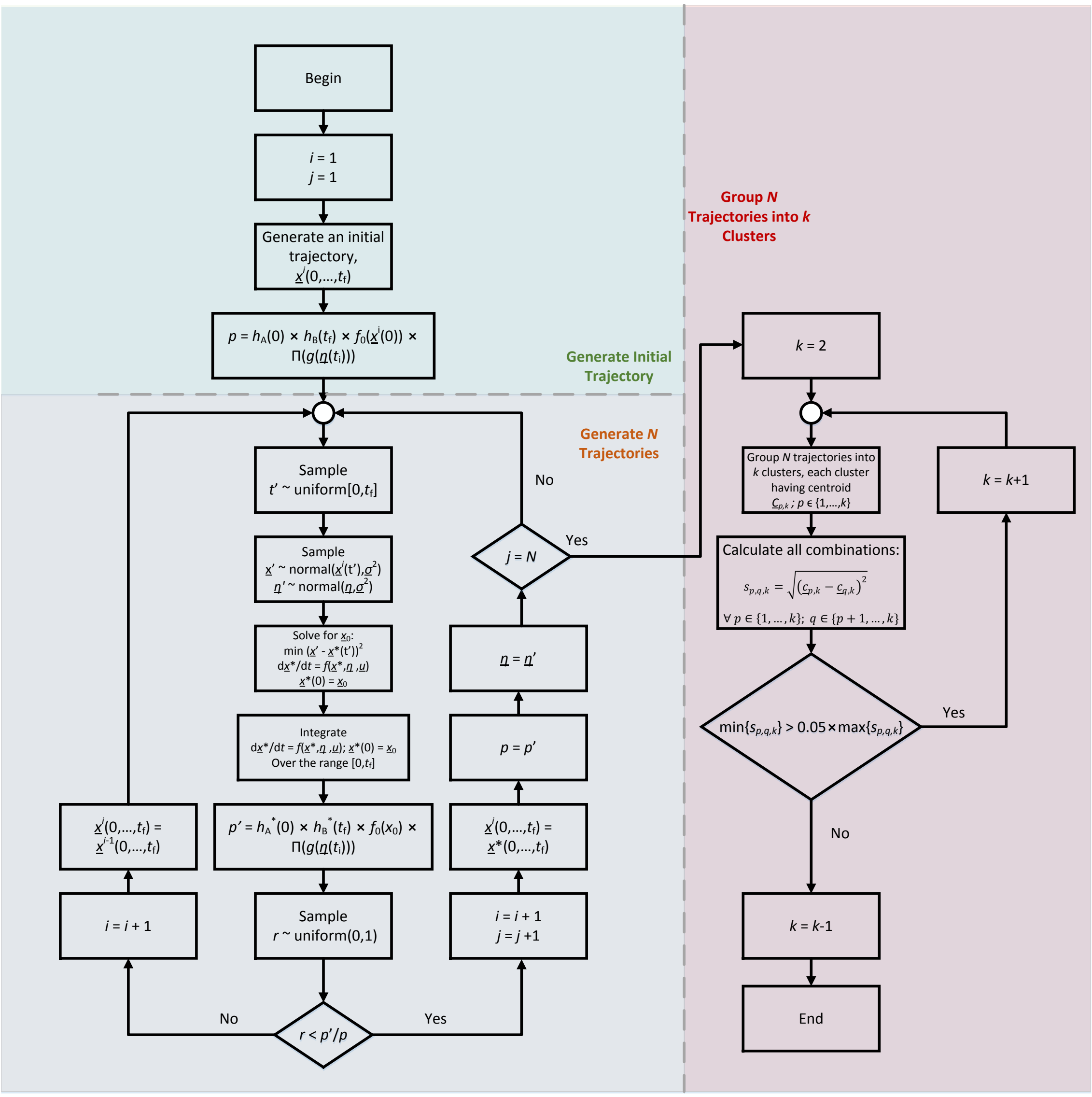
where $h_A(0)$, a binary variable, is unity when $x(0)$ lies within normal operation conditions (green-belt zone) at $t = 0$, $h_B(t_f)$, a binary variable, is unity when $x(t_f)$ satisfies the criteria of unsafe or unreliable conditions (red-belt zone) at $t = t_\mathrm{f}$, $f(x_0)$ is the likelihood of the initial conditions, $\eta(t_i)$ are the stochastic variables at time t_i, $g\big(\eta(t_i)\big)$ is the likelihood of stochastic variables at time t_i, and m is the number of stochastic samples taken along $[0, t']$ [8,9]. The likelihood of initial conditions can be estimated using process data, with $f(x_0)$ increasing as the population of x_0 increases. In large part, stochastic variables are related to noise (parametric or non-parametric), often expressed as probability distributions (commonly, normal distributions). In these cases, $g\big(\eta(t_i)\big)$ are simply the likelihood of noise η at time, t_i. When η is a normal distribution with mean 0 and variance, σ^2:

$$g\big(\eta(t_i)\big) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\eta^2}{2\sigma^2}\right) \tag{4.7}$$

Note that when these likelihood probabilities are very small, they are often expressed on a log scale.

4.2.3. Full TPS Algorithm

The full TPS algorithm for sampling process safety-events is shown in Figure 4.5, with the three phases: generating an initial trajectory, generating N unique trajectories, and grouping the N trajectories into k clusters. The algorithm begins with counters $i = 1$ and $j = 1$, an initial trajectory, $x^1(0,\ldots,t_\mathrm{f})$, and its associated likelihood p, calculated using Eq. (6). At time t', the $x^1(t')$ trajectory is perturbed using normal distributions to develop the state x'. The boundary-value problem in Eq. (4) is solved to obtain x_0, which is then integrated over $[0,t_\mathrm{f}]$ to obtain the trial trajectory, $x^*(0,\ldots,t_\mathrm{f})$. The likelihood of this trajectory, p', is calculated using Eq. (6). A random number, r, in the range [0,1] is sampled; when less than $\frac{p'}{p}$,the trial trajectory is rejected, otherwise, it is accepted. If $x^*(0,\ldots,t_\mathrm{f})$ is accepted, $i = i + 1$, $j = j + 1$, $x^i(0,\ldots,t_\mathrm{f}) = x^*(0,\ldots,t_\mathrm{f})$, and $p = p'$. If $x^*(0,\ldots,t_\mathrm{f})$ is rejected, $i = i + 1$ and $x^i(0,\ldots,t_\mathrm{f}) = x^{i-1}(0,\ldots,t_\mathrm{f})$. When N unique trajectories have been calculated ($j = N$), they are grouped into $k = 2$ clusters using the k-means clustering technique (Hartigan et al., 1979). The Euclidean distance, s, is calculated between all pairs of centroids (the center of each cluster) and if the distance between the closest pair is greater than 5% of the distance between the furthest pair, *k* is increased by one and new centroids are calculated. Eventually, the additional new cluster will be sufficiently close to an existing cluster, and the loop will be terminated.



Begin
$i = 1$
$j = 1$
Generate an initial trajectory, $\underline{x}^{i}(0,\ldots,t_f)$
$p = h_A(0) \times h_B(t_f) \times f_0(\underline{x}^{i}(0)) \times \Pi(g(\underline{\eta}(t_i)))$
Group N Trajectories into k Clusters
Generate Initial Trajectory
Generate N Trajectories
$k = 2$
Sample $t' \sim \text{uniform}[0,t_f]$
No
Yes
$j = N$
Group N trajectories into k clusters, each cluster having centroid $\underline{C}_{p,k}$; $p \in \{1,\ldots,k\}$
$k = k+1$
Sample $\underline{x}' \sim \text{normal}(\underline{x}^{i}(t'),\underline{\sigma}^2)$ $\underline{\eta}' \sim \text{normal}(\underline{\eta},\underline{\sigma}^2)$
Calculate all combinations:
$s_{p,q,k} = \sqrt{(\underline{c}_{p,k} - \underline{c}_{q,k})^2}$
$\forall\, p \in \{1,\ldots,k\};\ q \in \{p+1,\ldots,k\}$
Solve for $\underline{x}_0$: $\min\,(\underline{x}' - \underline{x}^*(t'))^2$ $d\underline{x}^*/dt = f(\underline{x}^*,\underline{\eta},\underline{u})$ $\underline{x}^*(0) = \underline{x}_0$
$\underline{\eta} = \underline{\eta}'$
Integrate $d\underline{x}^*/dt = f(\underline{x}^*,\underline{\eta},\underline{u})$; $\underline{x}^*(0) = \underline{x}_0$ Over the range $[0,t_f]$
$p = p'$
Yes
$\min\{s_{p,q,k}\} > 0.05 \times \max\{s_{p,q,k}\}$
$\underline{x}^{i}(0,\ldots,t_f) = \underline{x}^{i-1}(0,\ldots,t_f)$
$p' = h_A{}^{*}(0) \times h_B{}^{*}(t_f) \times f_0(x_0) \times \Pi(g(\underline{\eta}(t_i)))$
$\underline{x}^{j}(0,\ldots,t_f) = \underline{x}^*(0,\ldots,t_f)$
No
$i = i + 1$
Sample $r \sim \text{uniform}(0,1)$
$i = i + 1$
$j = j + 1$
$k = k-1$
End
No
Yes
$r < p'/p$

Figure 4.5. TPS algorithm for calculating trajectories of process safety-events.

4.3. Exothermic CSTR Example

The TPS algorithm is demonstrated using a familiar example – that of the exothermic CSTR. With only two differential balance equations, a heat and material balance, and PI-control, this example has the benefit of being low-dimensional. The existence of multiple (high- and low-conversion) stable steady-states (Balakotaiah et al., 1983), and in particular, the infrequent transitions between them, provide an excellent example of the potential of using TPS to study rare, yet important, safety-events.

Consider a model for the jacketed exothermic CSTR with reaction:

$$\mathrm{A} \rightarrow \mathrm{P} \ , \tag{4.8}$$

a schematic of which is shown in Figure 4.6. The temperature and inlet concentration of A are T_f, $C_{\mathrm{A,f}}$, and the outlet temperature and concentration of A, T, C_A, are calculated as a function of time. The reactor is assumed to have perfect level control, with equal inlet and outlet volumetric flow rates, F. The cooling jacket is assumed to be sufficiently large such that the temperature change of the cooling fluid, T_c, is negligible. The reaction has elementary kinetics and an Arrhenius rate constant; i.e.,

$$r = k_0 \exp\left(-\frac{E_a}{RT}\right) C_\mathrm{A} \tag{4.9}$$

where r is the intrinsic reaction rate, k_0 is the pre-exponential factor, E_a is the activation energy, and R is the gas constant. The derivatives of C_A and T with respect to time, t, are:

$$V\frac{dC_{\mathrm{A}}}{dt}=\frac{V}{\tau}\left(C_{\mathrm{A,f}}-C_{\mathrm{A}}\right)-Vk_0\exp\left(-\frac{E_{\mathrm{a}}}{RT}\right)C_{\mathrm{A}} \tag{4.10a}$$

$$\rho VC_{\mathrm{p}}\frac{dT}{dt}=\frac{\rho VC_{\mathrm{p}}}{\tau}(T_{\mathrm{f}}-T)+UA(T_{\mathrm{c}}-T)+V\Delta Hk_0\exp\left(-\frac{E_{\mathrm{a}}}{RT}\right)C_{\mathrm{A}} \tag{4.10b}$$

where τ is the residence time, U is the overall heat-transfer coefficient, A is the area for heat transfer, V is the reactor volume, ρ is the density, C_{p} is the heat capacity, and ΔH is the heat of reaction. Typical parameters are listed in Table 4.2.

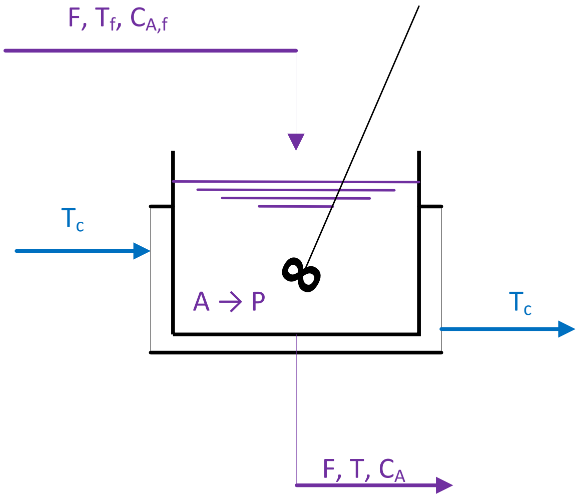


Figure 4.6. Schematic of the exothermic CSTR.

Table 4.2. Parameters for the dynamic CSTR model

Parameter	Value	Unit
A	30	m^2
$C_{A,f}$	2	$kmol/m^3$
C_p	4	kJ/(kg-K)
E_a	1.50E+04	kJ/kmol
k_0	1.7038	1/min
R	8.314	kJ/(kmol-K)
T_c	300	K
T_f	300	K
U	100	kJ/(min-K-m^2)
V	10	m^3
ΔH	2.20E+05	kJ/kmol
ρ	1,000	kg/m^3
τ	0.5	min
K_c	-0.02	min/K
τ_I	0.05	min

This reactor exhibits S-shaped dependences of conversion and temperature on residence time, as shown in Figure 4.7a,b. Consider that it is desired to operate along the high-conversion branch, but at lower temperature, 800K, for safety reasons, with a residence time of 0.5 minutes, having a conversion of 0.5 and a temperature of 800K.

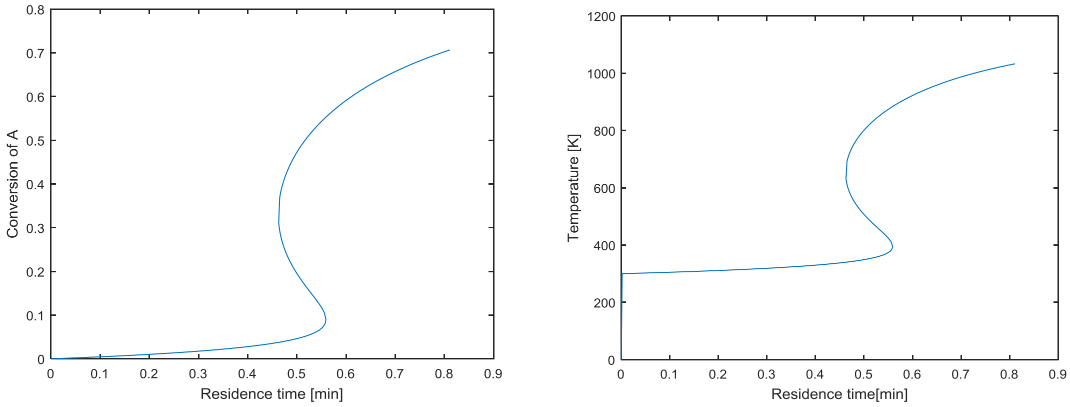


Figure 4.7. Conversion in the exothermic CSTR.

It is important to understand the mechanisms by which the CSTR can move from operation on the high- to the low-conversion branch. From a reliability perspective, this CSTR would likely be shutdown when it moves to the low-conversion branch, especially considering that ignition to the high-conversion branch during operation may pose safety risks (e.g., large overshoot that is difficult to avoid). To move from the high- to low-conversion branch, non-parametric noise is introduced to each of the two balance equations.

$$V\frac{dC_{\mathrm{A}}}{dt} = \frac{V}{\tau}\left(C_{\mathrm{A,f}} - C_{\mathrm{A}}\right) - Vk_0\exp\left(-\frac{E_{\mathrm{a}}}{RT}\right)C_{\mathrm{A}} + \frac{\eta_1 V}{\tau} \tag{4.11a}$$

$$\rho V C_{\mathrm{p}}\frac{dT}{dt} = \frac{\rho V C_{\mathrm{p}}}{\tau}(T_{\mathrm{f}} - T) + UA(T_{\mathrm{c}} - T) + V\Delta H k_0\exp\left(-\frac{E_{\mathrm{a}}}{RT}\right)C_{\mathrm{A}} + \frac{\eta_2 \rho V C_{\mathrm{p}}}{\tau} \tag{4.11b}$$

Each η is sampled every minute ($t_{\mathrm{s}} = 1\ min$) from an independent normal distribution with mean 0 and variance, σ_i^2. The non-parametric noise terms in Eq. (11) must be scaled carefully to predict plant shutdowns and accidents. Each noise term was scaled to represent noise in convective flux – the noise introduced to each equation is of the same order of magnitude as fluctuations in feed material or energy. These terms are scaled to yield either noise in convective flux or unpostulated safety events arising from phenomena having similar magnitudes to convective noise; possibly modest side-reactions or water leaks. When $\sigma_2 = 0$ (i.e., η_2 is sampled from a delta function centered at zero, $\eta_2 = 0$) at all times, with $\sigma_1^2 = 0.02$, the dynamic trajectory over an hour is shown in Figure 4.8a. However, when the variance is increased to $\sigma_1^2 = 0.2$, the system moves to its low-conversion region, as shown in Figure 4.8b. These figures motivate control for the CSTR – a relatively modest noise drives the system to its low-conversion region, and for very small noise, the temperature still fluctuates over a 50K range. Note that parametric noise could have been introduced to a parameter appearing in both balances, such as V, by modifying $V = V_0 + \eta$, where V_0 is the original choice for V (listed in Table 4.2).

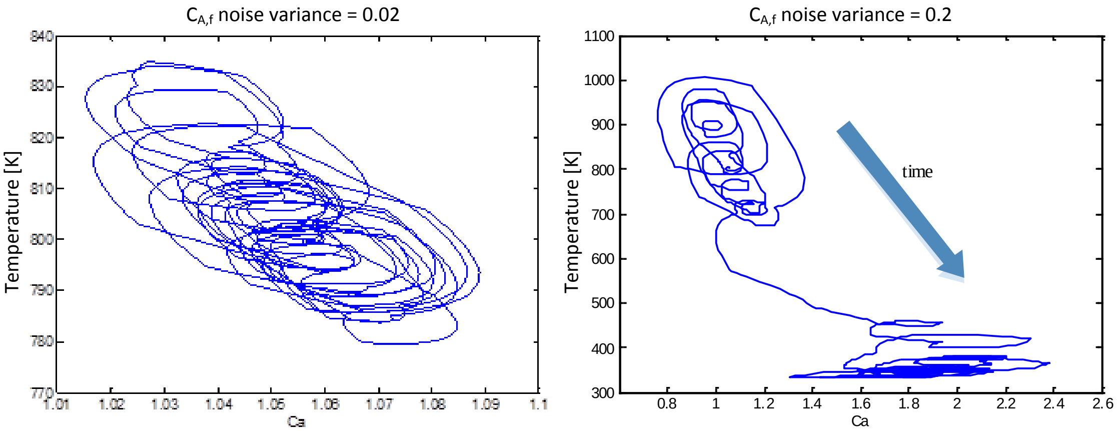


Figure 4.8. Effect of introducing noise to an uncontrolled CSTR.

In typical operation, PI-control is used to maintain the reactor in its high-conversion region. Herein, the residence time is manipulated to maintain the temperature at 800K:

$$V\frac{dC_{\mathrm{A}}}{dt} = \frac{V}{\tau}\left(C_{\mathrm{A,f}} - C_{\mathrm{A}}\right) - Vk_0 \exp\left(-\frac{E_{\mathrm{a}}}{RT}\right)C_{\mathrm{A}} + \frac{\eta_1 V}{\tau} \tag{4.12a}$$

$$\rho V C_{\mathrm{p}}\frac{dT}{dt} = \frac{\rho V C_{\mathrm{p}}}{\tau}(T_{\mathrm{f}} - T) + UA(T_{\mathrm{c}} - T) + V\Delta H k_0 \exp\left(-\frac{E_{\mathrm{a}}}{RT}\right)C_{\mathrm{A}} + \frac{\eta_2 \rho V C_{\mathrm{p}}}{\tau} \tag{4.12b}$$

$$\frac{de_i}{dt} = T_{SP} - T \tag{4.12c}$$

$$\tau = K_c\left(T_{sp} - T + \frac{e_i}{\tau_i}\right) \tag{4.12d}$$

where T_{SP} is the temperature setpoint, K_c is the controller gain, e_i is the integral of the error, and τ_i is the integral time constant, shown in Table 4.2. With control, the noise term has far less impact, as shown in Figure 4.9.

While control makes moving to the low-conversion region far less likely, it remains a plausible rare event with safety implications. In other words, even with control, a noise pattern can move the reactor to its low-conversion region. Using TPS, rare paths from the high- to the low-conversion regions are shown next.

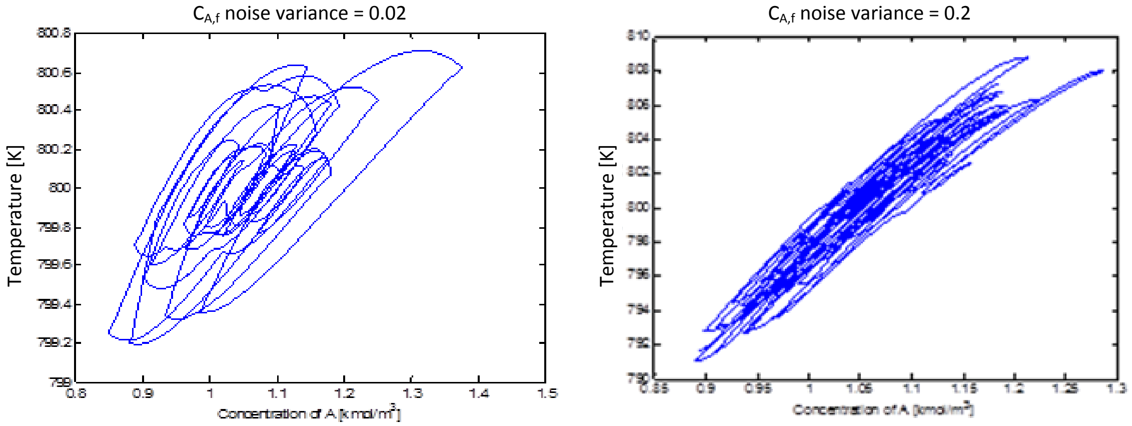


Figure 4.9. Effect of introducing noise to a controlled CSTR.

4.3.1. TPS to Generate Rare-Event Trajectories

As shown in the TPS algorithm in Table 4.1, an initial rare-event trajectory must be generated. Such a low-likelihood trajectory can be generated by prescribing the noise, η_1, to take a high magnitude over the full hour trajectory. An initial trajectory is generated by setting η_1 = -1.5 at each sampling interval (one minute), and setting $\eta_2 = 0$. The trajectory begins at steady-state. As the temperature initially decreases, with less reactant available, the PI-controller increases the residence time, increasing the conversion of the reaction and generating more heat. The trajectory is shown below in Figure 4.10. With the system at a low concentration of A and a temperature below the setpoint (800K), the PI-controller continues to increase the residence time, but the reaction heat release is offset by the cooling jacket. Said differently, at the end state of this trajectory, the PI-controller is incapable of returning the reactor to the desired high-conversion region. Subsequently, the region of rare-safety events is initially demarked by regions in the ranges:

$$790 \leq T_0 \leq 810 \text{ K} \quad ; 0.9 \leq C_{\text{A},0} \leq 1.2 \text{ kmol/m}^3 \qquad (4.13)$$

and final conditions bounded by:

$$T_\text{f} \leq 750 \text{ K} \quad ; \quad C_{\text{A,f}} \leq 0.8 \text{ kmol/m}^3 \qquad (4.14)$$

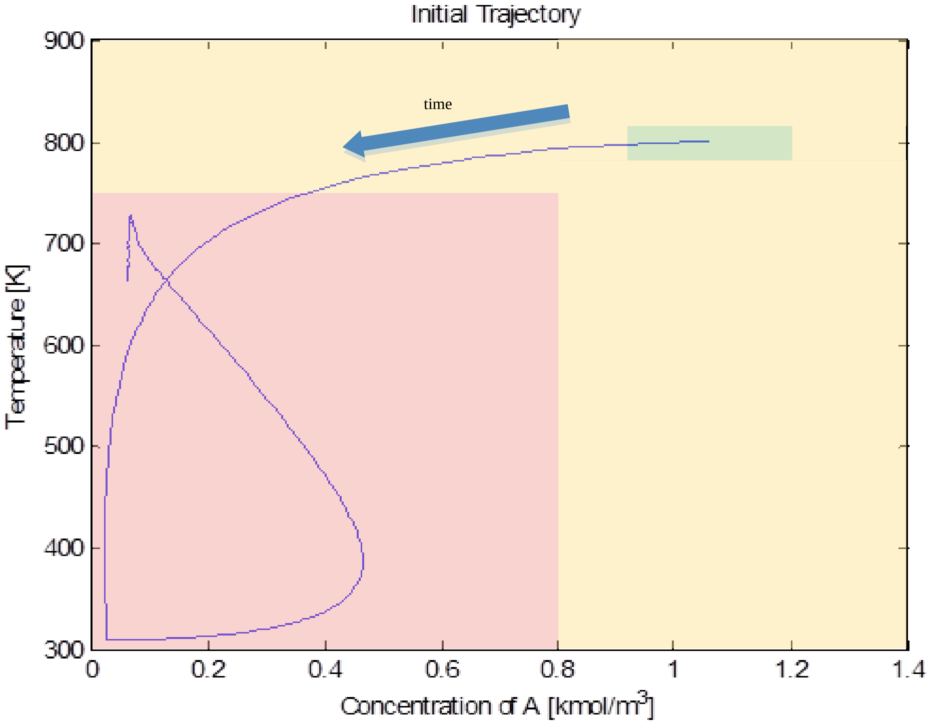


Figure 4.10. Initial rare-event trajectory.

From this initial trajectory, the TPS technique generates more trajectories – specifically, those having higher likelihoods (i.e., less noise). At random time, t', perturbations are made to the state variables, T, C_A, and e_i, and the stochastic variables; η_1, and η_2. These are sampled from normal distributions:

$$T^{'} \sim N(T(t^{'}), 5) \tag{4.15a}$$

$$C_{\mathrm{A}}^{'} \sim N(C_{\mathrm{A}}(t^{'}), 0.05) \tag{4.15b}$$

$$e_i^{'} \sim N(e_i(t^{'}), 5) \tag{4.15c}$$

$$\eta_1^{'} \sim N(0, 0.1) \tag{4.15d}$$

$$\eta_2^{'} \sim N(0, 10) \tag{4.15e}$$

Given sampling distributions for the perturbations, the likelihood distributions $f(x_0)$ and $g\big(\eta(t)\big)$ are needed. A simple, uniform likelihood distribution for the initial conditions is used, with all trajectories that meet the initial rare-event criteria (Eqs. (12) and (13)) equally likely; i.e.,

$$f\big(T_0, C_{A,0}\big) = \frac{1}{(810-790)\times(1.2-0.9)} \tag{4.16}$$

The likelihoods of the noise variables are:

$$g(\eta_1) = N(0, 0.1) \tag{4.17a}$$

$$g(\eta_2) = N(0, 10) \tag{4.17b}$$

With these distributions established, the TPS methodology iteratively generates many random rare-event trajectories. In Figure 4.11, a few such trajectories are illustrated.

In the TPS algorithm, the length of the trajectory, t_{f}, must be specified a priori. Both t_{f} and the sampling interval, t_{s}, for stochastic variables, η_1 and η_2, must be selected carefully. When t_{f} is too long, trajectories reside in either the high- or low-conversion regions too long – rather than moving from region-to-region in transition; i.e., along pathways of interest. Figure 4.12 shows a poor choice for t_{f}, where the reactor moves too quickly to the low-conversion region; i.e., low-temperature region. In this case, t_{f} was set at 24 hr. But, when t_{f} is too small, the low-conversion region is not reached; i.e., no rare-event trajectories are computed.

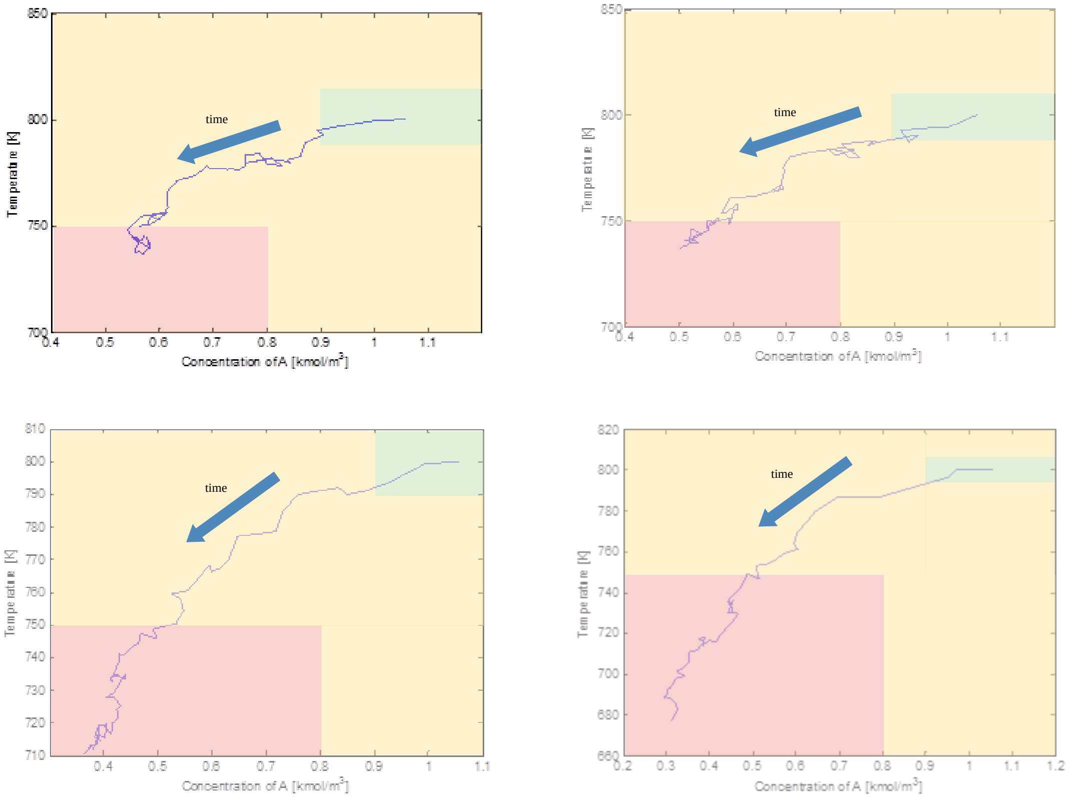


Figure 4.11. Rare-event trajectories generated using TPS.

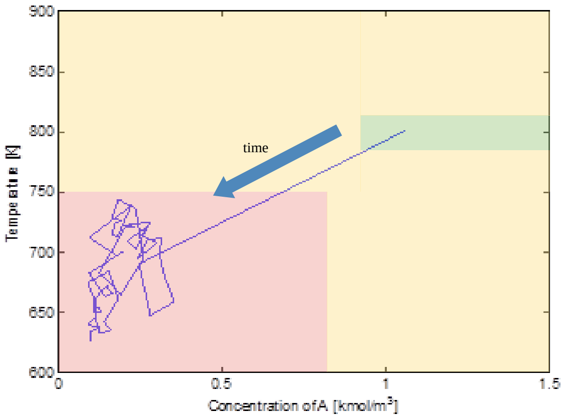


Figure 4.12. Example of a simulation that is too long.

The stochastic sampling time is especially crucial when control is implemented. Here, the PI-controller moves toward a steady-state, at its setpoint, in the face of disturbances. When effective, in response to η_1 and η_2 noise, the reactor *equilibrates* at its desired steady-state. But, after equilibration, backward integration cannot be restarted. Consequently, when responding to stochastic noise, sampling intervals must be smaller than controlled reactor response times; i.e., when sufficiently small noise sampling intervals are used, dynamic behavior is achieved. Note that more effective controllers require shorter noise sampling time intervals. For this example, a sampling time of one minute is sufficiently short.

The TPS strategy continually yields trajectories having probabilities greater than or of similar magnitude to the probabilities of the previous trajectory (Step 6 in Table 4.1). The TPS algorithm was run for N = 100,000. The likelihood probabilities, p, of the first 350 unique trajectories are shown in Figure 4.13, in sequence. From the least-likely initial trajectory, the sampling strategy moves towards more likely trajectories. These probabilities, of course, remain small as they represent the most likely of rare safety-events – and the first 150 are rejected as atypical.

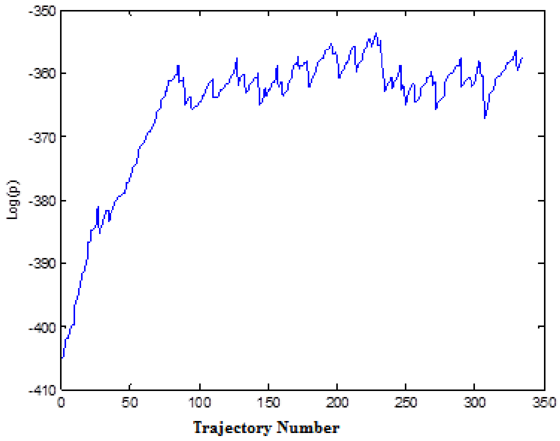


Figure 4.13. First 350 TPS trajectories.

As the trajectories are generated, it is desirable to cluster them to better understand the transitions between the high- and low-conversion states. In Figure 4.14, the time-averaged noise, $\widehat{\eta_1}$, is displayed as a function of the time-averaged noise, $\widehat{\eta_2}$, noting that each data point represents a single trajectory, with each time-averaged noise:

$$\hat{\eta} = \frac{1}{t_f}\int_0^{t_f} \eta(t)dt \tag{4.18}$$

As shown, there are two distinct clusters of trajectories, A and B, which are identified using the k-means clustering technique (Hartigan et al., 1979). Here, each trajectory is clustered about one of the centroids, the centers-of-mass of the clusters. Using this strategy, two distinct paths to the safety-event have been identified. While other clusters may not have been identified, these two clusters of likely trajectories are excellent candidates for protection in the form of reliability and safety systems.

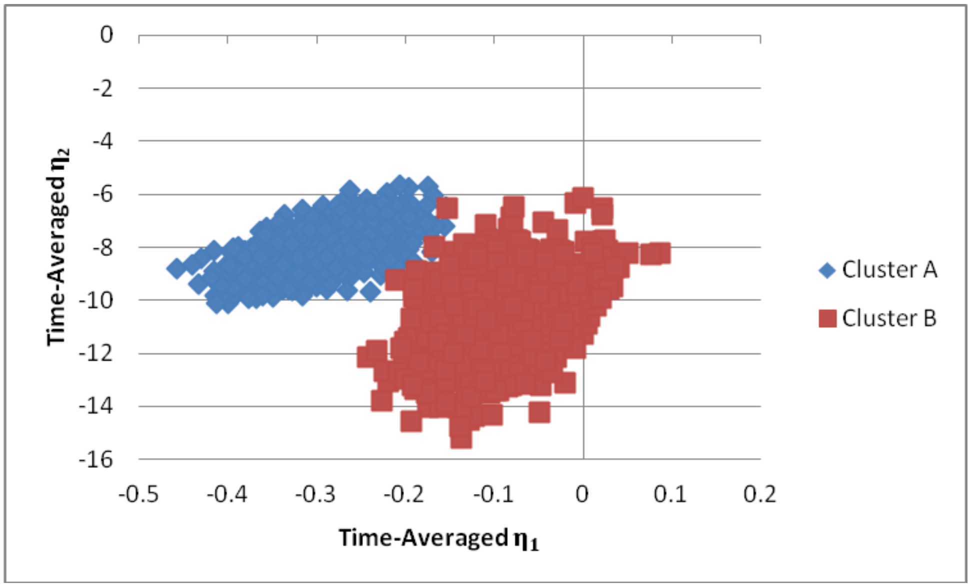


Figure 4.14. The trajectories displayed in two clusters.

Figure 4.15 shows the natural log of trajectory likelihood for 100,000 unique trajectories, with values displayed for trajectories 201, 301, 501, …, in sequence. Note: the first 200 trajectories are atypical and not displayed. Initially, Cluster A is populated, involving data for approximately 20,000 trajectories, having likelihoods near $\ln(p) = -355$. Eventually, the trajectories move towards the bridge between the two clusters, with those in Cluster B displayed, having likelihoods near $\ln(p) = -340$. When the TPS algorithm samples from within a cluster, it tends to yield trajectories within the cluster. However, some perturbations allow the algorithm to move from one cluster to the other. The movements between clusters are possible when sampling occurs from sections close to the other cluster – permitting transfers to occur with just small perturbations. In fact, a third Cluster C may exist, which is not populated because it lies too far from Clusters A and B.

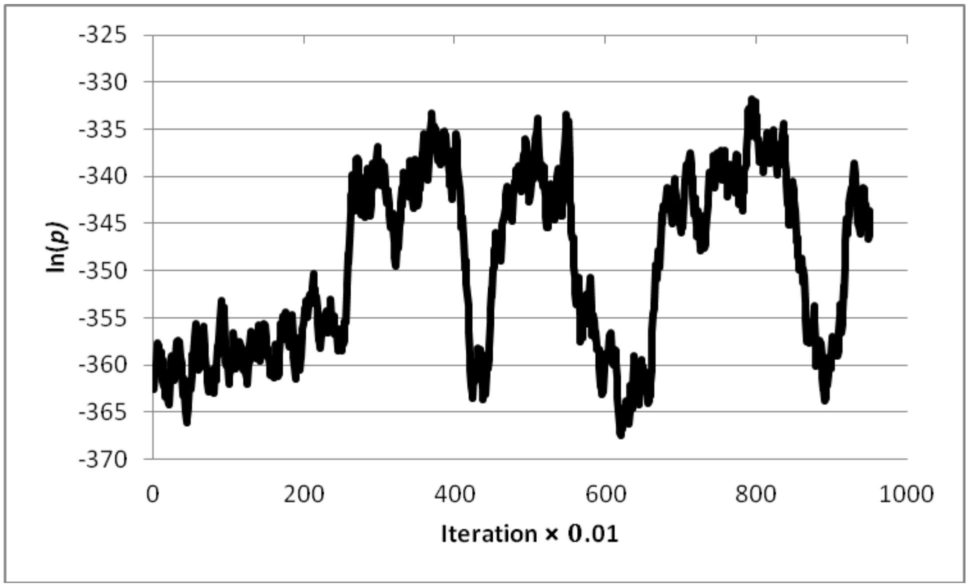


Figure 4.15. Trajectory likelihood in sequence.

When larger perturbations are allowed, the movement between clusters is more frequent – occurring across larger distances. Let the variances of the perturbations include a factor θ:

$$T' \sim N(T(t'), 5\theta) \tag{4.19a}$$

$$C_{\mathrm{A}}' \sim N(C_{\mathrm{A}}(t'), 0.05\theta) \tag{4.19b}$$

$$e_i' \sim N(e_i(t'), 5\theta) \tag{4.19c}$$

$$\eta_1' \sim N(0, 0.1\theta) \tag{4.19d}$$

$$\eta_2' \sim N(0, 10\theta) \tag{4.19e}$$

When the TPS algorithm is run, starting with the same initial trajectory, the number of movements between the clusters are shown as a function of θ in Figure 4.16. The greater the perturbation variance, the more movements between clusters. It should be noted that even when $\theta = 2$, only Clusters A and B were populated (i.e., Cluster C was not discovered). But, even larger θ may yield Cluster C – with significantly longer computation times. Alternatively, starting from a number of very different initial trajectories may be a fruitful avenue for discovering new trajectory clusters, because with larger perturbations, the trajectories are less likely to satisfy the rare safety-event criteria (i.e., $h_A(0) \times h_B\left(t_f\right) = 0$), or the acceptance criteria. Figure 4.17 shows that as θ increases, the probabilities of acceptance decrease significantly. While $\theta = 2$ allows for 25 movements between Clusters A and B, on the order of 1,000,000 trial trajectories are generated to capture 100,000 unique trajectories.

Clearly, the number of trial trajectories to yield a new, uncorrelated trajectory is sensitive to the choice of θ. While small θ yields a high probability of acceptance, as shown in Figure 4.17, the accepted trajectories are quite similar to the original trajectories – and many trajectories must be accepted before a new, *uncorrelated* trajectory is generated.

The autocorrelation function quantifies the correlation between trajectories κ iterations apart. As follows, a κ_{crit} is determined to locate sufficiently different trajectories. First, a midpoint for the respective trajectories is selected. For the CSTR process, within the green-belt and red-belt zones, temperatures and concentrations differ significantly, whereas in the yellow-belt zone, temperatures and concentrations follow similar paths. A midpoint is selected at $T_m = 770$K [(750 + 790)/2] – the temperature midpoint between the high and low conversion regions (see Eqs. (12, 13)). Note that for a trajectory i, the $i+1$ trajectory is similar at its midpoint, the $i+2$ trajectory is less similar, and so on. As the iteration distance κ increases, the midpoint of trajectory $i+\kappa$ becomes less similar to the midpoint of trajectory i. The autocorrelation function averages the correlation between the midpoints of trajectory i and $i+\kappa$ over all trajectories (Gelman et al., 2014). The autocorrelation function is:

$$R(\kappa) = \frac{E\left[\left(C_{A,m}^{i}-\mu_{m}^{i}\right)\left(C_{A,m}^{i+\kappa}-\mu_{m}^{i}\right)\right]}{{\sigma_{m}^{i}}^{2}} \tag{4.20}$$

where $R(\kappa)$ is the autocorrelation value at κ iteration distance, E is the expected value function, $C_{A,m}^{i}$ is the concentration of trajectory i at T_m, μ_m^i and ${\sigma_m^i}^2$ are the mean and variance of the midpoint concentration amongst trajectories in the cluster associated with iteration i, respectively. A critical iteration distance, κ_{crit}, is defined as the smallest κ for which

$$R(\kappa_{\text{crit}}) < 0.1 \tag{4.21}$$

noting that the autocorrelation function has the properties:

$$R(0) = 1 \;\; ; \;\; \lim_{A\to\infty} R(A) = 0 \tag{4.22}$$

Figure 4.18 shows κ_{crit} as a function of θ noting that the minimum, $\theta = 1.0$ is the most efficient computational choice.

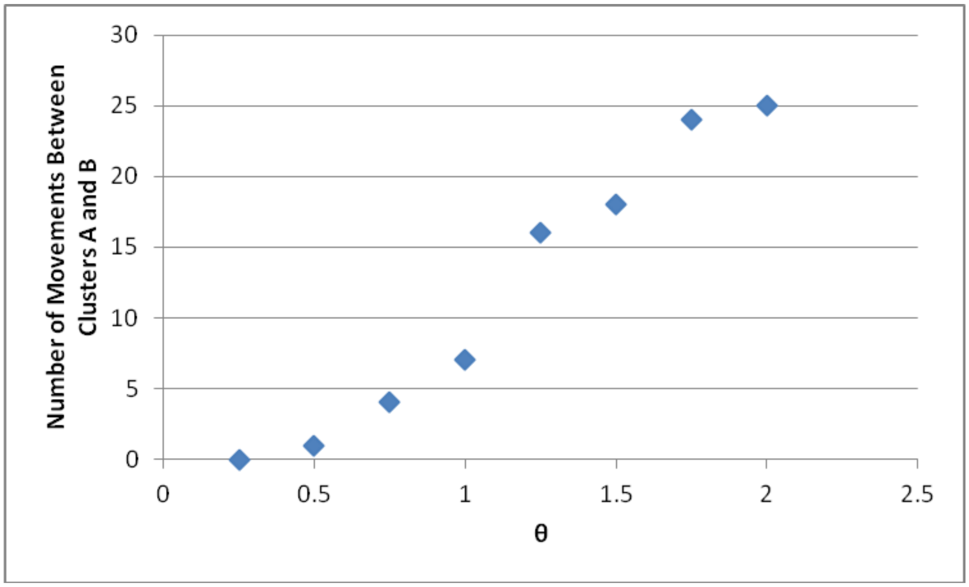


Figure 4.16. Number of movements between clusters as a function of perturbation size.

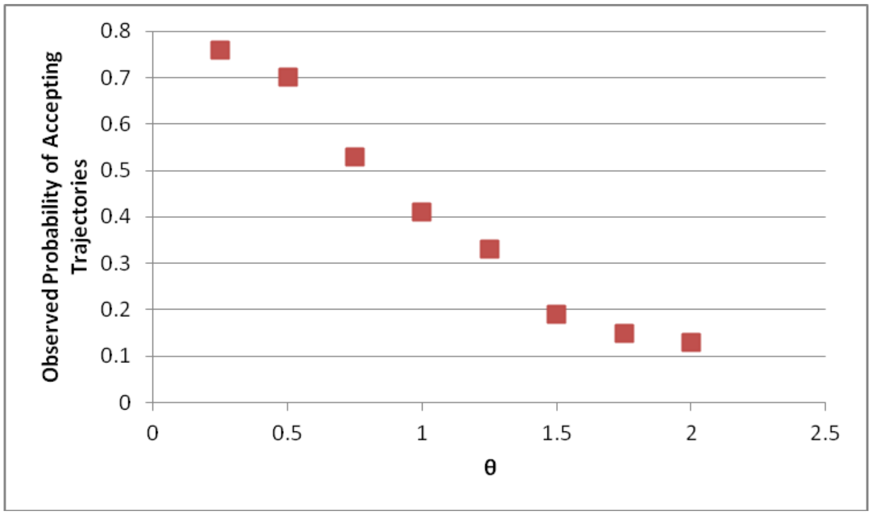


Figure 4.17. Probability of accepting trajectories as a function of perturbation size.

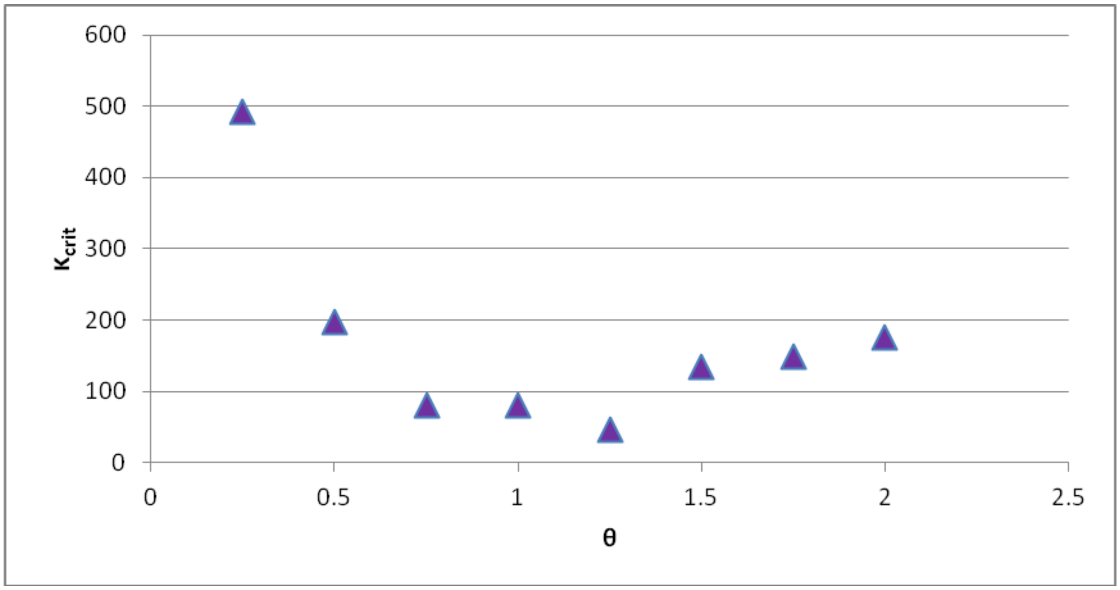


Figure 4.18. κ_{crit} **as a function of** $\boldsymbol{\theta}$**.**

With the two clusters identified, an opportunity exists to better understand each cluster. In Figure 4.19, the concentration of A is displayed as a function of temperature for 1,000 random trajectories associated with Cluster B. All trajectories fit the same pattern, which can be helpful in creating an alarm. Movement in this pattern suggests that a reliability or safety-event is impending. Such an alarm could alert operators in time to prevent such an event. This is a very powerful use for TPS – reliability and safety systems can be aided by the quantitative simulation analysis to mitigate rare safety events more frequently.

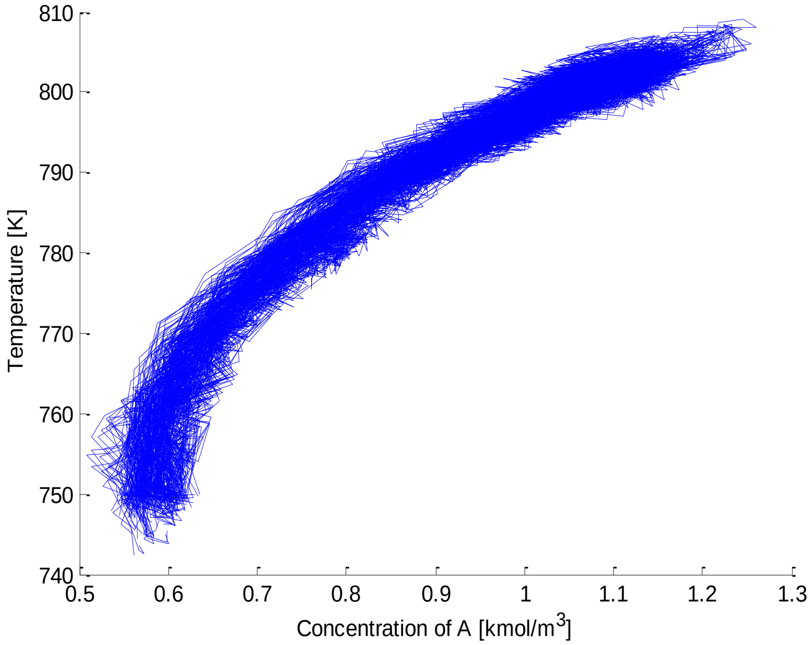


Figure 4.19. Concentration of A as function of temperature for all trajectories in Cluster B.

4.4. Air Separation Unit (ASU) Example

The exothermic CSTR example involves just two state variables in a simple, familiar model to demonstrate the TPS method and its associated challenges and opportunities. But, industrial application of TPS is likely to involve significantly more complex processes. In this example, TPS is applied to an ASU model, having 480 state variables. This model uses a modified version of a process flow diagram proposed by the NETL (“Commercial Technologies for Oxygen Production”), and uses mathematical formulae proposed by Huang et al. (Huang et al., 2009) Because the process operates at cryogenic conditions, it has many heat recycle loops. These loops create complex process interactions, some of which may propagate to rare safety and reliability events. The power of TPS is demonstrated in this example – with events that occur due to process-scale interactions captured and better understood, as compared with events that arise just due to unit operation disturbances and failures.

Pretreated air (stream 1) is separated into product liquid oxygen (LOX – stream 23), gaseous oxygen (GOX – stream 22), liquid nitrogen (LIN – stream 12), gaseous nitrogen (GAN – stream 14), and liquid argon (LAR – stream 20). Its pretreatment involves removing water, carbon dioxide, methane, and cooling to saturated vapor and saturated liquid feed streams with only oxygen, argon, and nitrogen present (in order of increasing volatility). In this example, three distillation columns are used – a high-pressure column (HPC) at 5.5 bar beneath a low-pressure column (LPC) at 1.25 bar, with a crude argon column (CAR) taking a sidedraw (stream 18) from the LPC, as shown in Figure 4.20. At these pressures, the columns operate cryogenically at temperatures on the order of 85K. The HPC vapor overhead (stream 8) is condensed by vaporizing the LPC bottoms liquid, and the CAR vapor overhead (stream 18) is condensed by vaporizing the HPC bottoms liquid (stream 5, cooled to stream 6). As streams leave the HPC and expand from 5.5 bar to 1.25 bar, they are cooled by GAN (stream 14) and waste nitrogen (stream 15) to maintain saturated liquids at the same temperature. The HPC and CAR has 40 trays, and the LPC has 80 trays. The feed and side-draw locations were chosen to provide products having impurities <1 mol%, as specified in the Huang et al. (Huang et al., 2009) ASU model.

The trays in each column are modeled at equilibrium using the MESH (m̲ass balance, phase e̲quilibrium, s̲ummation of mole fractions, and h̲eat balance) equations, an empirical equation relating the liquid holdup to the liquid flow rate, and the Peng-Robinson equation of state[17]. This model is modified from the Huang et al. (Huang et al., 2009) for model-predictive control. The overall material balance of tray i is represented by:

$$\frac{d}{dt}(M_i) = L_{i-1} + V_{i+1} - L_i - V_i + F_i \qquad (4.23)$$

where M_i is the liquid molar holdup, $\mathbb{L}$ is the liquid molar flow rate, $\mathbb{V}$ is the vapor molar flow rate, and F_i is the molar feed flow rate (zero for most trays). It should be noted that only the liquid holdup is considered here – the vapor, being far less dense, has a negligible holdup. The component material balances describe the composition of each tray, and are shown in:

$$\frac{d}{dt}\left(x_{i,j}M_i\right) = x_{i-1,j}L_{i-1} + y_{i+1,j}V_{i+1} - x_{i,j}L_i - y_{i,j}V_i + z_{i,j}F_i \ \forall\ j \in \{N_2, Ar, O_2\} \quad (4.24)$$

where x is the liquid molar fraction, y is the vapor molar fraction, and z is the feed molar fraction (phase unspecified). The heat balance is:

$$\frac{d}{dt}(M_i h_i^L) = h_{i-1}^L L_{i-1} + h_{i+1}^V V_{i+1} - h_i^L L_i - h_i^V V_i + h_i^F F_i \qquad (25)$$

where h^L is the liquid molar enthalpy, h^V is the vapor molar enthalpy, and h^F is the feed molar enthalpy (phase unspecified). An empirical relationship for the liquid flow rate as a function of the liquid molar holdup is assumed:

$$L_i = k_d M_i \tag{4.26}$$

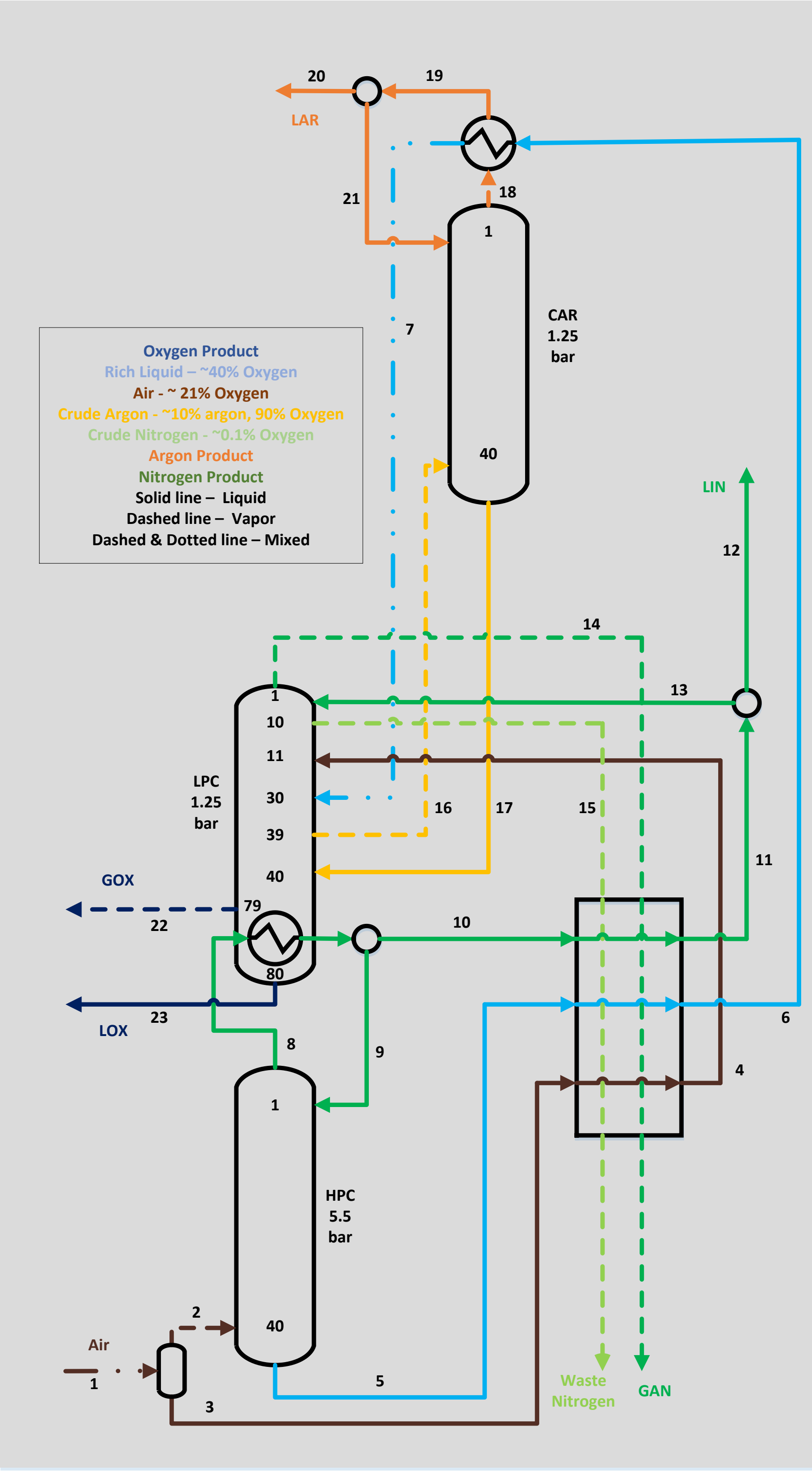


Figure 4.20. Air Separation Unit process flow diagram.

where k_d was specified by Huang et al. (Huang et al., 2009) at 0.5 min^{-1}. The vapor-liquid equilibrium in each tray is modeled by equating the mixture fugacities of the species in the vapor and liquid phases:

$$\phi_{i,j}^{L} x_{i,j} = \phi_{i,j}^{V} y_{i,j} \tag{4.27}$$

where $\phi_{i,j}^{L}$ is the liquid fugacity coefficient of species *j* on tray i, and $\phi_{i,j}^{V}$ is the vapor fugacity coefficient of species *j* on tray *i*, each calculated using the Peng-Robinson equation of state, at the tray temperature, T_i, the column pressure, and the mole fractions of the associated phases. The enthalpies are also calculated using the equation of state as described by Peng and Robinson (Peng et al., 1976). Lastly, the mole fractions sum to unity in the vapor on each tray:

$$1 = \sum_j y_{i,j} \tag{4.28}$$

The feed air is assumed to have a constant composition (78% nitrogen, 21% oxygen, 1% argon). Its flow rate, F_{air}, is determined by the product demand of each product i, d_{i}:

$$F_{air} = \max\left(\frac{d_{\mathrm{LIN}}+d_{\mathrm{GAN}}}{0.78}, \frac{d_{\mathrm{LOX}}+d_{\mathrm{GOX}}}{0.21}, \frac{d_{\mathrm{LAR}}}{0.01}\right) \tag{4.29}$$

There are six more operation decision variables, β_1 through β_6: the LIN split fraction fed to the subcooler, the LIN split fraction fed to the LPC, the LAR split fraction to the CAR reflux, the side-draw split fraction of waste nitrogen, the side-draw fraction of crude argon, and the side-draw fraction of GOX. These decision variables are manipulated using PID controllers to maintain the six set points, α_1 through α_6: N_2 mole fraction of GAN, N_2 mole fraction of LIN, O_2 mole fraction of GOX, O_2 mole fraction of LAR, ratio of GAN to LIN, and ratio of GOX to LOX. Note, the O_2 mole fraction of LAR is controlled, rather than the Ar mole fraction, because a fourth column normally handles the Ar-N_2 separation in industry – not included herein to reduce the computational load. The set points and α-β pairs are shown in Table 4.3.

4.4.1. TPS Process-Scale Demonstration

To demonstrate TPS, at least one rare-event must be identified. The rare events of interest for this example involve the nitrogen mole fraction in the LAR stream 20, $x_{\mathrm{N_2}}^{\mathrm{LAR}}$. In this model, the LAR typically contains on the order of 0.001 mole fraction nitrogen. As its nitrogen content increases, its condensation load increases – with the rich liquid oxygen stream (6) increasingly vaporized. The effluent liquid oxygen (stream 7), a feed to the LPC (on tray 30 – typically the largest feed stream), is crucial to the operation of the column, and as it becomes increasingly vaporized, α_3, d_{GOX}, and d_{LOX} may not be met. Additionally, as more nitrogen is introduced to the LAR, a greater burden is placed on the column responsible for the N_2-Ar separation before the argon product can be sold (column not modeled). When the nitrogen fraction in the LAR exceeds a critical level, the process typically undergoes a shutdown (reliability-event), which has the potential to become a safety-event (as shutdown and restart can provide more challenges). Therefore, the rare-event is defined by:

$$0.002 > x_{\mathrm{N_2}}^{\mathrm{LAR}}(0) \text{ and } 0.01 < x_{\mathrm{N_2}}^{\mathrm{LAR}}(t_\mathrm{f}) \tag{4.30}$$

Table 4.3. Control logic of the ASU Model.

Controlled Variable	**Manipulated Variable**	**Set point**
α_1 (N_2 mole fraction of GAN)	β_4 (side-draw fraction of waste nitrogen)	0.995
α_2 (N_2 mole fraction of LIN)	β_1 (LIN split fraction to subcooler)	0.995
α_3 (O_2 mole fraction of GOX)	β_5 (side-draw fraction of crude argon)	0.985
α_4 (Ar mole fraction of LAR)	β_3 (LAR split fraction to CAR)	0.99
α_5 (ratio of GAN to LIN)	β_2 (LIN split fraction to LPC)	$d_{\mathrm{GAN}}/d_{\mathrm{LIN}}$
α_6 (ratio of GOX to LOX)	β_6 (side-draw fraction of GOX)	$d_{\mathrm{GOX}}/d_{\mathrm{LOX}}$

An initial rare-event trajectory begins with little nitrogen in the CAR column – thus, there is little nitrogen entering the column from the LPC in stream 16. The vapor mole fraction profile along the LPC is shown in Figure 4.21a at the start of this trajectory – notice that the argon composition peaks close to tray 39, the tray whose vapor sidedraw is fed to the CAR column, while the nitrogen fraction is sufficiently low. At the beginning of this trajectory, overly aggressive set points (which are described by the set point likelihood, described later) for the production rate of LOX and LAR are simulated. With large amounts of vapor from the bottom of the LPC column, liquid higher in the column vaporizes to replace it, and the argon bubble begins dropping along the column. The mole fraction profiles at 1-, 4-, and 12-hr are shown in Figures 4.21b, 4.21c, and 4.21d, respectively. As the nitrogen enters the CAR, it rises into the LAR product, whose nitrogen mole fraction along the trajectory is shown in Figure 4.21e. This trajectory, which satisfied the criteria in Eq. (30), represents a rare-event trajectory.

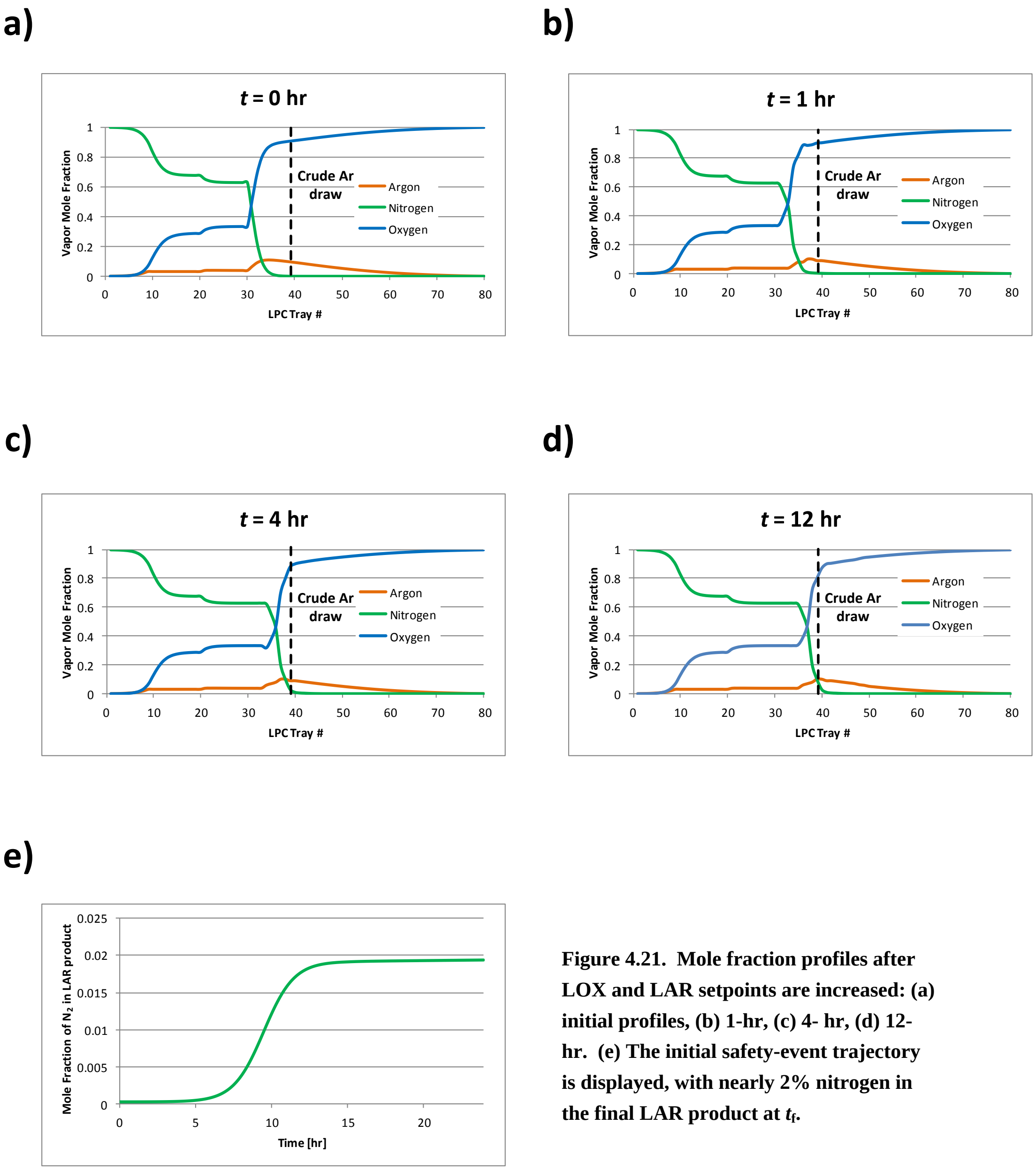


Figure 4.21. Mole fraction profiles after LOX and LAR setpoints are increased: (a) initial profiles, (b) 1-hr, (c) 4- hr, (d) 12-hr. (e) The initial safety-event trajectory is displayed, with nearly 2% nitrogen in the final LAR product at t_f.

From this initial trajectory, perturbations are made to each state variable, $M_i(t')$, $x_{i,\mathrm{N}_2}(t')$, and $x_{i,\mathrm{O}_2}(t')$. The perturbations to liquid mole fractions are particularly challenging – in regions where species j mole fractions, $x_{i,j}$, are near zero or unity, only small perturbations can be handled (that is, large perturbations may not allow the system of equations to be solved). However, in regions where $x_{i,j}$ are near 0.5, much larger perturbations can be handled. Therefore, the variances of the perturbations are scaled by factors $x_{i,j}$(1-$x_{i,j}$). The stochastic variables in the model are the demand of each product, d_{i}, sampled at each half-hour (with 48 samples over the day-long trajectory). The perturbations are sampled from normal distributions with mean and variance in parentheses:

$$M'_i(t^{'}) \sim N(M_i(t^{'}), 100\ \mathrm{kmol}^2) \tag{4.31a}$$

$$x'_{i,\mathrm{N}_2}(t^{'}) \sim N\left(x_{i,\mathrm{N}_2}(t^{'}), x_{i,\mathrm{N}_2}(t^{'}) \times \left(1 - \ x_{i,\mathrm{N}_2}(t^{'})\right) \times 10^{-4}\right) \tag{4.31b}$$

$$x'_{i,\mathrm{O}_2}(t^{'}) \sim N\left(x_{i,\mathrm{O}_2}(t^{'}), x_{i,\mathrm{N}_2}(t^{'}) \times \left(1 - \ x_{i,\mathrm{N}_2}(t^{'})\right) \times 10^{-4}\right) \tag{4.31c}$$

$$d'_i(t^{'}) \sim N(d_i(t^{'}), 1{,}000\ \mathrm{kmol}^2) \tag{4.31d}$$

The other important user-defined function is the likelihood of each trajectory, *p*. The density of initial conditions, $f_0(x_0)$, is calculated using computational data, collected from a 100-day run of the system with d_i sampled at 30 minute intervals:

$$d_i^{'}(t^{'}) \sim N(\overline{d_\iota}, 100\ kmol^2) \tag{4.32}$$

where $\overline{d_\iota}$ are shown in Table 4.4. A multivariate normal distribution, having 480 dimensions (one for each state variable) is constructed from this data, and is used as f_0. Two dimensions (oxygen and argon purity of the LPC sump and CAR condenser, respectively) of this distribution are shown in Figure 4.22.

Table 4.4. Product Demand for this Simulated Example.

Product	Demand Load [kmol/day]
GAN	20,000
LIN	20,000
GOX	10,000
LOX	10,000
LAR	500

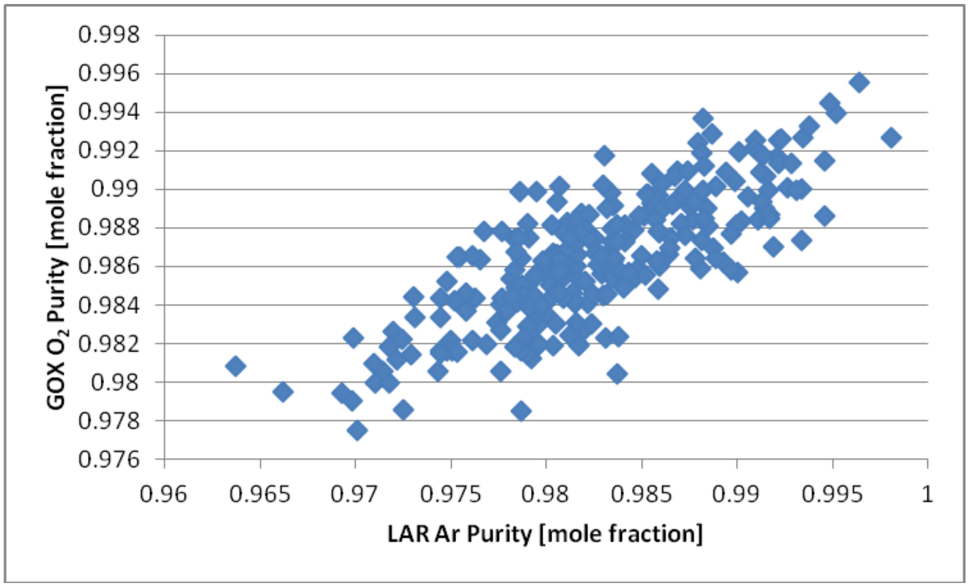


Figure 4.22. Initial condition simulated data.

The likelihood of the product demand at each sampling interval is calculated by:

$$g\big(d_i(t_i)\big) = N(\overline{d_\imath}, 1{,}000\ \mathrm{kmol}^2) \tag{4.33}$$

Having defined all factors in Eq. (6), the likelihoods of the trajectories for this example are calculated.

With an initial trajectory, a sampling distribution for perturbing the state variables at *t*', and a trajectory likelihood function, the TPS algorithm is run to investigate the rare-event. Because the boundary-value problem and integration of the model equations are more computationally taxing than for the CSTR example, *N* is reduced to 5,000. When this algorithm is executed, the trajectories move into one of two clusters. Said differently, two clusters (at least) of trajectories are identified, but trajectories are not observed to move between the clusters as in the CSTR example. The clusters are shown in Figure 4.23, with trajectories 50, 100, 150, …, 5,000, displayed. Here, cluster A contains trajectories that have a high average demand for LOX, whereas Cluster B is occupied by trajectories having a high average demand for LAR. For the trajectories in Cluster A, as LOX is withdrawn from the LPC, more N_2 is drawn into its lower trays and the waste nitrogen withdrawn is reduced Consequently, the crude argon sidedraw (stream 16) becomes increasingly concentrated in nitrogen, and as nitrogen enters the CAR column, the rare-event is realized. For the trajectories in Cluster B, a similar effect occurs where more material from the top of the column (typically N_2 rich) is drawn to the crude argon sidedraw. With more LAR production, the rich liquid stream has a higher condensing duty, with the rich liquid stream increasingly vaporized, resulting in more oxygen leaving through the waste nitrogen stream. Liquid nitrogen, from the LIN reflux, drops into the middle trays in the LPC to replace liquid oxygen, the crude argon stream becomes increasingly concentrated with nitrogen, and the rare-event trajectory is realized.

Clearly, rare-events occur when either LOX or LAR contain overdrawn nitrogen. The two nitrogen accumulations do not occur together in the rare events observed. Also, as shown in Figure 4.24, the sequence of TPS trajectories does not show movement between the two clusters in Figure 4.23 as the trajectories are generated.

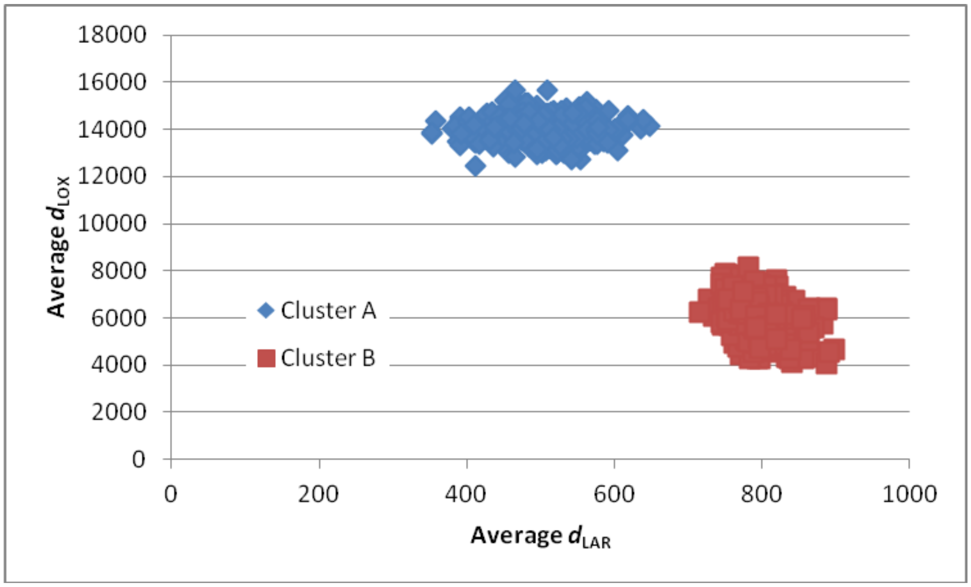


Figure 4.23. Clusters of rare-event trajectories.

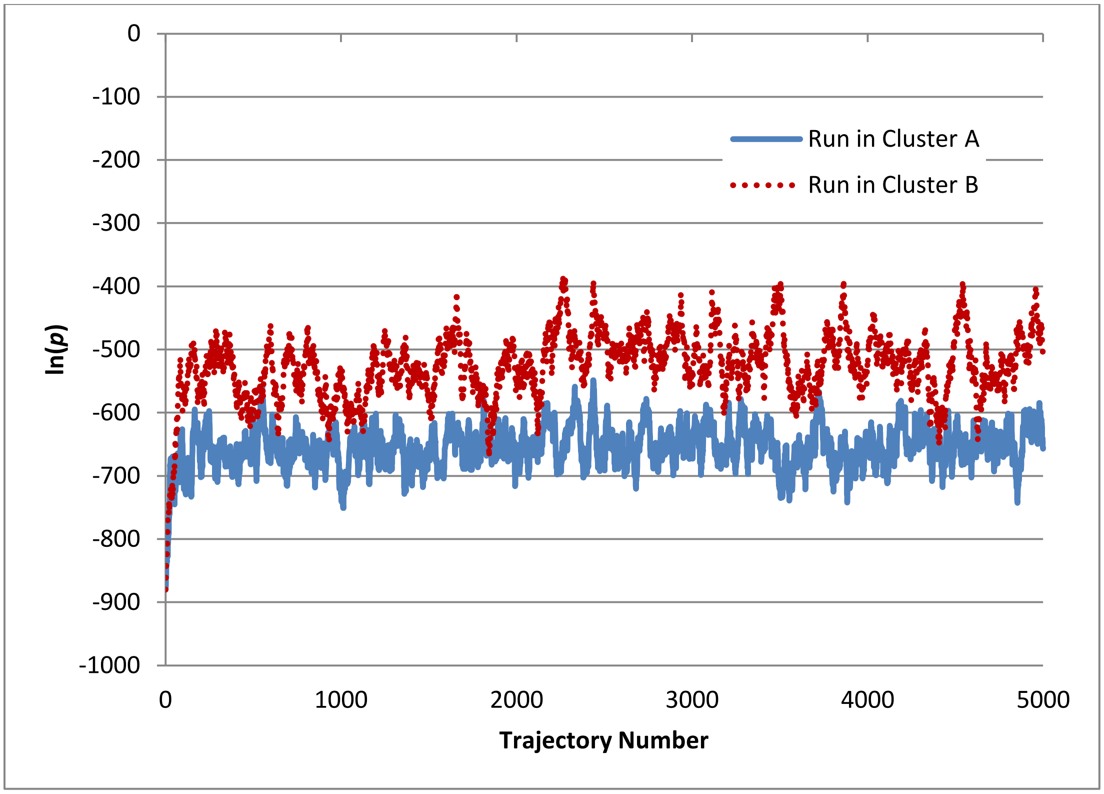


Figure 4.24. Likelihood of rare-event trajectories in succession.

4.5. Conclusions

While rare molecular dynamics events have been studied using TPS methods, this paper extends the techniques to apply to process dynamics in the study of rare product reliability and safety events. For process dynamics, a boundary-value problem is solved in lieu of performing backwards integration, and the likelihoods of trajectories are formulated. Two boundary-value solution methods, shooting (Bock et al., 2000) and orthogonal collocation (Cuthrell et al., 1989), are investigated. While both are sufficient for the CSTR process, the orthogonal collocation method is much better suited to handle the larger ASU process model. Other sampling algorithms, such as forward-flux sampling (Escobedo et al., 2009) or milestoning (Kuczera et al., 2009), do not require backwards integration, and may prove to be effective in understanding rare-event trajectories as well. The likelihood distributions are formed using simulated data, with the incorporation of process data to be investigated in the future. For the exothermic CSTR, two clusters of trajectories are generated by the TPS technique. For an ASU process model, having far more variables and process interactions, two separate clusters of trajectories are generated. In both examples, the discovery of two clusters was not expected – demonstrating that TPS can yield unanticipated rare-event trajectories. Possibly most interesting is the separation of clusters when applied to the ASU process. Because the clusters are sufficiently far apart, this indicates that in the operation of this ASU model, the LOX and LAR demand rate changes can be considered separately. Said differently, reaching the upper limit of acceptable LOX draw should not influence the upper limit of acceptable LAR draw. This sampling strategy benefits from the randomness in state variable perturbations and trajectory acceptances, allowing clusters of rare-event trajectories to be better understood and for the potential discovery of unanticipated trajectories.

Chapter 5
Conclusions and Future Work

5.1 Summary

This thesis has presented two methods, the generation of informed prior distributions (IPD) and transition path sampling (TPS), for predicting the failure probabilities of rarely activated alarm and safety systems. These are difficult to estimate using classical statistical approaches. Commonly, an alarm or safety system is activated just a handful of times, on the order of one to ten, over the lifetime of a process, yielding confidence intervals too large to allow meaningful design or operational decisions. Research on dynamic risk analysis, using copulas, reduced the variance of their predictions (Meel et al., 2006; Pariyani et al., 2012b). However, even using the most advanced statistical techniques, their variances depend upon the amount of data collected. Alarm and safety systems are vital to the proper operation of a chemical process, and meaningful estimates of their failure probabilities are extremely useful, even in the design and commissioning phase (when few data are available), or over the lifetime of the process despite relatively few data points where alarm and safety interlock systems activate.

Chapter 2 describes the IPD technique, which estimates failure probabilities of alarms rarely activated. In Chapter 3, operator behavior models are introduced, enhancing the predictions calculated using IPDs. Large amounts of L- and H-alarm data, resulting from more frequent, less severe, special-cause events, are systematically applied to improve the predictions resulting from less frequent, more severe special-cause events.

When applied to a SMR process, this methodology is shown to provide more reliable failure probability estimates.

The second method, TPS, developed initially by the molecular simulation community, has been modified and applied to study rare process dynamic events. Similar to IPDs, the probabilities of "trajectories" leading to alarm and safety system failures are estimated. With this method, many trajectories are calculated using random perturbations, with statistical weight given to the most likely trajectories. With a fuller understanding of the trajectories that lead to alarm and safety system activations and failures, using both methods, operators and plant managers can better protect processes from transitioning toward unsafe operating conditions.

5.2 A Systematic Approach for Simulation-Based Safety Analysis

While IPDs and TPS can be used individually to improve failure probability predictions, acting together, they can provide a fuller understanding of rare safety events. Each method relies on dynamic process models, which can be cumbersome to construct. Usually, beginning with steady-state models, appropriate dynamic terms are added, then controllers, and eventually alarm and safety systems, are modeled. Alarm and safety systems often involve operator decisions – usually involving stochastic modeling. The construction of these models often requires substantial time and effort. However, modeling for various design and control decisions is normally carried out in the chemical manufacturing industries, so these models are often not built solely for alarm and safety system analyses. With dynamic process and operator models, IPDs and TPS can be applied, both during the design and commissioning phases.

Beyond the dynamic modeling in generating IPDs, TPS provides a framework for generating more likely paths leading to alarm and safety system actions. Random perturbations allow for various trial trajectories to be calculated, and then accepted or rejected based upon their likelihood. This approach is structured to encourage potentially un-postulated trajectories to be 'discovered'. Its results may aid process managers during the design phase, where HAZOP is performed to assess potential process accidents. HAZOP analysis focuses on individual process units, the chemical compounds that may enter the unit, and the possible failures encountered by the unit. Without accounting for more complex process interactions (such as in material or heat recycle), HAZOP does not identify the most probable special-cause event trajectories on the process scale. Said differently, as the fluctuations within a process unit influence all downstream units, failure probabilities of similar units vary amongst different processes. TPS is well-suited to quantify the paths leading to special-cause events, possibly terminating with safety interlock shutdown or an accident. This method has the potential to calculate path trajectories that are either not postulated during HAZOP analysis, or events determined to be of far less significance than may be envisioned in the specific process design.

Even when TPS does not uncover un-postulated special-cause events, it often extends our understanding of postulated events. TPS helps plant managers identify operational conditions that render events more dangerous – possibly during a demand rate shift, or in the presence of another process disturbance. With process interactions leading to special-cause events well understood, the failure of alarm and safety-interlock systems may be prevented through safer operations that avoid their activation.

The TPS technique calculates many trajectories of special-cause events, and their associated probabilities. Using this information, probability distributions of special-cause events can be constructed. Note that IPDs in Chapters 2 and 3 are generated to estimate the failure probabilities of alarm and safety systems from *particular* special-cause events. The special-cause events, and the distributions of special-cause event magnitudes, can be quite challenging to quantify. With few such events, few data are available to directly calculate these distributions. Rather, TPS can be used to generate distributions of the most probable special-cause events along with their magnitudes, which can be input to generate the IPDs for the failure probabilities of alarm and safety systems – even those that involve complex human factors. The synergy of these two methods is quite powerful, even with few data (or no data, in the process design phase), permitting the probabilities of special-cause events and the associated failure probabilities of alarm and safety systems to be better estimated.

While TPS calculates the trajectory probabilities leading to alarm and safety system failures, the IPD is better suited for this purpose. The backwards integration feature required in TPS, accomplished through the solution of a boundary-value problem, does not permit stochastic operator response times to be modeled. In the simulations to obtain IPDs, operator response times are calculated upon the activation of alarm thresholds, with response times a function of the derivatives of variables as they cross their thresholds, and the number of other active alarms in the process. Other factors were investigated (the demand rate of the process, reactant feed composition), but did not correlate well with operator response times. The factors were used to generate a distribution of response times from which a response time is sampled and simulated.

With backwards integration, this approach cannot be taken the response time of the operator must be purely deterministic. Therefore, the IPD technique is capable of accounting for stochastic operator responses, and thus is better suited for estimating alarm and safety system failure probabilities.

5.3 Future Work

In future work, three areas are worthy of consideration, as discussed briefly in the next subsections.

5.3.1 Rare-Event Sampling Strategies

TPS was developed by the molecular simulations community as a technique for studying rare-events. This thesis adapted the technique to handle rare process dynamic events. While TPS has provided many exciting opportunities in the molecular simulations community, other sampling strategies have been developed to handle similar problems. Three such techniques are forward flux sampling (FFS), milestoning, and transition interface probabilities (Allen et al., 2009).

The main computational effort when applying TPS to process dynamics is in solving the boundary-value problem, often accounting for well over 90% of the computation time in the two examples in Chapter 4. This limitation is circumvented by FFS, which does not require the calculation of backward-integrated trajectories. The general approach of this method involves an order parameter, λ, that spans the two regions of interest from $\lambda = 0$ at the interface to region A and $\lambda = 1$ at the interface to region B. The probability of a trajectory transitioning from:

$$\lambda = \alpha_i \text{ to } \lambda = \alpha_j; \quad \alpha_i < \alpha_j; \quad \alpha_i \geq 0; \quad \alpha_j \leq 1$$

(5.1)

is calculated for all α_i in [0,1]. In this manner, the probability of a trajectory spanning from region A to region B can be calculated. Additionally, the rate of formation of trajectories is calculated.

This method may provide further insights toward understanding special-cause events in process dynamics. Trajectories spanning the two regions (from green- to red-belt zones) are calculated using FFS without solving a boundary-value problem. Thus, the computational efficiency of FFS shifts to the efficiency of forward integration. Similar to TPS, random perturbations are introduced, retaining the potential to discover un-postulated trajectories. Furthermore, in TPS, the event time must be fixed, whereas in FFS the time length is allowed to vary. This can be important in setting alarm thresholds that give operators sufficient time to take corrective action.

5.3.2 Operator Decision Modeling

The informed prior distribution technique utilizes operator decision models, which seek to quantify operator response times as a function of various factors within the process. These include the state of the process during a special-cause event, which has significant impact on the operator's ability to diagnose and appropriately respond to events. Also, information pertaining to the specific operators describing those that are most effective and better trusted to handle more challenging tasks (larger demand rate shifts or operational shifts) is helpful. Furthermore, the time-of-day or time-of-year may play a critical role in determining operator's successes. A better understanding of these phenomena may influence scheduled shutdown periods or rate shifts at a plant.

5.3.3 Alarm and Safety System Design

This thesis has presented techniques for estimating the failure probabilities of rarely activated alarm and safety systems in chemical processes. These estimates permit engineers and plant managers to install more effective systems. Through the use of transition path sampling and informed prior distributions, better choices of alarmed variables, alarm thresholds, operator training, operator decisions, and automatic safety systems, can be selected. As the dynamic risk analysis community continues to develop sophisticated methods for understanding the performance of alarm and safety systems, better operational and design decisions will be implemented (Khan, 2015).

REFERENCES

Commercial Technologies for Oxygen Production. Nat'l. Ener. Tech. Lab., website. http://www.netl.doe.gov/research/coal/energy-systems/gasification/gasifipedia/commercial-oxygen (accessed 07/18/2016).

U.S. Chemical Safety and Hazard Investigation Board, http://www.csb.gov/ (accessed 09/29/2014).

U.S. Chemical Safety Board Report: BP America Refinery Explosion. http://www.csb.gov/bp-america-refinery-explosion/ (accessed 07/18/2016).

U.S. Chemical Safety Board Report: BP Deepwater Horizon. http://www.csb.gov/csb-board-approves-final-report-finding-deepwater-horizon-blowout-preventer-failed-due-to-unrecognized-pipe-buckling-phenomenon-during-emergency-well-control-efforts-on-april-20-2010-leading-to-environmental-disaster-in-gulf-of-mexico/ (accessed 07/18/2016).

U.S. Chemical Safety Board Report: Kleen Energy Natural Gas Explosion. http://www.csb.gov/kleen-energy-natural-gas-explosion/ (accessed 07/18/2016).

Worker Fatalities Reported to Federal and State OSHA. https://www.osha.gov/dep/fatcat/dep_fatcat.html (accessed 07/18/2016).

Agarwal, A., Biegler, L.T. and Zitney, S.E., 2008. Simulation and optimization of pressure swing adsorption systems using reduced-order modeling.*Industrial & Engineering Chemistry Research*, *48*(5), pp.2327-2343.

Ahooyi, T.M., Soroush, M., Arbogast, J.E., Seider, W.D. and Oktem, U.G., 2014. Maximum-likelihood maximum-entropy constrained probability density function estimation for prediction of rare events. *AIChE Journal*, *60*(3), pp.1013-1026.

Allen, R.J., Valeriani, C. and ten Wolde, P.R., 2009. Forward flux sampling for rare event simulations. *Journal of physics: Condensed matter*, *21*(46), p.463102.

Andrews, J.D. and Dunnett, S.J., 2000. Event-tree analysis using binary decision diagrams. *IEEE Transactions on Reliability*, *49*(2), pp.230-238.

Balakotaiah, V. and Luss, D., 1983. Multiplicity features of reacting systems: dependence of the steady-states of a CSTR on the residence time. *Chemical Engineering Science*, *38*(10), pp.1709-1721.

Baybutt, P., 2002. Layers of protection analysis for human factors (LOPA-HF). *Process Safety Progress*, *21*(2), pp.119-129.

Baybutt, P., 2003. Layers of protection analysis for human factors (LOPA-HF): an improved method for addressing human failures in process hazard analysis. In *Proceedings of the symposium on Human Error in Occupational Safety, American society of Safety Engineers* (pp. 163-175).

Bock, H.G., Diehl, M.M., Leineweber, D.B. and Schlöder, J.P., 2000. A direct multiple shooting method for real-time optimization of nonlinear DAE processes. In *Nonlinear Model Predictive Control* (pp. 245-267). Birkhäuser Basel.

Bolhuis, P.G., Chandler, D., Dellago, C. and Geissler, P.L., 2002. Transition path sampling: Throwing ropes over rough mountain passes, in the dark.*Annual review of physical chemistry*, *53*(1), pp.291-318.

Chen, J. and Patton, R.J., 2012. *Robust model-based fault diagnosis for dynamic systems* (Vol. 3). Springer Science & Business Media.

Clemen, R.T. and Reilly, T., 1999. Correlations and copulas for decision and risk analysis. *Management Science*, *45*(2), pp.208-224.

Cockshott, J.E., 2005. Probability bow-ties: a transparent risk management tool. *Process Safety and Environmental Protection*, *83*(4), pp.307-316.

Crowl, D.A. and Louvar, J.F., 2001. *Chemical process safety: fundamentals with applications*. Pearson Education.

Cuthrell, J.E. and Biegler, L.T., 1989. Simultaneous optimization and solution methods for batch reactor control profiles. *Computers & Chemical Engineering*, *13*(1), pp.49-62.

Bolhuis, P.G., Chandler, D., Dellago, C. and Geissler, P.L., 2002. Transition path sampling: Throwing ropes over rough mountain passes, in the dark.*Annual review of physical chemistry*, *53*(1), pp.291-318.

Dowell III, A.M., 1998. Layer of protection analysis for determining safety integrity level. *Isa Transactions*, *37*(3), pp.155-165.

Ericson, C.A., 2015. *Hazard analysis techniques for system safety*. John Wiley & Sons.

Escobedo, F.A., Borrero, E.E. and Araque, J.C., 2009. Transition path sampling and forward flux sampling. Applications to biological systems.*Journal of Physics: Condensed Matter*, *21*(33), p.333101.

Garcia, C.E., Prett, D.M. and Morari, M., 1989. Model predictive control: theory and practice—a survey. *Automatica*, *25*(3), pp.335-348.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B., 2014. *Bayesian data analysis* (Vol. 2). Boca Raton, FL, USA: Chapman & Hall/CRC.

Gowland, R., 2006. The accidental risk assessment methodology for industries (ARAMIS)/layer of protection analysis (LOPA) methodology: A step forward towards convergent practices in risk assessment. *Journal of hazardous materials*, *130*(3), pp.307-310.

gPROMS (Version 3.6.1) [Computer Software]. London, United Kingdom: Process System Enterprise.

Hartigan, J.A. and Wong, M.A., 1979. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 28*(1), pp.100-108.

Hendershot, D.C., 2006. An overview of inherently safer design. *Process safety progress, 25*(2), pp.98-107.

Hollifield, B. and Habibi, E., 2010. *Alarm Management Handbook*. Plant Automation Services, Incorporated.

Hopkins, A, 2009. Thinking about process safety indicators. *Safety Sci.*, 47(4), pp. 460-465.

Hosseini, S.H. and Takahashi, M., 2007, September. Combining static/dynamic fault trees and event trees using Bayesian networks. *International Conference on Computer Safety, Reliability, and Security* (pp. 93-99). Springer Berlin Heidelberg.

Hottel, H.C. and Sarofim, A.F., 2011. Radiative Transfer (McGraw-Hill Series in Mechanical Engineering).

Huang, R., Zavala, V.M. and Biegler, L.T., 2009. Advanced step nonlinear model predictive control for air separation units. *Journal of Process Control*,*19*(4), pp.678-685.

Jones, S., Kirchsteiger, C. and Bjerke, W., 1999. The importance of near miss reporting to further improve safety performance. *Journal of Loss Prevention in the process industries*, *12*(1), pp.59-67.

Kalantarnia, M., Khan, F. and Hawboldt, K., 2009. Dynamic risk assessment using failure assessment and Bayesian theory. *Journal of Loss Prevention in the Process Industries*, *22*(5), pp.600-606.

Kennedy, R. and Kirwan, B., 1998. Development of a hazard and operability-based method for identifying safety management vulnerabilities in high risk systems. *Safety Science*, *30*(3), pp.249-274.

Khajuria, H. and Pistikopoulos, E.N., 2011. Dynamic modeling and explicit/multi-parametric MPC control of pressure swing adsorption systems.*Journal of Process Control*, *21*(1), pp.151-163.

Khakzad, N., Khan, F. and Amyotte, P., 2012. Dynamic risk analysis using bow-tie approach. *Reliability Engineering & System Safety*, *104*, pp.36-44.

Khakzad, N., Khan, F. and Amyotte, P., 2011. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches.*Reliability Engineering & System Safety*, *96*(8), pp.925-932.

Khan, F., Rathnayaka, S. and Ahmed, S., 2015. Methods and models in process safety and risk management: Past, present and future. *Process Safety and Environmental Protection*, *98*, pp.116-147.

Kletz, T.A., 1999. *HAZOP and HAZAN: identifying and assessing process industry hazards*. IChemE.

Kletz, T., 2009. *What went wrong?: case histories of process plant disasters and how they could have been avoided*. Butterworth-Heinemann.

Kuczera, K., Jas, G.S. and Elber, R., 2009. Kinetics of Helix Unfolding: Molecular Dynamics Simulations with Milestoning†. *The Journal of Physical Chemistry A*, *113*(26), pp.7461-7473.

Latham, D.A., McAuley, K.B., Peppley, B.A. and Raybold, T.M., 2011. Mathematical modeling of an industrial steam-methane reformer for on-line deployment. *Fuel processing technology*, *92*(8), pp.1574-1586.

Leveson, N.G. and Stephanopoulos, G., 2014. A system-theoretic, control-inspired view and approach to process safety. *AIChE Journal*, *60*(1), pp.2-14.

Luyben, W.L., 1989. *Process modeling, simulation and control for chemical engineers*. McGraw-Hill Higher Education.

Mannan, M.S., O'Connor, T.M. and West, H.H., 1999. Accident history database: An opportunity. *Environmental Progress*, *18*(1), pp.1-6.

Marsh, D.W.R. and Bearfield, G., 2008. Generalizing event trees using Bayesian networks. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *222*(2), pp.105-114.

Meel, A. and Seider, W.D., 2006. Plant-specific dynamic failure assessment using Bayesian theory. *Chemical engineering science*, *61*(21), pp.7036-7056.

Oh, M. and Pantelides, C.C., 1996. A modelling and simulation language for combined lumped and distributed parameter systems. *Computers & Chemical Engineering*, *20*(6), pp.611-633.

Oktem, U.G., Seider, W.D., Soroush, M. and Pariyani, A., 2013. Improve process safety with near-miss analysis. *Chemical Engineering Progress*,*109*(5), pp.20-27.

Palmer, C. and Chung, P.W.H., 2008. A computer tool for batch hazard and operability studies. *Journal of Loss Prevention in the Process Industries*,*21*(5), pp.537-542.

Pariyani, A., Seider, W.D., Oktem, U.G. and Soroush, M., 2010. Incidents Investigation and Dynamic Analysis of Large Alarm Databases in Chemical Plants: A Fluidized-Catalytic-Cracking Unit Case Study†. *Industrial & Engineering Chemistry Research*, *49*(17), pp.8062-8079.

Pariyani, A., Seider, W.D., Oktem, U.G. and Soroush, M., 2010. Incidents Investigation and Dynamic Analysis of Large Alarm Databases in Chemical Plants: A Fluidized-Catalytic-Cracking Unit Case Study†. *Industrial & Engineering Chemistry Research*, *49*(17), pp.8062-8079.

Pariyani, A., Seider, W.D., Oktem, U.G. and Soroush, M., 2010. Incidents Investigation and Dynamic Analysis of Large Alarm Databases in Chemical Plants: A Fluidized-Catalytic-Cracking Unit Case Study†. *Industrial & Engineering Chemistry Research*, *49*(17), pp.8062-8079.

Pariyani, A., Seider, W.D., Oktem, U.G. and Soroush, M., 2012. Dynamic risk analysis using alarm databases to improve process safety and product quality: Part II—Bayesian analysis. *AIChE Journal*, *58*(3), pp.826-841.

Peng, D.Y. and Robinson, D.B., 1976. A new two-constant equation of state.*Industrial & Engineering Chemistry Fundamentals*, *15*(1), pp.59-64.

Phimister, J.R., Oktem, U., Kleindorfer, P.R. and Kunreuther, H., 2003. Near-miss incident management in the chemical process industry. *Risk Analysis*,*23*(3), pp.445-459.

Reason, J., 1990. *Human error*. Cambridge university press.

Rosenthal, I., Kleindorfer, P.R. and Elliott, M.R., 2006. Predicting and confirming the effectiveness of systems for managing low-probability chemical process risks. *Process safety progress*, *25*(2), pp.135-155.

Rothenberg, D.H., 2009. *Alarm management for process control: a best-practice guide for design, implementation, and use of industrial alarm systems*. Momentum Press.

Ruilin, Z. and Lowndes, I.S., 2010. The application of a coupled artificial neural network and fault tree analysis model to predict coal and gas outbursts. *International Journal of Coal Geology*, *84*(2), pp.141-152.

Seborg, D.E., Mellichamp, D.A. and Edgar, T.F., 2010. FJD III, Process Dynamics and Control.

Seider, W.D., Seader, J.D. and Lewin, D.R., 2009. *Product & Process Design Principles: Synthesis, Analysis And Evaluation*. John Wiley & Sons.

Soroush, M. and Kravaris, C., 1992. Discrete-time nonlinear controller synthesis by input/output linearization. *AIChE Journal*, *38*(12), pp.1923-1945.

Srinivasan, R. and Natarajan, S., 2012. Developments in inherent safety: a review of the progress during 2001–2011 and opportunities ahead. *Process Safety and Environmental Protection*, *90*(5), pp.389-403.

Stavrianidis, P. and Bhimavarapu, K., 2000. Performance-based standards: safety instrumented functions and safety integrity levels. *Journal of hazardous materials*, *71*(1), pp.449-465.

Stavrianidis, P. and Bhimavarapu, K., 1998. Safety instrumented functions and safety integrity levels (SIL). *ISA transactions*, *37*(4), pp.337-351.

Stephanopoulos, G., 1984. Chemical process control: an introduction to theory and practice.

Summers, A.E., 2003. Introduction to layers of protection analysis. *Journal of Hazardous Materials*, *104*(1), pp.163-168.

Tanaka, H., Fan, L.T., Lai, F.S. and Toguchi, K., 1983. Fault-tree analysis by fuzzy probability. *IEEE Transactions on Reliability*, *32*(5), pp.453-457.

Vaidhyanathan, R. and Venkatasubramanian, V., 1995. Digraph-based models for automated HAZOP analysis. *Reliability Engineering & System Safety*, *50*(1), pp.33-49.

Venkatasubramanian, V. and Vaidhyanathan, R., 1994. A knowledge-based framework for automating HAZOP analysis. *AIChE Journal*, *40*(3), pp.496-505.

Venkatasubramanian, V., Zhao, J. and Viswanathan, S., 2000. Intelligent systems for HAZOP analysis of complex process plants. *Computers & Chemical Engineering*, *24*(9), pp.2291-2302.

Xu, J. and Froment, G.F., 1989. Methane steam reforming, methanation and water-gas shift: I. Intrinsic kinetics. *AIChE Journal*, *35*(1), pp.88-96.

Yeomans, H. and Grossmann, I.E., 1999. A systematic modeling framework of superstructure optimization in process synthesis. *Computers & Chemical Engineering*, *23*(6), pp.709-731.

Yi, W. and Bier, V.M., 1998. An application of copulas to accident precursor analysis. *Management Science*, *44*(12-part-2), pp.S257-S270.