UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
ENERGÍA – SISTEMAS ENERGÉTICOS

QUANTITATIVE MEASURES OF HUMAN PERFORMANCE & RELIABILITY IN SUPPORT OF NUCLEAR POWER PLANT SAFETY CULTURE

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN INGENIERÍA

PRESENTA:
PAMELA FRAN NELSON EDELSTEIN

TUTORA PRINCIPAL
Cecilia Martín del Campo Márquez, Facultad de Ingeniería
COMITÉ TUTOR
Juan Luis François Lacouture, Facultad de Ingeniería
Carlos Chávez Mercado, Facultad de Ingeniería
Bruce P. Hallbert, Idaho National Laboratory
Ali Mosleh, University of California, Los Angeles

MÉXICO, D. F., FEBRERO 2016

JURADO ASIGNADO:

Presidente:	Dr. Juan Luis Francois Lacouture
Secretario:	Dr. Carlos Chávez Mercado
Vocal:	Dra. Cecilia Martín del Campo Márquez
1 er. Suplente:	Dr. Ali Mosleh
2 do. Suplente:	Dr. Bruce P. Hallbert

Lugar o lugares donde se realizó la tesis: Facultad de Ingeniera

TUTOR DE TESIS:

Dra. Cecilia Martín del Campo Márquez

Acknowledgements

I would like to acknowledge all the people who supported my endeavor to finally complete my dream of getting a Ph.D. and becoming another Dr. Nelson.

Special thanks to my committee members, Dr. Juan Luis Francois, Dr. Carlos Chavez, Dr. Ali Mosleh, Dr. Bruce Hallbert, and chair, Dr. Cecilia Martín del Campo. Thank you to Cecilia for reminding me that I am an expert in the topic and that the completion of this dissertation allows me to prove it.

Thank you to my mother and Bill, my father and Suzanne, Bruce and Melissa for always being there for me.

Thank you to my friends; in particular, Chío and Adriana for the celebrations and Tere for the coaching.

I acknowledge the support by Rick Grantom who never gave up believing in my project and made it possible to gather much of the data.

An extra special thank you to Dr. Mike Boldrick for all his support.

I dedicate this dissertation with love to my wonderful children, Adrián and Elena.

.

Table of Contents

List of Figures

List of Tables

Acronyms

BBN	Bayesian Belief Networks
BN	Bayesian Networks
CAP	Corrective Action Program
CAQ-L1	Condition Adverse to Quality-Level 1 (same as STCR)
CAQ-L2	Condition Adverse to Quality-Level 2 (same as DLCR)
CEA	Commissariat à l'Énergie Atomique
CR	Condition Report
DLCR	Department Level Condition Report (same as CAQ-L2)
EPIX	Equipment Performance and Information Exchange (an equipment failure database)
HEP	Human Error Probability
HRA	Human Reliability Analysis
I&C	Instrumentation and Control
ICES	INPO Consolidated Event System
INPO	Institute of Nuclear Power Operations
INSAG	International Nuclear Safety Advisory Group
IPE	Individual Plant Examination
KINS	Korea Institute of Nuclear Safety
LER	Licensee Event Report
NPP	Nuclear Power Plant
NPRDS	Nuclear Plant Reliability Data System
CNAQ	Condition Not Adverse to Quality
NRC	Nuclear Regulatory Commission
OPIS	Operational Performance Information System
PED	Plant Events Database
RPS	Reactor Protection System
SCAQ	Significant Condition Adverse to Quality (same as SCCR)
SCCR	Significant Contributor to Condition Reports (same as SCAQ)
SCRAM	Safety control rod axe man. The sudden shutting down of a nuclear reactor usually by rapid insertion of control rods either automatically or manually by the reactor operator.
SCWE	Safety Conscious Work Environment
SECY	Staff position papers before the Commission
STCR	Station Level Condition Report (same as CAQ-L1)
V&V	Verification and Validation

Abstract

Quantitative measures of human performance and reliability are fundamental in the support of nuclear power plant safety culture. These measures can be used not only for updating human reliability data, but also to quantify organizational performance factors. This dissertation is focused on human and organizational reliability, with the objective of developing a process to monitor and reduce the number of consequential errors in nuclear power plants. With this end in mind, this dissertation develops a new analysis method and associated capabilities to evaluate and predict organizational resilience levels of the nuclear power plant. It identifies the human and organizational errors associated with causes of consequential events. New leading performance indicators are developed that provide insights into organizational stress levels, leading to and facilitating the development of compensating measures to reduce stress levels (i.e., increase organizational resilience levels).

The development of operational performance indicators is of utmost importance for nuclear power plants, since they measure, track and trend plant operation. Leading or predictive indicators are ideal for reducing the likelihood of consequential events. This dissertation describes the operational data analysis of the information contained in ten years of Condition Reports generated by one plant's Corrective Action Program (CAP). The methodology considers human error and organizational factors because of their large contribution to consequential events. The results include a tool to be used for the identification, prediction and reduction of the likelihood of significant consequential events. This tool is based on the resilience curve, a stress-strain curve that was built from the plant's operational data. The stress is described by the number of unresolved Condition Reports. The strain is represented by the number of preventive maintenance tasks and other periodic work activities (i.e., baseline activities), as well as ongoing corrective actions to resolve the Condition Reports (i.e., corrective action workload). The use of the Condition Reporting Program is appropriate since for this plant it represents both permanent, repetitive work activities (i.e., baseline) and emergent work activities. The resilience threshold is determined for the facility. When this threshold is exceeded, the resultant organizational stress exceeds the station's ability to operate successfully, with a corresponding increased likelihood that a consequential event will occur. A leading performance indicator is developed to reduce the likelihood of consequential events at nuclear power plants through the recognition of plant specific situations leading to or contributing to excessive organizational stress levels (i.e., reduced organizational resilience margins).

When the performance indicator shows a decrease in resilience, a Bayesian Belief Network (BBN) is used to determine the best barrier to install. The methodology developed in the dissertation to build and evaluate the BBN is based on the interactions among causes in the CAP database. The model is expanded to be able to compare the barriers based on their costs and savings in order to aid the plant in choosing the most appropriate barrier. These barriers translate into plant specific compensatory measures that can be utilized by station management or personnel to offset reduced organizational resilience margins.

Introduction

Background

The Nuclear Regulatory Commission's "Policy Statement on the Conduct of Nuclear Power Plant Operations" (NRC, 1989) refers to safety culture as "the necessary full attention to safety matters" and the "personal dedication and accountability of all individuals engaged in any activity which has a bearing on the safety of nuclear power plants. A strong safety culture is one that has a strong safety-first focus."

The Commission has referenced the International Nuclear Safety Advisory Group's (INSAG) definition of safety culture as follows: "Safety Culture is that assembly of characteristics and attitudes in organizations and individuals which establishes that, as an overriding priority, nuclear plant safety issues receive the attention warranted by their significance."

The Commission's policy statement "Freedom of Employees in the Nuclear Industry to Raise Safety Concerns Without Fear of Retaliation," May 14, 1996, describes the Safety Conscious Work Environment (SCWE) as "a work environment where employees are encouraged to raise safety concerns and where concerns are promptly reviewed, given the proper priority based on their potential safety significance, and appropriately resolved with timely feedback to the originator of the concerns and to other employees." SCWE is described as an attribute of safety culture in SECY-04-0111, "Recommended Staff Actions Regarding Agency Guidance in the Areas of Safety Conscious Work Environment and Safety Culture," August 30, 2004. The NRC has developed Guidance for Establishing and Maintaining a Safety Conscious Work Environment.

However, human error cannot be avoided, as mentioned in an assessment conducted by the Commissariat à l'Énergie Atomique (CEA) in France that concluded that no amount of technical innovation can eliminate the risk of human-induced errors associated with the operation of nuclear power plants. Two types of mistakes were deemed most serious: errors committed during field operations, such as maintenance and testing, that can cause an accident or cause a malfunction or failure of important equipment; and human errors made during operational events that cascade to complete failure of safety functions or systems (Evans, 2011).

The concern about human errors is not only that they can impact initiating event frequency, but also that they can cause unexpected failures in the plant that can cause plant downtime or worker injury (which also affects the safety of the plant). For example, human errors in test and maintenance activities of NPPs have the potential for inducing unplanned reactor trips. The Korean regulatory organization for nuclear and radiological systems, Korea Institute of Nuclear Safety (KINS), provides a list of the major events, including unplanned reactor trips and unplanned initiations of safety systems that have occurred in Korean nuclear power plants, on a public website, the Operational Performance Information System (OPIS) (KINS, 2013). According to OPIS, about 23% of the events that occurred during 2002 ~ 2006 were caused by human error. More recently, during 2004 ~ 2005, the contribution of human error to the unplanned reactor trip events had grown to about 34%, a significant increase. Recently the data indicates that there were 150 trip events from 2004-2013, 32 of which (21%) were assigned as due to human error.

It appears that the coding rules (i.e., for assigning cause of an event as being the result of human error) were stable over the coding periods; that is: 2002-2006: 23/102(23%), 2004-2005 13/39 (34%), 2002-2003:8/44 (18.2%), 2005-2006: 8/40 (20%). In data from the Licensee Event Reports (LERs) in the United States, the contribution to unplanned SCRAMs during maintenance and surveillance activities was shown to be almost 40% (Wegner, 1999). INPO also issued several Significant Operating Event Reports (SOERs) that further point to declining human performance.

Interest in analyzing and reducing the human-induced or human-related unplanned reactor trip events has been increasing gradually in response to the increased number of human-induced unplanned reactor trip events (Kim & Park, 2008; Kim & Park, 2011).

Test and maintenance activities performed in nuclear power plants are essential for sustaining the safety of the power plant and maintaining the reliability of plant systems and components. However, on the other hand, the potential of human errors during test and maintenance activities has also the possibility of inducing unplanned reactor trips or power derate in an active error mode, or inducing latent failures that render safety-related systems or functions unavailable when they are demanded for incidents/accidents (Reason, 1990; Dhillon & Liu, 2006). Often, human errors are related to problems in establishing the maintenance or testing boundaries (i.e., equipment clearances) to allow these activities to be performed in a safe manner without inadvertent actuations of equipment or endangering plant personnel.

Generally in conventional probabilistic safety assessments (PSAs), human actions leading to initiating events, i.e., unplanned reactor trips or power reductions, have not been modeled explicitly, while maintenance human errors have only been modeled in PSA on the aspect of system unavailability (IAEA, 1992; IAEA, 1995). Laakso, Pyy & Reiman (1998) and Pyy (2001) introduced an analysis of maintenance human failures and their effects, and discussed their safety significance from the PSA point of view; however, the effects and safety significance of maintenance related human failures mainly included equipment unavailability or equipment malfunction, very few were related to unplanned reactor trips. But, there is a growing need to analyze the mechanism associated with human-induced unplanned reactor trips and the organizational constraints and characteristics associated with often performed maintenance and testing activities leading to consequential human errors such as unplanned reactor trips, in order to provide a basis for managing maintenance and operations related human errors as well as to incorporate human-induced initiators more explicitly in PSA models (Hirschberg, 2004; Canavan & Hannaman, 2004). INPO/WANO track consequential operational events and have noted a trend upwards over the last several years prompting their issuance of SOER 10-02 where the balance between rule based and knowledge based procedural guidance is being further evaluated resulting in questions related to the risk significance of performing routine stations activities (e.g., surveillance tests, preventive maintenance activities, etc.).

Here, the concern is not only induced reactor trips, but also human errors that can cause undesirable (i.e., consequential) events or complications such as an inadvertent Safety Injection actuation or inadvertent actuation of equipment. An event of this type was seen in the nuclear industry where a human error can cause a complication, given an accident (e.g., Three Mile Island).

The accident started at 4:00 a.m. on Wednesday March 28, 1979 with the loss of normal water supply to the steam generators. The primary transient caused emergency shutdown, which gradually lowered pressure in the primary cooling system. After 12 seconds the relief valve received as normal the command to close but this valve remains jammed open. The primary cooling system continued to discharge into the pressurizer relief tank, located in the containment, at a flow-rate of 60 metric tons per hour (there are approximately 200 metric tons of primary coolant).

The steam generator auxiliary feedwater system pumps started up normally after 30 seconds, but the connecting valves between the pumps and the steam generators were closed instead of open, due to a maintenance error. The generators dried out in 2 to 3 minutes, stopping all cooling of the primary system. Although the position indicator for these valves located in the control room signal this fault, eight minutes passed before the operators identified the fault and gave the command manually to open the valves. Twenty-five minutes passed before the situation of the secondary cooling system stabilized, after numerous operations, no doubt commanding all the attention of the operating team (IAEA, 2012).

In addition, there have been events that do not cause a nuclear safety event (i.e., a core damaging event); however, their consequential cost is high due to lost generation, equipment damaged due to maintenance human errors (e.g., improper/incorrect lubrication), radiological cleanup and associated costs, lost time accidents, and equipment/plant damage due to inappropriate operation of equipment (e.g., flooding a room by opening a wrong valve, etc.).

Large amounts of plant event data exist from nuclear power plant Corrective Action Programs (CAP). These corrective action programs are the primary mechanism where station employees and contractors identify problems and issues that need to be addressed. In most cases, items in the CAP are minor or administrative and planned events such as a work order that can also be identified through CAP. The CAP is intended to provide station personnel with a means to identify problems no matter how big or small and CAP also satisfies regulatory requirements for Problem Identification and Reporting processes. However, this data is not uniform from station to station and has not been assembled in a manner that is helpful to fully understand human error rates, their associated human error classification, and their risk significance. Although thousands of events are reported each year (most of which are administrative low level items and not risk significant, the classification is done differently at each plant, thus making difficult the formation of a generic database or any subsequent higher level analyses or research, such as that typically existing for equipment (e.g., Equipment Reliability Programs). For this reason alone, it is important to reactor safety to design a human error data base that provides the process, data and information to facilitate quantitative analyses and future human performance research, based on actual plant events as documented in Corrective Action Programs and other industry programs (e.g., NRC Licensee Event Reports, INPO Significant Operating Experience Reports). Once developed such a database and associated computational algorithms could be maintained and updated to provide insightful trends into human performance and formal Human Reliability Analyses (HRA) not only on a plant specific basis but also on a fleet or industry basis; and these are the goals of the research presented in this dissertation. This potentially could allow human error precursors to consequential operational events to be better identified and allow risk management methods to reduce the likelihood of events.

Human Error Data

Human error is almost always involved in one way or another in accidents in any industry, and the nuclear industry is not an exception. While it is possible to refer to the major accidents that have occurred and investigate the human causes, this study concentrates rather on human actions that are carried out daily, such as during maintenance activities. Also, in this study we are concentrating on the occurrence of consequential events during normal operation and outages, and not solely core melt accidents. This type of information is recorded in the Condition Reports (CRs) of the Corrective Action Programs (CAPs). The database created from these events is used in this study to investigate the relationships between the causes and their effect on consequential event frequency.

Human error and organizational performance is of special interest in any industry, and the nuclear industry has developed methods for performing Human Reliability Analysis (HRA) to calculate the contribution of human error to accidents. While built on human factors, HRA distinguished itself early on from human factors due to its emphasis on predicting human performance. While one of the major focus areas of human factors has been improving the design of novel systems to optimize human performance, HRA has largely focused on predicting human performance for as-built systems. Over time, as HRA became closely tied particularly to the nuclear energy industry, it increasingly became a field associated more with reliability engineering than human factors. Yet, the similarity to human factors has not abated, nor has the opportunity for the two fields to cooperate. Human factors research provides the empirical basis to support predicting human performance in HRA. Importantly, HRA continues to benefit human factors by providing: (1) a framework for modeling human performance, (2) an example of how a human factors discipline can be seamlessly integrated with an engineering field, and (3) insights on how predictive modeling may be used as a system design tool (Boring, R.L., Roth, E., Straeter, O., Laumann, K., Blackman, H.S., Oxstrand, J., & Persensky, J.J., 2009).

There have been projects to collect data to inform quantification in HRA, starting with the work done for the THERP methodology by Swain (1983). These efforts have continued to the present time, with efforts like the US Nuclear Regulatory Commission's Human Event Repository and Analysis (HERA) system (Hallbert, et al., 2007) and the UK's CORE Database (Kirwan, 1997). However, many experts in HRA are of the opinion that there should have been more effort on collecting human error data for the purpose of quantifying the probabilities of human error (Boring, 2012); however, there does exist a wealth of information in the Condition Reports, products of the Corrective Action Programs, at most nuclear plants. The corrective action process includes formal mechanisms to report, capture, assess, and correct organizational failures or shortcomings. Typically the focus is placed on identifying root causes and implementing corrective actions to ensure organizational learning and improvement.

In fact, in the nuclear industry, we propose that the CAPs are a source of metadata. The information contained in the CRs at every nuclear power plant is invaluable, and while the reports for each event may be lengthy, there should be an efficient way to record, store and retrieve data and feedback continually. The statistics of the data can tell us much about the trends in failures, whether system or human failures; however, if the information is not codified to work for the intended database, the results may be inaccurate. For this reason, this dissertation describes the review and work done to extract benefit from the root cause analysis done on any abnormal occurrence at a nuclear plant, and presents a model to include this wealth of information in a structure that furthers the knowledge management about human errors in nuclear power plants.

The cause coding together with study of written descriptions in the failure and repair work orders helps to identify candidates of human errors related to maintenance activities. From a sample containing thousands of Condition Reports labeled human-related it was possible to utilize it for evaluating the effect of introducing barriers to the plant. Once validated, the model can serve to predict events, risk-inform the procedures to reduce the reoccurrence of the events, and to avoid consequential events.

Problem Statement

As was mentioned in the previous section, maintenance and testing (e.g., surveillance testing) of reactor systems are important causes of unplanned reactor trips, turbine trips, down-power events, inadvertent system actuations, damage to the plant equipment and even harm to workers and possibly to the public. For this reason it is essential to find ways to reduce undesired events. Although there is an entire discipline entitled Human Reliability Analysis used in Probabilistic Safety Analyses which analyzes human errors and their probabilities for PRA, there are several limitations. This analysis only considers errors on components that are modeled in the PRA. Also, the research and in-depth analysis has been concentrated on the human errors after an initiating event presents itself (i.e., post-initiator). The end state considered is core melt or large, early release.

In addition, human errors are usually classified into three types:

A. Pre-initiator human actions,
B. Human errors that cause initiating events and,
C. Post-initiating event human actions.

Type C analysis is the center of many present Human Reliability Analysis (HRA) research projects as well as nearly all of the HRAs that have been done for commercial plants. Type B events are typically considered to be included in the initiating event frequency. Finally, Type A errors are included in the HRA, that is, actions that can be performed erroneously and cause an equipment misalignment or miscalibration; however, these events are usually found to be not risk significant[1] for PSA purposes and thus not highly scrutinized in the review processes. They can, in fact, be highly risk significant from the perspective of a consequential operational event which is the point being emphasized by nuclear oversight organizations such as INPO.

A quantitative method for establishing the contribution of human performance and reliability to consequential operational events leading to, but prior to an initiating event is needed to quantitatively correlate human performance and reliability of often repeated tasks and activities to operational events considered undesirable during nuclear power plant operation leading to, but prior to, an initiating event. It is important to establish quantitative measures of human performance and reliability (i.e., figures-of-merit) prior to the occurrence of an initiating event, as well as to provide a method to evaluate organizational and operational practices and processes (i.e., procedures) to assess the risk contribution of those activities leading to initiating events.

Human performance and human reliability focused risk informed performance indicators are needed to monitor consequential human performance trends and measure effectiveness of station processes, procedures, and corrective actions. This could be a basis for a systematic approach for establishing the risk significance of procedural related human actions (i.e., procedure risk profiling) performed at nuclear power stations (e.g., Operations and Maintenance organizations).

Thus, for this thesis, it is proposed to develop a robust human performance monitoring and tracking methodology and tool that can be deployed to nuclear plant organizations for the purposes of quantitatively measuring and monitoring human performance events and trends for the purposes of reducing the occurrence of consequential human errors (i.e., prior to the initiating event).

Justification

Nuclear energy is necessary to fulfill energy demands supporting civilized societies as an environmentally clean source of baseload electric power, but has potential hazards associated with the use of fissionable materials. Thus, this is an important topic to study in more detail. Safety is the highest priority issue at nuclear power stations and one of the most important contributors to assuring safety is improving human performance. Due to the dramatic increase of significant operational events in recent years (INPO, 2010), there is a need to develop and deploy risk informed solutions that can be applied to procedures / work instructions to ensure that the right level of detail and human factor engineering principles are applied to critical activities to reduce the likelihood of active and latent errors that challenge reactor safety and equipment reliability, in effect, increase safety culture at nuclear power plants.

In fact, Magwood (2015) writes that safety culture has been identified as having played an important role in allowing precursor conditions at Fukushima to go unaddressed. Ensuring nuclear reactor safety is not only a question of physical protection against all credible

[1] This is defined as having a risk importance measure of RAW>2 or FV>.005.

threats, enhancing robustness of important safety systems and increasing redundancy of back-up power and water cooling systems, but also one of making certain that qualified and trained staff are supported by effective procedures. However, these assets are valued only in an organizational culture that places a premium on ensuring high levels of safety, or implementing what is called an effective 'nuclear safety culture'. In recognition of the importance of such factors, this disertation presents an approach to enhancing organizational resilience so that staff is better able to respond under increasing organizational stress due to an excess of work activities.

The structure of the dissertation is as follows:

Chapter 1 presents analysis of the data contained in the CAP database. The methodology followed during this study is also presented.

Chapter 2 presents de development of the resilience curve and leading performance indicator.

Chapter 3 presents the development of the tool for evaluating barriers, employing Bayesian Networks.

The Conclusions summarize the project and discuss future work.

Chapter 1 Methodology/Data Analysis

The methodology includes examining the data and finding ways to convert data to information and then to knowledge that can be shared and maintained using a knowledge management tool or system. In particular, given the current state of the nuclear industry due to many simultaneous retirements, there is a great need for nuclear knowledge management.

"Organizational resilience" can be considered as the ability of an organization to anticipate, prepare for, and respond and adapt to incremental change and sudden disruptions in order to survive and prosper. Resilience engineering, adapting materials science, the use of factor analysis, bayesian networks, etc.

The development of the resilience curve is used to identify when barriers should be evaluated to implement to avoid consequential events.

The curve is developed into a leading performance indicator to ease the interpretation for the plant personnel.

When the stress factor is in the white band, the plant should check upcoming maintenance programs to combine with the stress factor and start the barrier analysis

The barrier analysis consists of using the bayesian network for determining what the cost and benefit of each barrier.

The barrier is chosen for the problem at hand and is recommended for implementation. It may be a temporary or permanent barrier.

Analysis of effectiveness of barrier should be carried out, and if it is permanent, there should be periodic evaluations.

1.1 Use of Statistics for Pre-Initiators

1.1.1 Classification of Events

The observation of the nuclear power plants' Condition Reports makes it evident that human performance is important. For example, in several of the plant CAP databases, 20% of all Condition Reports are coded as human performance. In the Korean plants, the public database also indicates a 20% contribution due to human performance (KINS, 2013). In the INPO Consolidated Event System database (ICES), which is an effort to combine EPIX and NPRDS data, only a 7% contribution from human caused events is observed (INPO, 2013). This is likely due to the types of events reported to INPO; in general they are associated with equipment reliability. In order to test this hypothesis, an exercise was performed to relate the categories from the ICES database to the CAP event and cause categories defined for the pilot plant. While the categories correspond in the majority of the cases, the numbers of occurrences are substantially different. For example, the category for *Did not follow document sequence or steps correctly* from ICES reports 107 events for 104 NPPs from 2005–2013, while a similar category in the pilot plant *Procedure / instruction /*

step not performed or performed out of sequence reported 104 events for the same time period. The numbers are almost identical but the ICES includes all plants, thus implying that the ICES value should be on the order of 100 times larger. In fact, the 7% of ICES reports that identify a human-related cause is much less than the 62% contribution of human error to the most severe condition reports (SCAQ) in the CAP database of the pilot plant. Since we expect most plants to be similar to the pilot plant, the nuclear industry must be careful when using this INPO data to reflect human performance at NPPs.

While we can identify many processes prone to error in a nuclear power station, in order to count those that are important for pre-initiators, we can focus on the following: operation, maintenance, engineering and training. Reason (1990) coined the term general failure types (GFT) for human factors work, but did not specify them. Table 1-1 classifies the general failure types (GFT) for these four processes at a NPP into four categories (supervision, work practices, work instructions and procedures) and lists the human performance failures that occur. The GFTs for management related failures are not included; however there is a discussion of these organizational failure contributions in Section 1.7. After that discussion, the organizational failures are considered as an integral part of the development of the models in the thesis.

Table 1-1 General Failure Types and their Causes.

General Failure Type	Causes
Supervision	S1: pre-job preparation or briefing inadequate S2: prioritization of work activities inadequate S3: oversight/monitoring of task/personnel performance inadequate S4: supervisory practices promote/allow undesirable behaviors, S5: communication of information inadequate/untimely (leader) S6: worker capabilities not matched with task demands S7: insufficient time allotted for worker to perform task
Work practices	WP1: slip or lapse WP2: communication of information inadequate/untimely (worker) WP3: shortcuts used WP4: negligent WP5: knowledge, skills, and abilities not applied during task performance
Work instructions	WI1: document usage classification incorrect • *"Available" procedure should have been "In-Hand"* WI2: needed/necessary document does not exist WI3: document contents incorrect or missing WI4: document inadequacy (unclear/poor format)
Procedures	P1: procedure/instruction/step implemented incorrectly (intent not met) P2: procedure/instruction/step not performed – performed out of sequence P3: in-hand procedure not in use at work station during performance P4: procedure usage requirements not met (placekeeping, signoffs, verification, use of n/a-entry correction) P5: referenced procedure not readily available in work area/in-hand for infrequent or first time evolutions

1.2 Data Analysis

There are many inconsequential items included in the condition reports at the department level, called here Department Level Condition Reports (DLCR), but the interactions of the plant and the organizational responses sometimes adversely coincide and produce consequential events, called here: Significant Contribution Condition Reports (SCCR), while later on they will be referred to as SCAQs. These SCCRs usually result in plant trips, inadvertent operation of a safety system, or injury.

Figure 1-1 shows the human error rates (number of events per day), by cause (from Table 1-1), based on analysis of 7 years of DLCRs (2005 – 2011). The most important contributor is WP1 (slip or lapse), at a rate of more than one event per day. While this may initially seem alarming, it makes sense, in that maintenance and surveillance activities are continuously going on and many of them concurrently. Additionally, these slips or lapses have not caused any consequential outcome, such as plant trip, inadvertent operation of a safety system or injury. More generally, the work practice (WP) causes are by far the most frequent, followed by the work instruction (WI) causes.

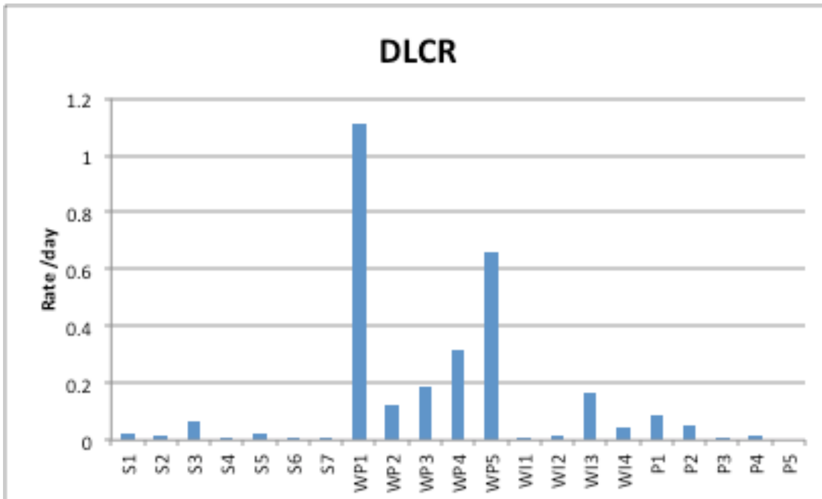

Figure 1-1 Human error rates per day from 7 years of Department Level Condition Reports (DLCRs).

Figure 1-2 graphs the rate of occurrences of events at a station level from the Station Level Condition Reports (STCR). While WP1 continues to be an important contributor, WP5, *knowledge, skills, and abilities not applied during task performance* has a higher rate of occurrence.

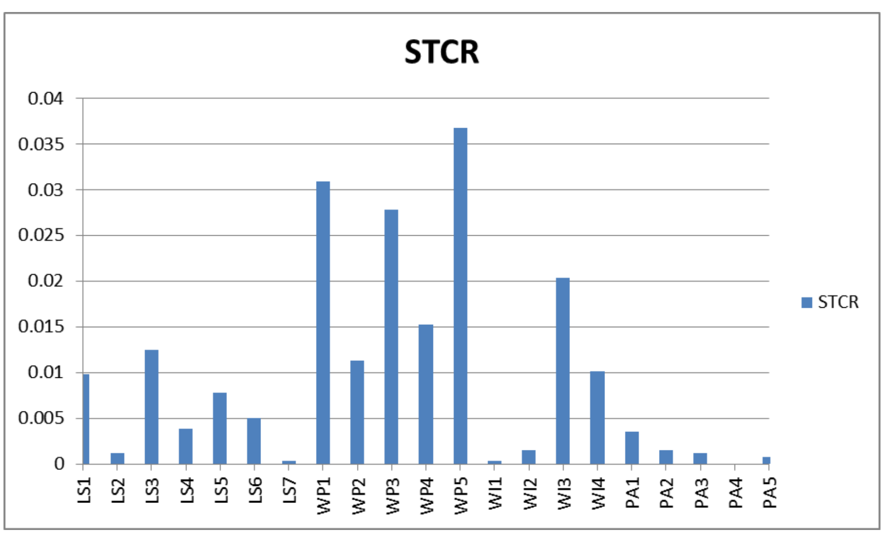

Figure 1-2 Human error rates per day from 7 years of Station Level Condition Reports (STCRs).

For a similar analysis performed using SCCRs (Figure 1-2), we observe a much lower the error rate, approximately one order of magnitude less. This is to be expected, since these events lead to consequences that are major problems for the plant. In many cases these correspond to Licensee Event Reports (LERs) that are also to be reported to the regulatory body (these are identified as NERs in the Corrective Action Program at Laguna Verde, Mexico). The WP and WI causes are still the most frequent, but WP is not as dominant, and supervision (S) events are nearly as frequent as the WPs.

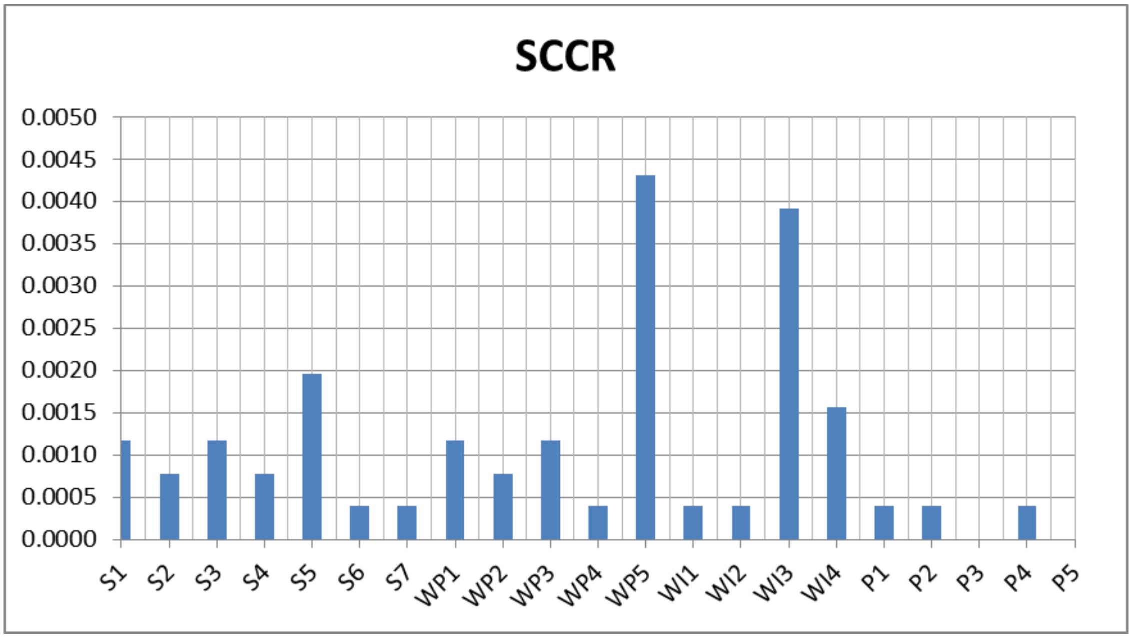

Figure 1-3 Human error rates per day from 7 years of Significant Contribution Condition Reports (SCCRs).

Figure 1-4 presents a comparison of the types of errors, the department level (DLCR), station level (SLCR), and Significant Contribution Condition Report (SCCR) events. The same pattern as observed in three figures above can be observed, that is, the work practices and work instructions are the main causes of the human errors.

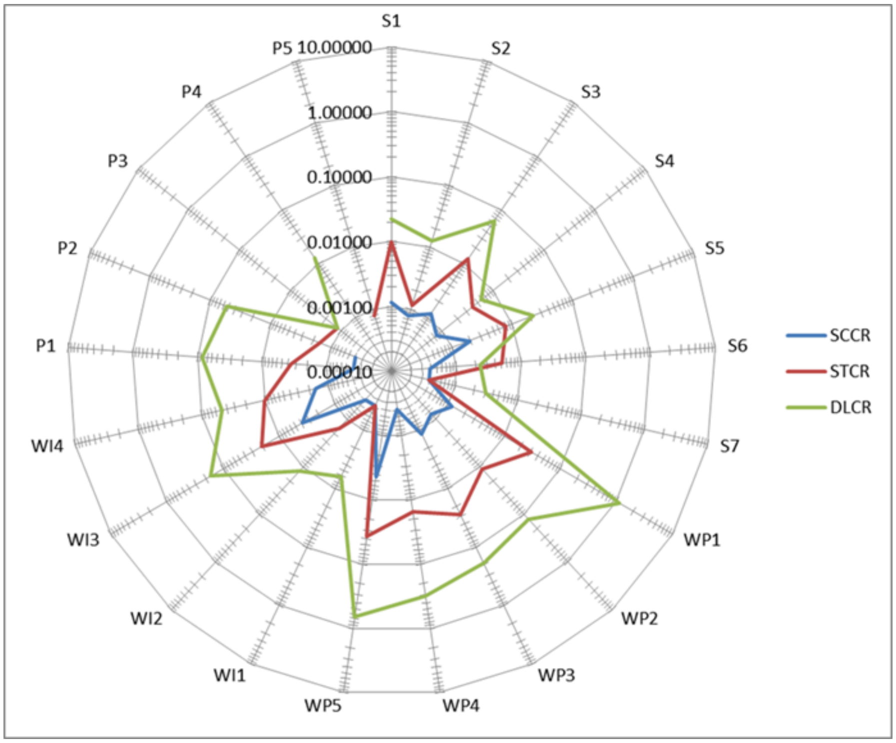

Figure 1-4 Logarithmic-scale comparison of DLCR, STCR and SCCR causes.

There are numerous activities occurring at a NPP during normal operation; for example, some components are tested once a month, and, others, once a week, some only receive preventive maintenance every 3 months and corrective maintenance is performed as needed for all components. The early morning manager meeting prepares the leaders for the scheduled activities that may affect their teams; however, typically many more activities occurs during the day than presented, and sometimes, even more than anticipated. The day at the plant never ends, because there is 24/7 movement, 365 days a year. When evaluating error cause data, we must place ourselves into a typical day at the plant.

Therefore, the continuous nature of plant operation enables us to consider that the error rate (the frequency of the 21 causes presented in Table 1-1) can be considered a probability. It may not be necessary to divide the number of errors by the number of opportunities for the specific error, because there is constant and continuous plant and organizational opportunity for these types of errors. That is, to convert the outputs from the CAP entries to frequencies, each data source is normalized (e.g., based on time). Over the course of a year, tens of thousands of CAP entries can be entered into the database.

$$F(\text{cause}) = \text{ Number of CAP entries involving the cause per year} \tag{1-1}$$

The probability of the cause can be calculated as follows:

$$P(\text{cause}) = \frac{\text{Number of CAP entries involving the cause per year}}{\text{Total number of CAP entries per year}} \tag{1-2}$$

This representation of the probability of the cause was later found to be consistent with the work by Pence, Mohaghegha, Kee, Hubenak, Billings, and Dang (2015).

This being the actual case for operating nuclear power plants, using equation 1-2 the probabilities result in a range from .01 to .0005, which coincide with these types of Human Error Probabilities (HEPs) in THERP (Swain, 1983). For example, slip or lapse (WP1) has a probability of .0015 from Figure 1-5 (for all the years) and a probability of .001 in THERP, Table 20-6, entry 1.

The calculation is as follows:

The rate was calculated per day for Figures 1-1, 1-2, 1-3:

Rate WP1 for SCCR = 11 WP5/7 years*365 = .0043/day

P(WP5) = 11/7 yrs/7071/7 yrs = .0015

In order for the error rate to be used as a probability, the rate should be constant over any time period. This was tested by comparing the rates over different time periods and the results are shown in Figure 1-5, showing a relative coincidence for each of the causes over the entire time period, and taken at intervals of two and three years. Thus, the data was verified for these time periods to show the constant rates during any time period.

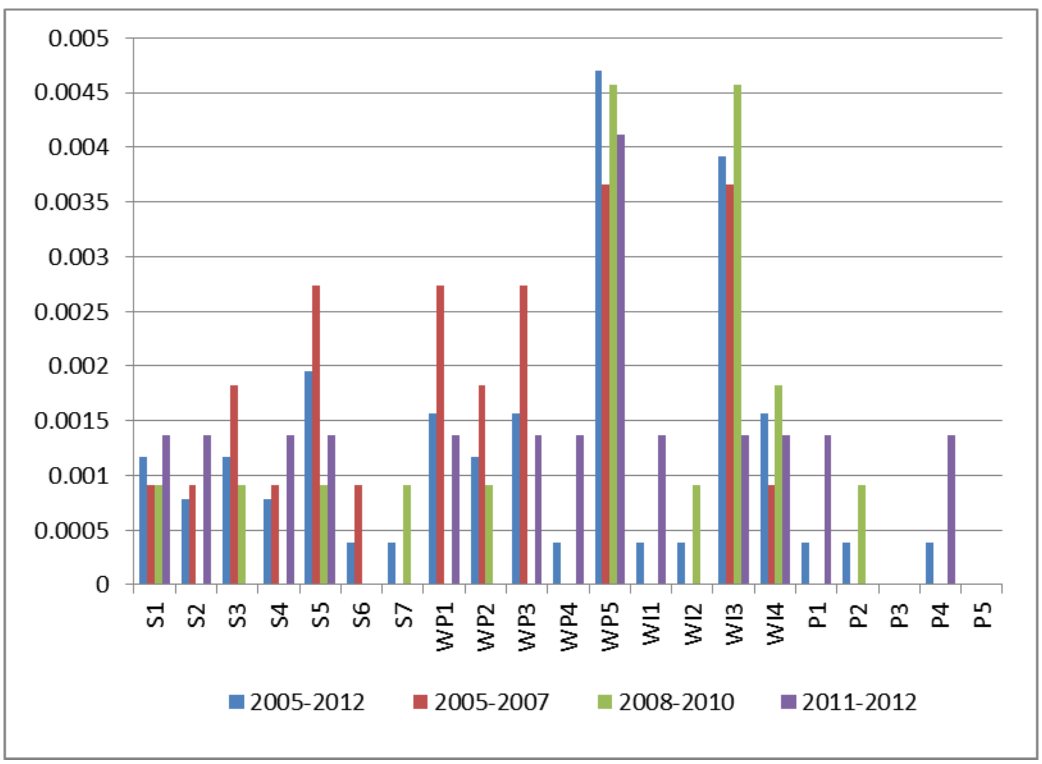

Figure 1-5 Comparison of error rates for several time periods.

Another result from the CRs can be seen in Table 1-2, where the combinations of the causes of the SCCRs are presented. There are combinations of from one to six causes for each of the events. The frequencies of the combinations are shown; each of which is taken from the number of times the combination appears in the SCCR events divided by the 7 years (2500 days). That is, generally the combination occurs once, except in the case that it occurs twice (WI4, WP5), or five times as in the cases of (WI3) and (WP5). The values for the combination of the first order can be compared to those in Figure 1-3, and it can be mentioned that they usually have a smaller frequency since in the figure the total number of appearances are counted and no differentiation is made if it occurs with another cause or not.

Table 1-2 Frequency of Cause Combinations for each SCCR Event.

Cause combinations						**Rate of combo**	**Rate from** Figure 1-3
S1	S5					.0004	
S1						.0004	.0012
S1	S2	S5	WP1			.0004	
S2	S5	WP1	WP2	WP3	WP5	.0004	
S3						.0004	.0012
S3	WI3					.0004	
S3	P2					.0004	
S4	WP5					.0004	
S4	WI4	WP3				.0004	
S5						.0004	.0019
S5	WI3	WP2	WP5			.0004	
S6	WP3					.0004	
S7						.0004	.0004
P1	WI1	WI3				.0004	
P4						.0004	.0004
WI2						.0004	.0004
WI3	WP2					.0004	
WI3						.002	.0039
WI3	WP1					.0004	
WI4	WP1					.0004	
WI4	WP5					.0008	
WP3	WP5					.0004	
WP4	WP5					.0004	
WP5						.002	.0043

This brings us to the concept of organizational reliability analysis and Reason's Swiss cheese model, first published by Reason (1990), an adapted version is shown in Figure 1-6. Accidents in complex systems occur through the accumulation of multiple factors and failures. J. Reason has famously developed a model based on the Swiss Cheese Metaphor that suggests multiple contributors (the holes in cheese slices) must be aligned for any adverse events to occur. Barriers in a system (the slices themselves) are intended to prevent errors that result in these adverse events. This Swiss cheese model is not without drawbacks, and is not accepted uncritically. With use over time even the author has acknowledged its limitations. Nevertheless it remains widely used and is employed as the main basis for new method development.

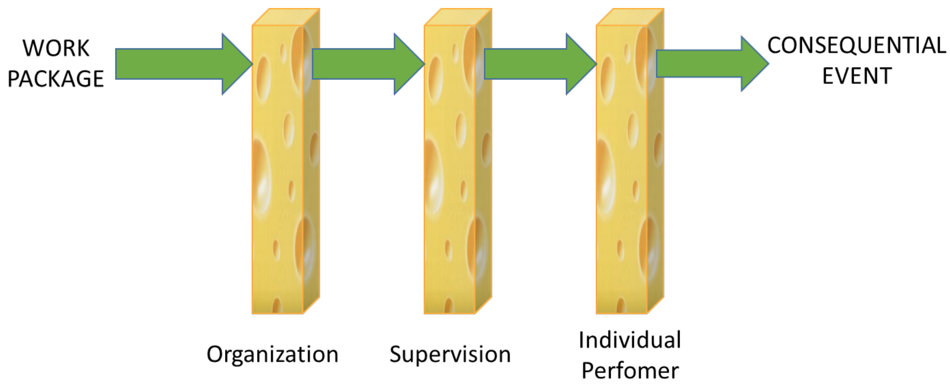

Figure 1-6 Representation of defenses in the organization, supervision and performer.

James Reason hypothesizes that most accidents can be traced to one or more of four levels of failure: Organizational influences, unsafe supervision, preconditions for unsafe acts, and the unsafe acts themselves. In this model, an organization's defenses against failure are modeled as a series of barriers, with individual weaknesses in individual parts of the system, and are continually varying in size and position. The system as a whole allows failures when all individual barrier weaknesses align, permitting "a trajectory of accident opportunity", so that a hazard passes through holes in all of the defenses, leading to a failure (Trepess, 2003). The model includes, in the causal sequence of human failures that leads to an accident or an error, both active failures and latent failures. Active failures encompass the unsafe acts that can be directly linked to an accident, such as turning off emergency cooling when it is needed. Latent failures are particularly useful in the process of nuclear power accident investigation as they encourage the study of contributory factors in the system that may have lain dormant for a long time (days, weeks, or months) until they finally contributed to the accident.

In order for an event to become consequential, we would expect to find at least one organizational breakdown or error; one supervision error, including lack of supervision, and one performer error in each significant event, as depicted in Figure 1-6. But rather what we find are the combinations of errors, as depicted in Table 1-2. This tells us that some of these causes either cover several defenses or the defenses in the other “slices of cheese” are already flawed.

We can assume that the “S” causes correspond to the supervision slice, the “WI” to the organization and the “WP” to the individual (human) performer. The P causes seem to correspond to both organizational deficiencies and human deficiencies. An example of this is the fact that we find the combination of S3 and P2. Clearly S3 corresponds to the supervision slice, thus P2 corresponds to both the organization and performer slices. This can be exemplified through an example of a reactor trip caused during a reactor trip breaker surveillance by following the procedure in the wrong order and there was flawed supervision. Thus, P2 can be seen as the intersection of Organization and Human performer defense breakdown. Another example is when WP5 causes an event, all by itself. This may mean that this failure to use knowledge, skills and abilities of the performer corresponds really to failure in the three defenses, or holes are opened and aligned in these cases.

1.3 Preliminary Results

The preliminary results of statistical data treatment show that the use of Condition Report information has the possibility of enabling the update of THERP data as well as providing insights into organizational error analysis. Because activities are ongoing at an NPP, it may possible to substantiate that the error rates can be considered probabilities and in this way compared to or used to complement the THERP data. This research was limited in the number of CAP databases reviewed and the process should be repeated many more times to enable the creation of HEPs for use in HRA. The correlation of causes to consequential events was examined through the relation of cause combinations to consequential events and through this, it was determined that the cause set represents breakdowns in the organization, supervision and human performer. Due to these organizational dependencies, it became clear that these dependencies must be explicitly incorporated into the model that will be used to quantify the HEPs, and into the organizational error analysis in general.

In the effort to design a method to standardize events and cause codes, to be used to register human error and to provide insight into the human performance tracking, trending and performance indicators to support safety culture at an NPP, the results could also be used to enable the calculation of up-to-date HEPs, which can then be used in PRAs. The CAP data can also be used for nuclear knowledge management, which is discussed in the next section.

1.4 Use of Corrective Action Programs for Knowledge Management

Almost every organization manages data in some way. Data is a major corporate resource; however it is frequently poorly documented. Descriptions of data and other resources are metadata, which are part of the corporate memory for the organization, and preserving corporate memory is one of the basic features of knowledge management. At present, many countries are experiencing a large percentage of the personnel at nuclear power plants reaching retirement age. It is estimated that in the next 5 years 50% of the existing workforce in the U.S. will reach retirement age. As a result, recording the experiences of these workers, including the history behind the data, is increasingly important (i.e., knowledge retention). Preserving metadata is crucial for understanding data years after the events, lessons learned and the data were created. Human error or human performance is of special interest in any industry, and the nuclear industry has developed methods for performing human reliability analysis to calculate the contribution of human error to accidents. There have been attempts to collect data to inform quantification in HRA, starting with the work done for the THERP methodology by Swain. Many experts in HRA have related the opinion that there should have been more effort on collecting human error data for the purpose of quantifying the probabilities of human error; however, there does exist a wealth of information in the Condition Reports contained in the Corrective Action Programs at most nuclear power plants.

In fact, in the nuclear industry, the Corrective Action Programs are a source of this metadata. The information contained in the Condition Reports (CR) at every nuclear power plant is invaluable and while the reports for each event may be pages long, there needs to be an efficient way to record it, store it and have it provide data and feedback continually in order to support analyses, trending, and pertinent human performance indicators. The statistics of the data can tell us much about the trends in failures, whether system or human failures; however, if the information is not codified to work for the intended database, the results may be uncertain and inaccurate. This dissertation describes the methodology to extract benefit from the root and apparent cause analysis done on any abnormal occurrence at a nuclear plant, and presents a model to include this information in a process that furthers the knowledge management about consequential human errors in nuclear power plants.

The CR cause coding together with study of written CR descriptions in the failure and repair work orders helps to identify events of human errors related to maintenance activities from the failure and maintenance history. From a sample containing thousands of human performance related labeled condition reports it was possible to define a model of the factors that influence the occurrence of these events. The model provides the structure of the information necessary in the database to be able to better utilize it for conserving knowledge and lessons learned for future generations. This should be done in such a way as to facilitate the processing of the data to become continual input for the model, which was originally developed from the existing data. Once validated, the model can serve to estimate human failure rates for organizational processes (e.g., procedure adherence), predict the rate at which events may occur, as well as risk-inform procedures that require the manipulation of important safety systems to identify risk significant steps within the procedures, apply appropriate compensatory measures to reduce the reoccurrence of the events, as well as avoid future consequential events.

Data mining is the process of extracting nontrivial and potentially useful information, or knowledge, from the enormous data sets. These can be historical records, such as those contained in the Corrective Action Programs (CAP) at nuclear power plants.

1.5 Other Sources of Data

Several of the Human Reliability Analysis (HRA) methods, including THERP, have made attempts to collect data to inform quantification in HRA. These efforts have continued to the present time, with efforts like the US Nuclear Regulatory Commission's Human Event Repository and Analysis (HERA) system (Hallbert, Whaley, Boring, McCabe & Lois, 2007), a human performance database developed for NASA (Boring, Gertman, Whaley, Joe, Tran, Keller, Thompson, & Gerszewski, 2006), or the UK's CORE Database (Kirwan, Basra, & Taylor-Adams, 1997). However, as the primary author of THERP, Alan Swain, has noted in reference to the development of quantification in THERP (Boring, 2012):

> *"There should have been many more changes had the research been done to develop more hard data on HEPs for human tasks. That failure has been a disappointment to me. ... I always said that the data tables in [THERP] were not written in stone, and I was not Moses coming down from a mountain with these tables so inscribed."*

However, even if we were to be able to derive the quantitative HRA data, which focus entirely on the HEP, this does not necessarily provide the information about the performance measures desired by human factors. For example, knowing the error likelihood does not actually tell the human factors researcher or practitioner the expected performance or the level of performance degradation that may precede an actual error.

It is the opinion of the researcher that in order to obtain the information necessary to quantify the human errors and obtain the performance measures necessary to identify risk significant process steps for frequently performed activities (e.g., surveillance procedures) and interpret the degraded defenses at the nuclear plant, it is necessary to study and understand the events that actually occur at the plants. In the nuclear industry, the Nuclear Regulatory Commission requires a Problem Identification and Reporting program. Compliance with this requirement is performed through station specific Corrective Action Programs (CAP). These Corrective Action Programs generally use a reporting mechanism that is available to the general station employee population to identify and record problems, issues, or actions that need to be performed to accomplish the station's business and operational missions. Typically, a standard Condition Report (CR) form is used to document and enter this information into station databases. Thus, the CRs are significant and objective source of events and metadata. The information contained in the Condition reports is invaluable and while the reports can be pages long and be highly variable from CR to CR due to the many "authors" at a plant for a CR, it is important that the CR information be processed and evaluated to generate important data and analyses relative to plant and human performance. This offers a significant opportunity to associate consequential station events to those processes and activities that were being performed by station personnel and the time of the event as well as their associated causes. This provides important opportunities to develop human performance models from objective plant specific data that has the potential to reveal organizational weaknesses and those station activities with risk significant relative to consequential human failures (not just core damaging events). There needs to be a process model that provides an efficient and consistent way to record the data, store it, and have it provide the basis for follow-on technical analyses related to human performance. As described in the previous sections, the statistics of the data can tell us much about the trends in failures, whether system or human failures; however, if the information is not codified to work for the intended database, the results may be inaccurate and uncertain. For this reason, this next section describes the review of the coding at a pilot plant and the plan to develop the database fields necessary to fill the gap in the data process.

1.5.1 More on the Pilot Database

If we are to use the information from the CAP programs at the nuclear stations, it is necessary to continue to examine its content, and determine what is useful, as well as what is needed to be able to fully exploit the existing information as well as create the manner in which it can be most useful in the future to represent all that goes on at the plant in terms of errors and their causes, as well as the defense breakdowns.

Each plant codes their reporting system in their own way: one plant has 61 cause codes and 70 event codes, another uses 385 cause codes and 315 event codes, another plant's CAP program uses even more cause codes. For this report, the database with the 61 cause codes has been studied to develop the model from the data, which is presented in Chapter 3.

From the pilot plant 1 there are more than 121,000 events, 107 events from 2005-2014 that were categorized as SCAQ[2] and labeled SCCR in the previous sections. Of these events, we can group them in the following activities:

- Post maintenance testing
- Surveillance testing
- Maintenance Instrumentation & Calibration (I&C)
- Mechanical maintenance
- Other operational challenge
- Outage
- Tech spec compliance issue
- Latent condition
- Preventive maintenance
- Repeat maintenance

The equipment types involved in single human errors in relation to maintenance activities are grouped in the following manner.

Activity	**Quantity**
I&C	7
Electrical	5
Mechanical	8
Valves	3
Valve instrumentation	2
Other	7

The operating states during the activity and at detection should be identified as one of the following: power operation, refueling outage prior to start-up, hot shutdown, cold shutdown, plant start-up, plant shutdown.

Distribution of the detection activity types of human common cause failures:

- Functional testing
- Preventive maintenance
- Periodic testing
- Alarm
- Shift walkaround

[2] Note: Significant Condition Adverse to Quality – a condition adverse to quality that, if uncorrected, could have a serious effect on safety or operability. (Based on ASME NQA-1-1994, Part 1, Section 1, Introduction.) [NEI-08-02 rev3].

- Repair

It is important to observe when the errors occur and when they were detected, which provides knowledge about the adequacy of the built in defenses at the station. It is also necessary to record the severity of errors; for example, mechanical errors may be more frequent, but I&C may be more severe, causing more plant outages, for example.

1.6 The New Database Fields

As Johnson, C.W. (1999) mentions in his work on accident reporting,

> "*accident and incident reports are the primary means of ensuring future safety in many industries. It is, therefore, surprising that so little attention has been paid to the format and presentation of these documents.*"

A new database Excel format will be used to do several calculations that are necessary to enhancing the CR reporting in such a way as to be able to feed into the model built from the data. Thus, there is a need to develop a process to capture existing CRs in this format and then continue recording it this way. The fields are:

- Date
- CR#
- Organization
- Equipment
- Tag TPNS
- Error occurred
- Procedure
- Plant state
- Detection activity
- Detection plant state
- Factor
- PRA
- End state
- Cause

From the equipment type, a pie graph indicates where the errors occur.

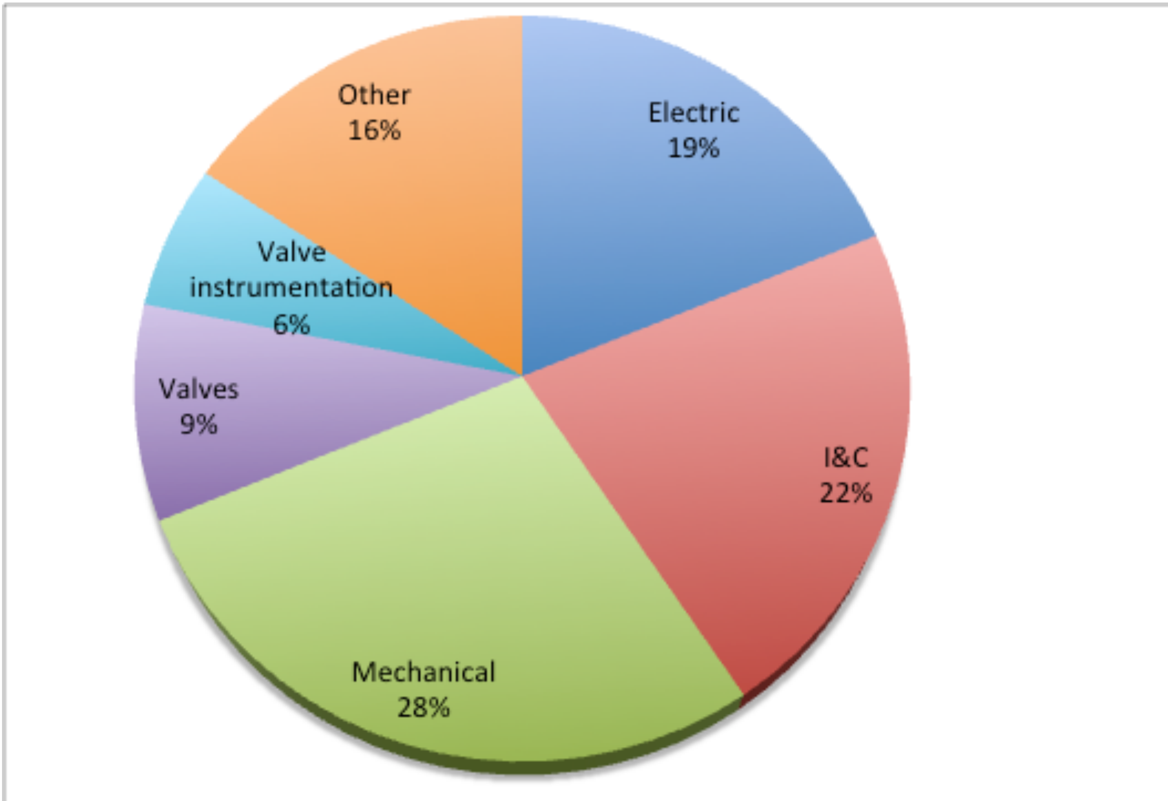

Figure 1-7 Equipment types involved in human errors.

The activity type describes where the errors occur. The column of component tags (i.e., equipment identifier) should be searched for repeat events. The same is done for procedures, if procedures are repeated, this should be noted in order to go to that procedure to see what the problem could be. This is in the case that the cause is categorized as P (for Procedure related cause code). Also, even if the procedure is not repeated, where the problem is in the procedure and put in a note as a corrective action, although this is already done at plants, it should be identified as a field in the new database. A very important aspect to learning from operational experience is the detection of the error WHEN and HOW is the error detected are important data to have and provide insights into the effectiveness of the existing defenses or barriers, or the non effectiveness, and the need for improvements in this area. These points are addressed in Chapter 3.

1.7 Organizational Factors

This section presents the results of the examination of the organizations (departments at the station) and their responsibility for the events reported in the CAP program. This section continues to delve into the information looking to use it in any way that enables enriching the knowledge obtained; that is, graphing different parts of the information and conducting statistical analysis to look for ways to model the data in ways useful for use in prediction as well as trending.

Figure 1-8 presents the distribution of Condition Reports (CRs) among the departments for the most significant events that were generated in a 7-year period and Figure 1-9 shows the CRs among the 87 departments at the plant for all the CRs. These represent a total of 7024 CRs. The Department of Health Physics has 1290 CRs, but is truncated at 600 in order to avoid having to scale the results (due to the large number of corresponding CRs in the less significant events). In fact, there are no significant CRs assigned to the Health Physics Department.

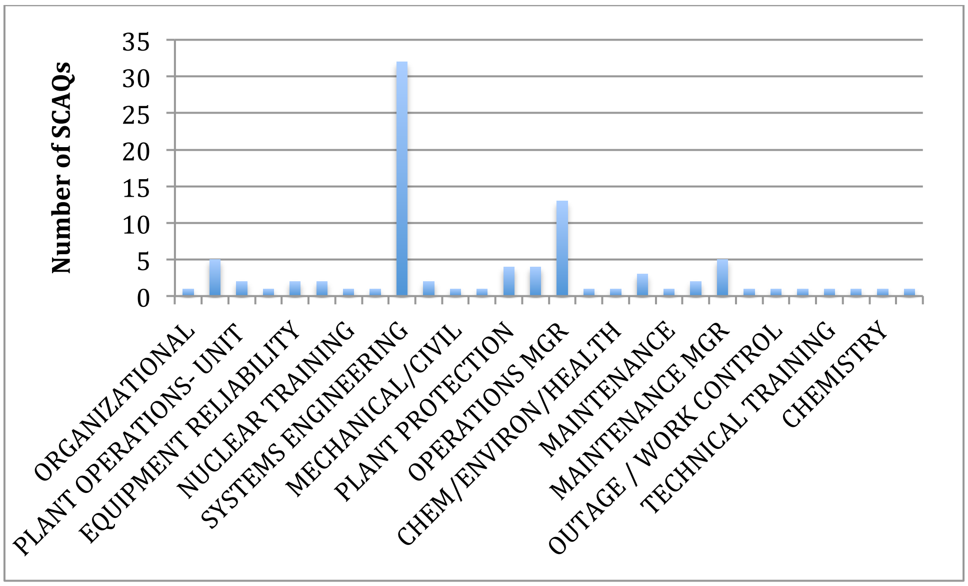

Figure 1-8 Number of significant CRs per department (7 years).

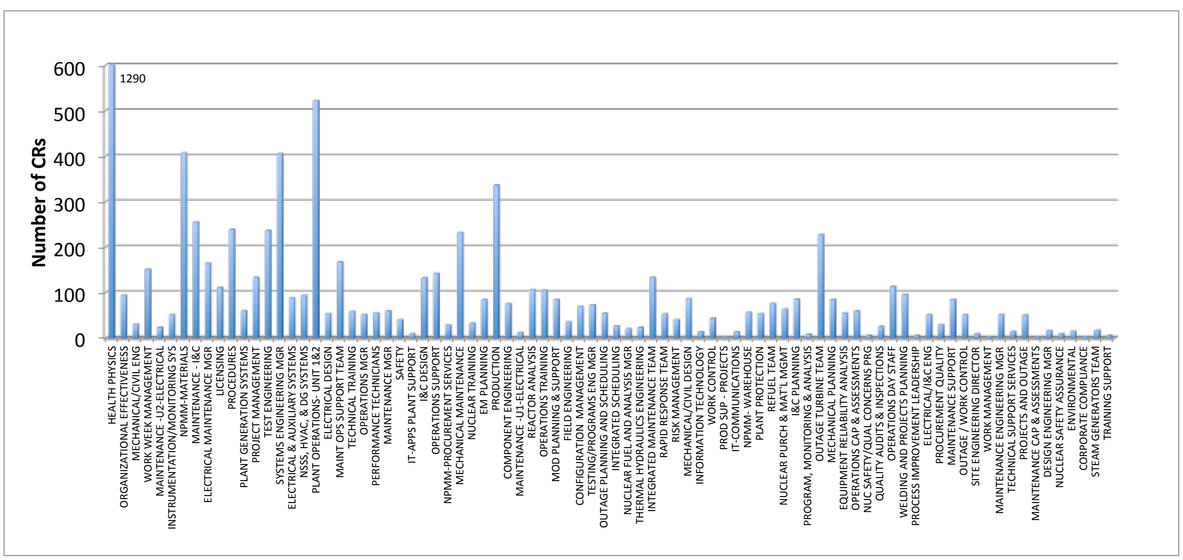

Figure 1-9 Number of CRs per department (7 years).

Table 1-3 contains the definitions of the causes not defined previously in Table 1-1.

Table 1-3 Extended Generic Failure Types and their Causes.

General Failure Type	Causes
Design	DE1: Acceptable Failure (Run to Failure) DE2: Design Inadequacy DE3: Manufacturer Material/Fabrication Deficiency ▪ *The vendor supplied equipment that was deficient caused the failure or condition* ▪ *Equipment condition or failure was the result of a manufacturing process defect* ▪ *Equipment condition or failure was the result of a fabrication deficiency* DE4: Improper Component/Material Selection for Application DE5: Equipment Performance Monitoring/Tracking Inadequacy ▪ *Failure to identify or take action on an adverse trend in performance* DE6: Predictive Maintenance Program Inadequacy ▪ *Equipment was not in the program and should have been* ▪ *Periodicity was inadequate to identify emerging problems* ▪ *Wrong or missing monitored parameters* DE7: Preventive Maintenance Program Inadequacy ▪ *Equipment was not in the program and should have been* ▪ *Periodicity or scope of PMs were inadequate to prevent problems* ▪ *Wrong or missing monitored parameters* DE8: Random Electronic Component Failure DE9: Design Inadequacy Outside of Design Control DE10: Preventative Maintenance Program Inadequacy
External	EX1: Weather Related EX2: Human Performance Beyond plant Control ▪ *Theft of station resources by personnel who are not station employees* ▪ *Error by vendor/manufacturer that was not working under plant oversight* EX3: Other factors beyond plant control
Human Factors	HF1: Human Factors Not Properly Addressed in Work

	Area/Equipment ▪ *Workplace design or layout does not take human limitations into account* ▪ *Inadequate lighting, alarms with insufficient volume, controls that give no feedback when actuated, etc.* HF2: Identification Method/Labeling Missing or Inadequate HF3: Required Equipment/Controls Inoperable/Unavailable HF4: Distractive/Uncomfortable Work Environment ▪ *Excessive heat or cold, noisy, cluttered work area, ust, etc.*
Leadership/ Supervision	LS: Same as S in Table 1-1.
Management Assessment/Corrective Action	MA1 Organization Not Sufficiently Self-Critical ▪ *Failing to perform self-assessments and/or not utilizing their results to improve* MA2 Cause Analysis for Known Problem Inadequate ▪ *The underlying cause was misidentified for a known problem* MA3 Corrective Action for Known Problem Untimely/Inadequate MA4 Industry/Vendor Information Not Used to Prevent Problem ▪ *External information was not proactively used to prevent similar problems at STP* MA5 Need to Improve Program/Process Not Identified ▪ *Program or process was considered adequate until an event identified that additional controls or enhancements were required* ▪ *Failure to recognize indications of poor or declining performance* MA6 Industry/Vendor Information Not Used to Minimize Recurrence/Mitigate Event Impact
MC: Change Management	MC1: Need for Change Not Recognized MC2: Change Not Implemented in a Timely Manner ▪ *No schedule for change or change not maintained on schedule (if change is due to a corrective action then preferentially use code MA3)* MC3: Impact of Change Not Properly Determined Prior to Implementation ▪ *No change management tools or principles used for*

	planning implementation ▪ *All affected working groups not apprised of/involved in change* ▪ *Procedures or documents impacted by change not revised in conjunction with change* ▪ *Change produced unanticipated/undesirable consequences*
Management Practices	MP1: Communication Within an Organization Inadequate/Untimely MP2: Communication Between Organizations Inadequate/Untimely MP3: Management Practices Promote/Allow Unacceptable Behaviors ▪ *Coaching- failing to identify and/or correct unacceptable habits/behaviors* ▪ *Coaching- no reward or positive acknowledgement of accepted/desirable behaviors* MP4: Human Performance Tools Inadequate or Not Provided MP5: Human Performance Monitoring/Tracking Inadequacy ▪ *Performance Indicators- failing to use quantifiable indicators of human performance* ▪ *Insufficient contact time with personnel* MP6: Management Expectations are Unclear/Inadequate/Not Defined ▪ *Personnel do not understand what is desired, how it is important, or what the expected level of performance is* ▪ *Conflicting expectations* MP7: Management Expectations Not Reflected in Process/Program/Procedure ▪ *Expectations are clear but conflict with how the process actually works* ▪ *Alignment- expectations contribute to a lack of alignment between the different parts of the process* MP8: Supervisor Capabilities Not Matched with Task Demands ▪ *Failure to ensure proper defenses were in place to ensure successful task performance by assigned personnel (TWIN Analysis)* MP9: Insufficient Time Allotted for Supervisor to Perform Task

	MP10: Economic Decisions
Resource Management	MR1: Manpower Inadequate MR2: Budget/Funding Inadequate MR3: Prioritization/Scheduling of Activities Inadequate (Management Level) ▪ *Failure to organize activities by relative importance* ▪ *Failure to understand the urgency of an activity* ▪ *Failure to properly deal with conflicting priorities or activities*
Procedure adherence	PA: Same as P in Table 1-1
TR: Training	TR1: Training Content Inadequate TR2: Method of Instruction/Presentation Inadequate TR3: Necessary Initial/Refresher Training Not Provided ▪ *Initial/continuing training does not cover task/situation* ▪ *Periodicity of training insufficient to maintain task proficiency* ▪ *Qualified personnel do not possess proficiency with task assumed to be "skill-of-the-craft"* TR4: Assessment of Task Proficiency Inadequate ▪ *Testing or practical exercise was either not performed or was insufficient to ensure student was adequately trained* TR5: Simulator or Mockup Fidelity Inadequate ▪ *Failure to accurately model the actual condition of the plant or equipment*

Figure 1-10 presents the distribution of CR causes of SCAQs in the period, including the management causes, which were not included in the graphs in Section 1.1. The largest contribution is from the MA5, which is the *Need to Improve Program/Process Not Identified.* This includes the fact that the program or process was considered adequate until an event occurred that identified that additional controls or enhancements were required or that there was the failure to recognize indications of poor or declining performance. Chapter 3 presents a model that was developed using more complete information about the causes of the events.

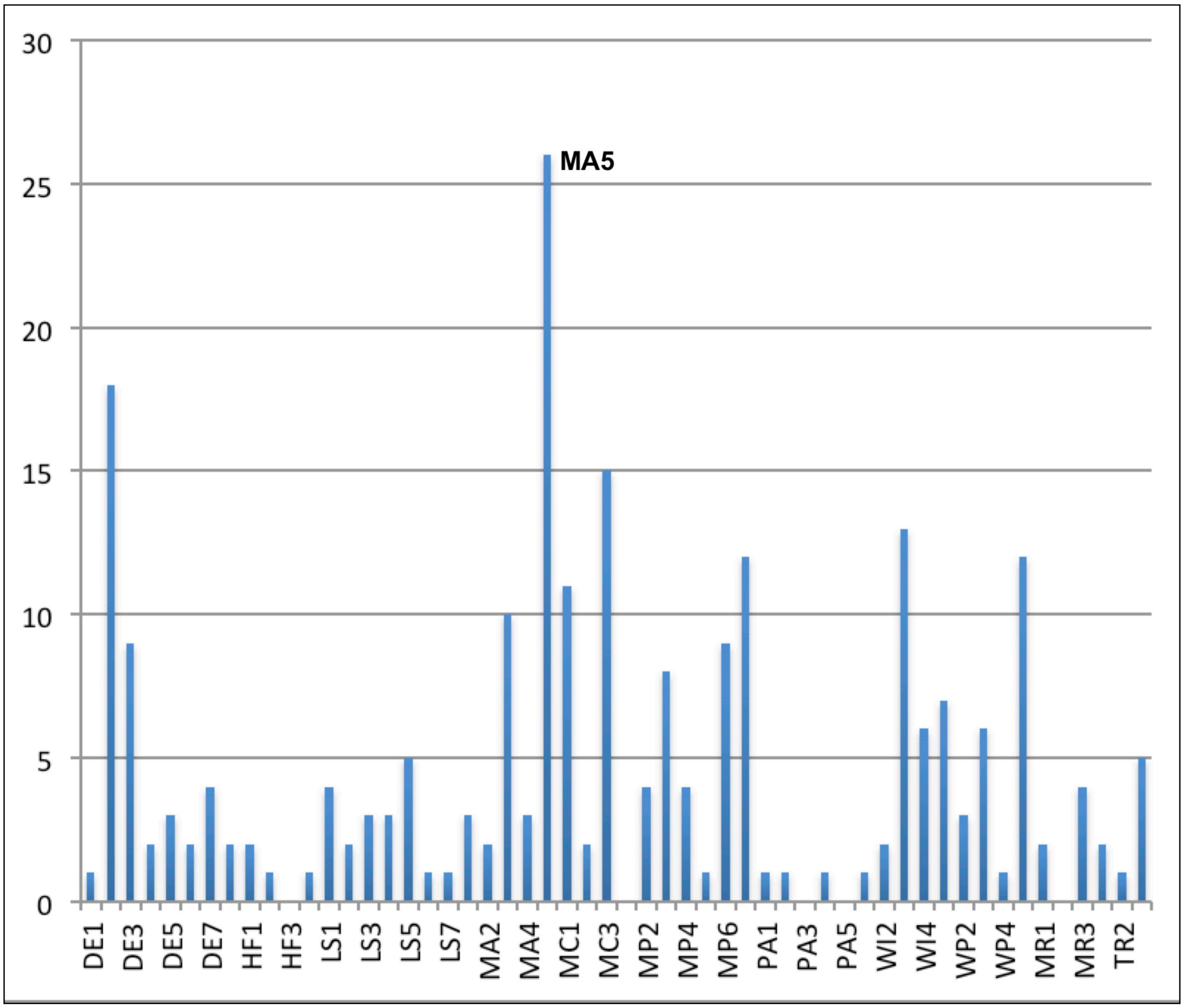

Figure 1-10 Distribution of causes of significant events (2005-2012).

Chapter 2 Resilience and Leading Performance Indicator

The development of operational performance indicators is of utmost importance for nuclear power plants, since they measure, track and trend plant operation. Leading indicators are ideal for reducing the likelihood of consequential events. This chapter describes the operational data analysis of the information contained in the Corrective Action Program. The methodology considers human error and organizational factors because of their large contribution to consequential events. The results include a tool developed from the data to be used for the identification, prediction and reduction of the likelihood of significant consequential events. This tool is based on the resilience curve that was built from the plant's operational data. The stress is described by the number of unresolved Condition Reports. The strain is represented by the number of preventive maintenance tasks and other periodic work activities (i.e., baseline activities), as well as, closing open corrective actions assigned to different departments to resolve the Condition Reports (i.e., corrective action workload). Beyond the identified resilience threshold, the stress exceeds the station's ability to operate successfully and there is an increased likelihood that a consequential event will occur. A performance indicator is proposed to reduce the likelihood of consequential events at nuclear power plants.

Every nuclear power station is subject to daily organizational stresses, which result from the cumulative strain of routine operation of the plant, maintaining regulatory and operating requirements, and supporting long-term reliable operations. In addition, operational conditions periodically change in order to refuel safely, perform shutdown maintenance activities, and restart for another cycle. The impact of these strains varies depending upon the age of the plant. One must also consider unexpected operational events that result in work that goes beyond normal plant operations, regulatory compliance, and typical maintenance activities. These conditions result in periods of time when individual and organizational workloads increase significantly, raising the likelihood of errors, which in turn, further increase personnel workloads.

Safety culture emphasizes the importance of developing and maintaining a strong Problem Identification and Resolution Program (NRC, 2015a), typically referred to as a Corrective Action Program (CAP) where all incidents, risk significant or not, are to be reported. The term 'safety culture' was first used in INSAG's 1988 'Summary Report on the Post-Accident Review Meeting on the Chernobyl Accident,' where it is described as "that assembly of characteristics and attitudes in organizations and individuals which establishes that, as an overriding priority, nuclear plant safety issues receive the attention warranted by their significance" (IAEA, 1988). All nuclear power stations in the US have a Problem Identification and Resolution Program as required by regulation.

A plant's CAP is provided to employees, who use it to identify problems or issues and to record them in a problem report, formally known as a Condition Report (CR). The events that trigger these reports serve as sources of organizational stress, as they represent additional scopes of work beyond those required for maintaining regulatory compliance and reliable plant operation. Increasing numbers of CRs accompanied by CRs with high severity levels indicate that organizational resilience levels are being exceeded. Here we define resilience as the intrinsic ability of an organization to adjust its functioning prior to, during, or following changes and disturbances, in order to sustain required operations for the current conditions of the plant (Hollnagel, 2011).

Some Condition Reporting programs are considered “low-level,” as the threshold required for generating a CR is very minor (e.g., editorial errors in procedures or minor errors in design drawings). Low-level CR programs are characterized by having high levels of granularity as criteria for the identification of a situation requiring the generation of a Condition Report (i.e., thousands of items are identified in a single year covering virtually all plant organizations). Alternatively, some Condition Reporting programs are considered “high-level,” as generation of a Condition Report must meet a certain, high criteria (e.g., only plant hardware issues are considered). Generally, most US plants are characterized as low-level Condition Reporting programs, such that each typically generates in excess of ten thousand CRs each year.

The fact that even minor incidents reported in low-level Condition Reporting programs can combine with others and cause an accident brings forward the concept of high reliability organizations (HROs), which include nuclear power-generation plants, naval aircraft carriers, air traffic control systems, and space shuttles. Studies of HROs have challenged the postulations of Perrow’s Normal Accident Theory (Perrow, 1984), in which he insists that “normal” or system accidents are inevitable in extremely complex systems. He states that given the characteristics of the system involved, multiple failures that interact with each other will occur, despite efforts to avoid them. He continues to say that operator error is a very common problem, many failures relate to organizations rather than technology, and big accidents almost always have very small beginnings. Such events appear trivial to begin with before unpredictably cascading through the system to create a large event with severe consequences.

HROs, and specifically nuclear power plants (NPPs), are complex, but have nonetheless maintained exceptional safety records over a long period of time. According to Weick, Sutcliffe and Obstfeld (2008), HROs are learning organizations characterized by a set of cognitive practices that enable people to work safely and eventually create mindfulness and reliability. These practices involve constantly tracking and investigating small errors, resisting oversimplification, sensitivity towards current operations and committing to resilience.

HRO research can be said to represent a focal shift in safety research, from a focus on failure to a focus on success. The HRO perspective represents a valuable addition to safety research, and we believe that combining the HRO perspective with data that is readily available, specifically from the CRs contained in the CAP database, provides the necessary elements to produce a resilience curve and associated resilience threshold. This can be applied at nuclear power plants in order to identify areas where human errors are more likely to result in consequential events, reduce human error rates, consider organizational interaction factors, and develop a leading performance indicator.

The application of resilience engineering is relatively new to the nuclear industry, but it has been used in general aviation, offshore oil and gas production, safety science, and healthcare, among others, and it has provided a substantial body of knowledge and experience (Sutcliffe, K.M. & Vogus, T.J., 2003, Hollnagel, E. 2006, Woods, D. & Leveson, N., 2006, Herrera, I.A. & Hovden, J., 2008, Hollnagel, E., 2010, Woods, D., Chan, Y. & Wreathall, J., 2013). In particular, Woods compared the demand-stretch model of an organization with the stress-strain curve and resilience property from materials science (Woods, D., Chan, Y. & Wreathall, J., 2013). This prior work is largely qualitative, whereas here we present a quantitative application.

Section 2 describes the data used. Section 3 identifies the sources of stress and strain and presents the methodology used to develop the resilience model. Section 4 presents the resulting organizational resilience curve and threshold. Section 5 shows the application of the resilience threshold to develop a leading performance indicator to predict situations where the likelihood of consequential events is increased. Section 6 contains the conclusions and describes future work.

2.1 Nuclear Power Plant Operational Data

We propose the use of the CAP database to evaluate human and organizational performance. Other studies have examined Licensee Event Reports (LERs), to evaluate human performance, types of events, etc. (Šimić, Zerger, & Banov, 2015; Groth, 2009; Hallbert, 2006). These studies provide valuable ways of looking at the historical events. We believe the inclusion of all plant specific events (LERs plus all the other events reported in the CRs) increases the statistical validity of the data and enables the specific and detailed study of a plant's operating experience and organizational behaviors.

In this study, the CAP database from an operating plant was analyzed to test the database's ability to yield measurable results with regard to assessing organizational resilience. Ten years of CRs (2005-2014) were analyzed, yielding not only interesting tendencies and insight into resilience, but also a basis for the construction of leading organizational performance indicators at NPPs.

In order to begin to understand the information contained in the Condition Reports, as well as the complex interdepartmental relationships in HROs such as nuclear power plants, it is necessary to define the most important administrative units, known as organizations, as well as the extent of their responsibilities in everyday activities.

A simplified flow diagram is shown in Figure 2-1, which outlines a typical process used for planning, executing and completing a work package. A work package can be considered an organizational activity directly impacting plant equipment or other hardware. The work package contains the necessary prerequisites, approvals, work steps, and hardware parts (consumables) that will be necessary to complete the activity on a component or set of components. The flow diagram shows the types of activities during which the events that are the focus of this paper occur. That is, when a problem (e.g., unplanned equipment failure) or a necessary work activity (e.g., preventive maintenance activity) is identified there are many opportunities for organizational errors. These errors can occur based on the organizational programs and procedures necessary to authorize and perform work on plant equipment. Since the actions recommended to resolve these errors are combined with other organizational work activities associated with low-level CR programs not directly associated with plant hardware, it can be seen that organizational workloads can vary greatly, as well as be significantly affected by the quantity and scope of CRs and scope of CAP programs.

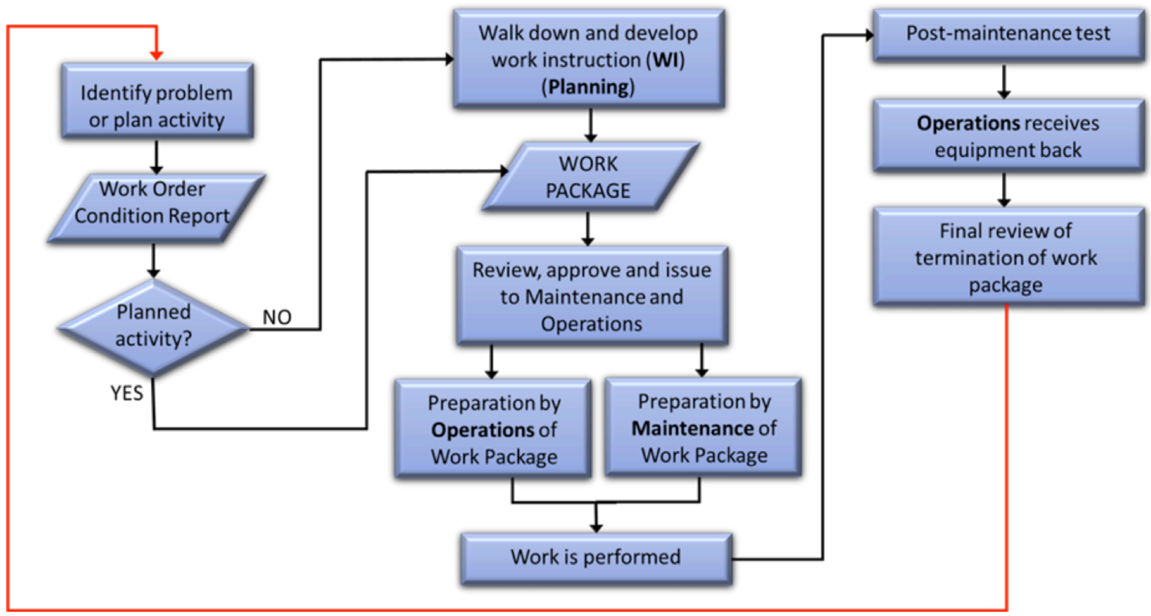

Figure 2-1 Typical organization process flow of work activities at an NPP.

As shown in Figure 2-1, a Work Order (WO) is written to trigger the work process. If the work is emergent or unplanned, a work planner ‘walks down’ the job per the WO and develops draft work instructions, which are then reviewed and finalized. A work package is then prepared and planned. This package is reviewed and approved and is issued to the appropriate maintenance discipline. The package is scheduled per the work scheduling process, and when the scheduled workday arrives the working discipline retrieves the package, gathers parts, materials, tools, etc. and begins the process of completing the activities required and described in the work package. The Operations organization ensures the proper equipment clearance tags are hung so that the equipment to be worked on is properly isolated such that work can be performed safely. Maintenance for the working discipline (e.g., Mechanical, Electrical) begins by obtaining work start approval from Operations (i.e., Operations releases the equipment to Maintenance), a pre-job briefing is typically held between Maintenance and Operations, then the working discipline is released to perform the work. After the work activities are completed, a post maintenance test is performed to ensure the equipment operates correctly and, if the test is passed (i.e., results are acceptable to Operations), then Maintenance releases the equipment back to Operations. Then, if applicable, the work process activities continue to obtain the necessary final reviews and approvals (e.g., engineering reviews) and the package is closed and archived. Follow-on activities include entries made in equipment history logs as well as other monitoring processes (PRA risk profile, Maintenance Rule, Equipment History, etc.).

This organizational process is performed thousands of times during an operating cycle and is also performed during planned and unplanned plant outages. This thesis analyzes the errors that occur during these processes, and demonstrates how this constant tracking becomes the data feedstock to produce methods that can become part of the solution for the plant to minimize similar errors, and most importantly, to avoid consequential outcomes (e.g., plant trip, inadvertent actuations).

As part of the effort to determine the organizational factors that lead to an event (Condition Report), a detailed review of the CAP data made it possible to better understand which plant organizations have greater exposure for consequential errors, given the number of CRs generated that identify that organization as the responsible party for resolving the condition described in the CR. Also, through analysis of the actions that are generated after the occurrence of an event, the creation of the Condition Report and the subsequent investigation, we gain more insight into the total organizational workload and how the organizations work together, or at times, do not work together to produce conditions of low resilience and higher likelihood of consequential events. The time series of the events provides insight into the cyclic behavior, particularly controlled by the outages. This can be used for predictive purposes and is presented in the next Section.

2.1.1 Operational Data Time Series

One way to observe the operational experience at the plant is to plot the events that occur at the plant over time. This graph is presented in Figure 2-2, using data from the operating nuclear power plant. In this graph, the events are plotted by level of severity, the red (SCAQ[3]) representing the most significant contributor to risk, next the gray (condition adverse to quality on an station level, CAQ-L1), and finally the green (condition adverse to quality on an department level, CAQ-L2). Although the more severe (red) events are plotted on an exaggerated scale – on the right side of the graph with between zero and 4 SCAQs a month – this does not detract from the fact that the peaks in number of events frequently coincide for all severity levels. Presumably, we will have more events during cold shutdowns, refueling and outages, because there is an increased amount of maintenance work, more people at the plant, especially contractors, and the peaks in the figure illustrate this.

[3] Significant Condition Adverse to Quality – a condition adverse to quality that, if uncorrected, could have a serious effect on safety or operability (based on ASME NQA-1-1994, Part 1, Section 1, Introduction) NEI-08-02 rev3].

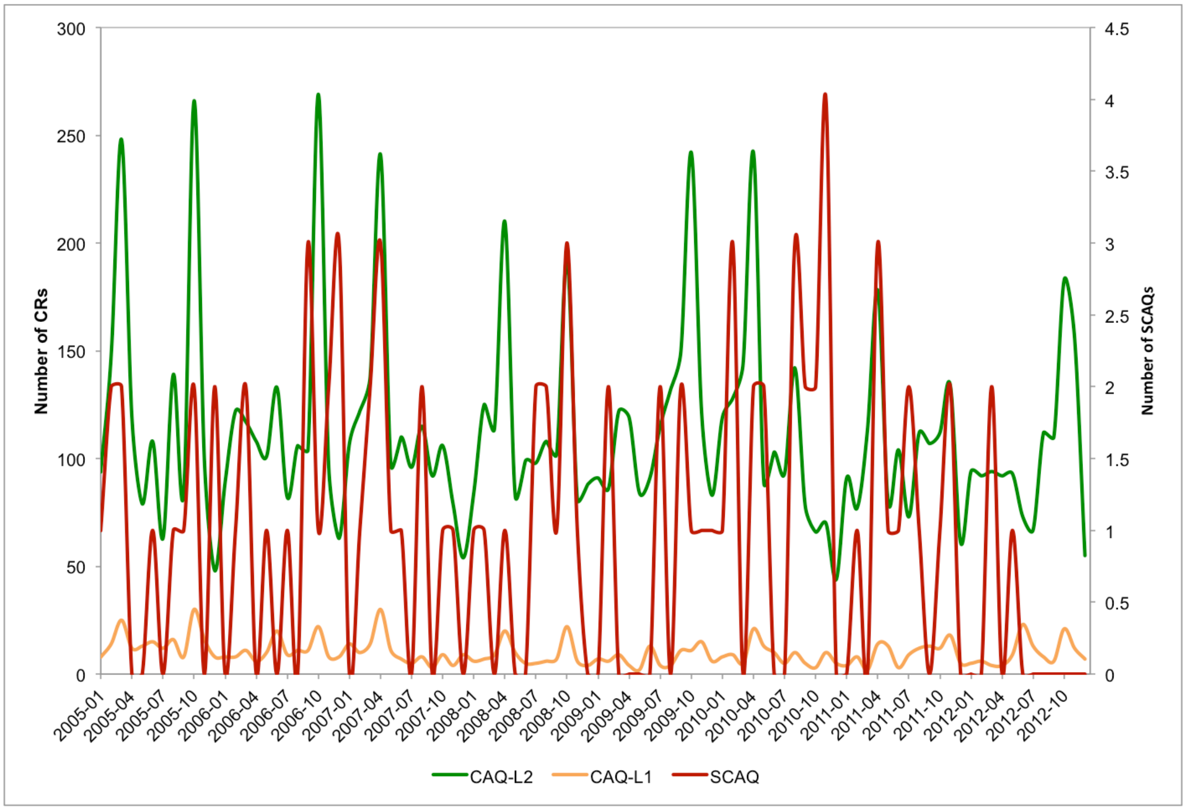

Figure 2-2 Events 2005-2012.

Figure 2-3 plots the events per month, but for only the period 2007-2008, allowing the relationship between the different CR severity levels to be observed in greater detail. In particular, the first and last peaks (April 2007 and October 2008) for this time period show that the peaks of all three severity level CRs coincide. Despite the fact that we see dips (i.e., lower total number of CRs), we can also observe that they, too, generally follow the same trend. In other words, in periods of time where the total number of CRs is low, the three highest severity level CRs are also at minimums. This may seem to be an obvious conclusion; however, the severity level of a single CR is independent of the number of CRs generated. It is determined by predetermined criteria, and therefore a CR's severity level is not related to the absolute number of CRs generated. Thus, based on Figure 2-2 we can conclude that there is a correlation between the number and scope of open CRs and the likelihood of occurrence of a more severe CR, up to and including the most severe, a SCAQ. Also, it is important to mention that even when the red peak (SCAQ) is not above the green (CAQ-L2), we are still seeing significant results, remembering that the scale is different. There may be only one significant event, as in April 2008; however, the three types of events are aligned, occurring simultaneously. This means that as more events of less severity occur, it is more likely it is that significant events may occur.

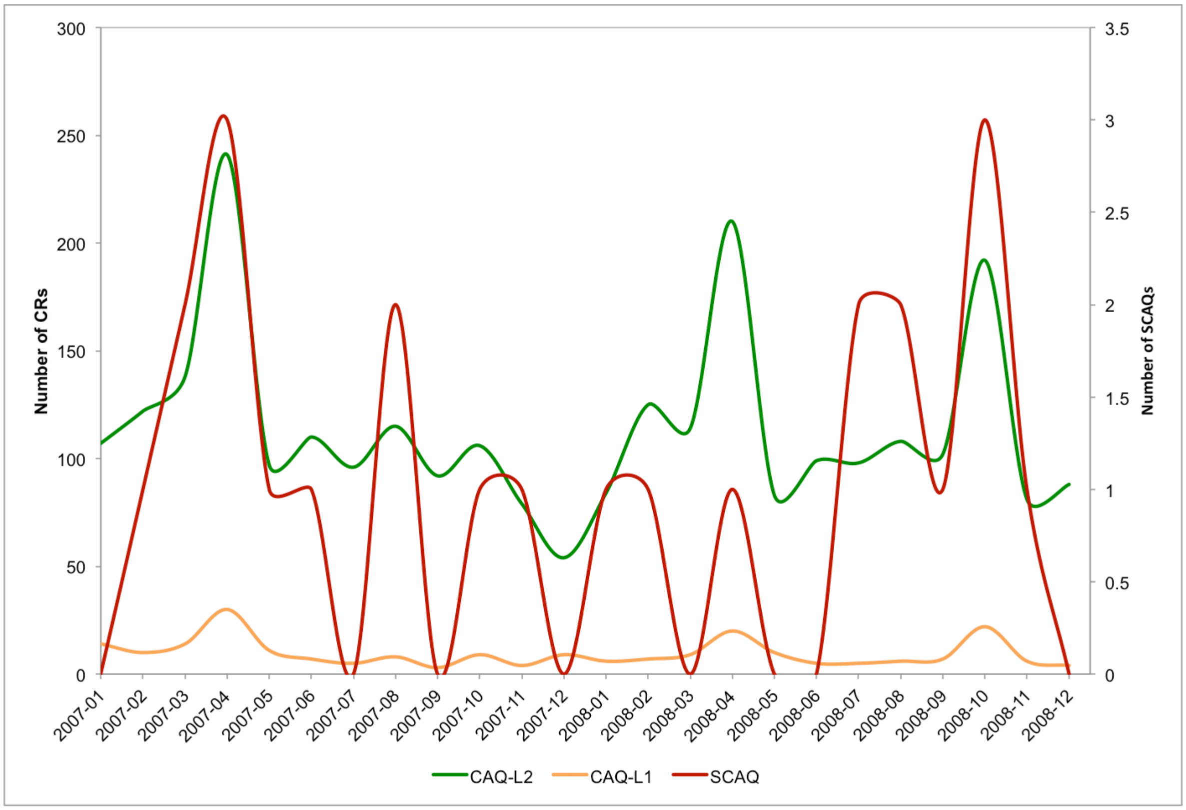

Figure 2-3 Events 2007-2008.

Figure 2-4, which shows events per week, includes the least severe events (Condition Not Adverse to Quality, CNAQ) in blue and locates the SCAQs by red dots. The higher red dots represent occasions when there were two SCAQs in one week. The importance of the CNAQs is their large number, and while they can be events that do not affect components, they sometimes generate as many as 2000 activities on top of the already large amount of work that each department must accomplish.

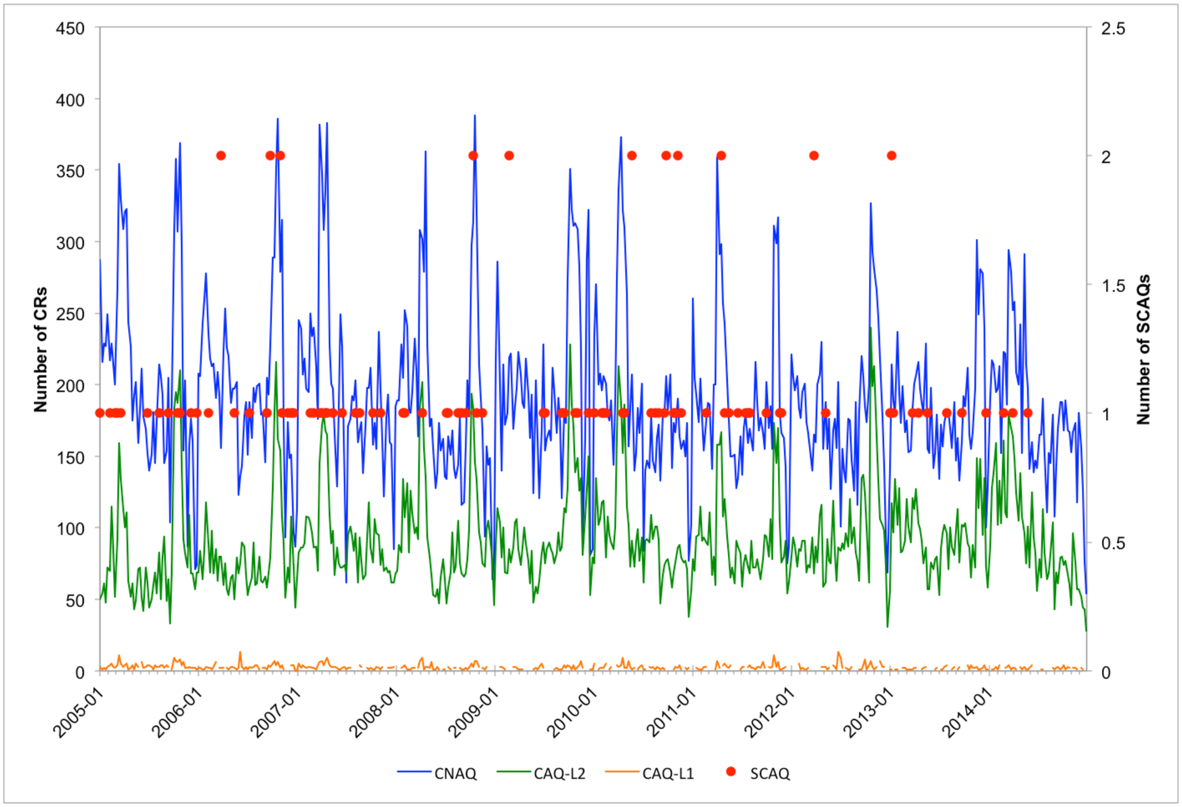

Figure 2-4 Events per week, 2005-2014.

2.1.2 Tools Developed from Time Series

From the CAP database, we can develop a simple planning tool, as presented in Figure 2-5. The cumulative frequency curve was developed for determining the probability of an SCAQ occurring given the number of CRs accumulated since the last SCAQ.

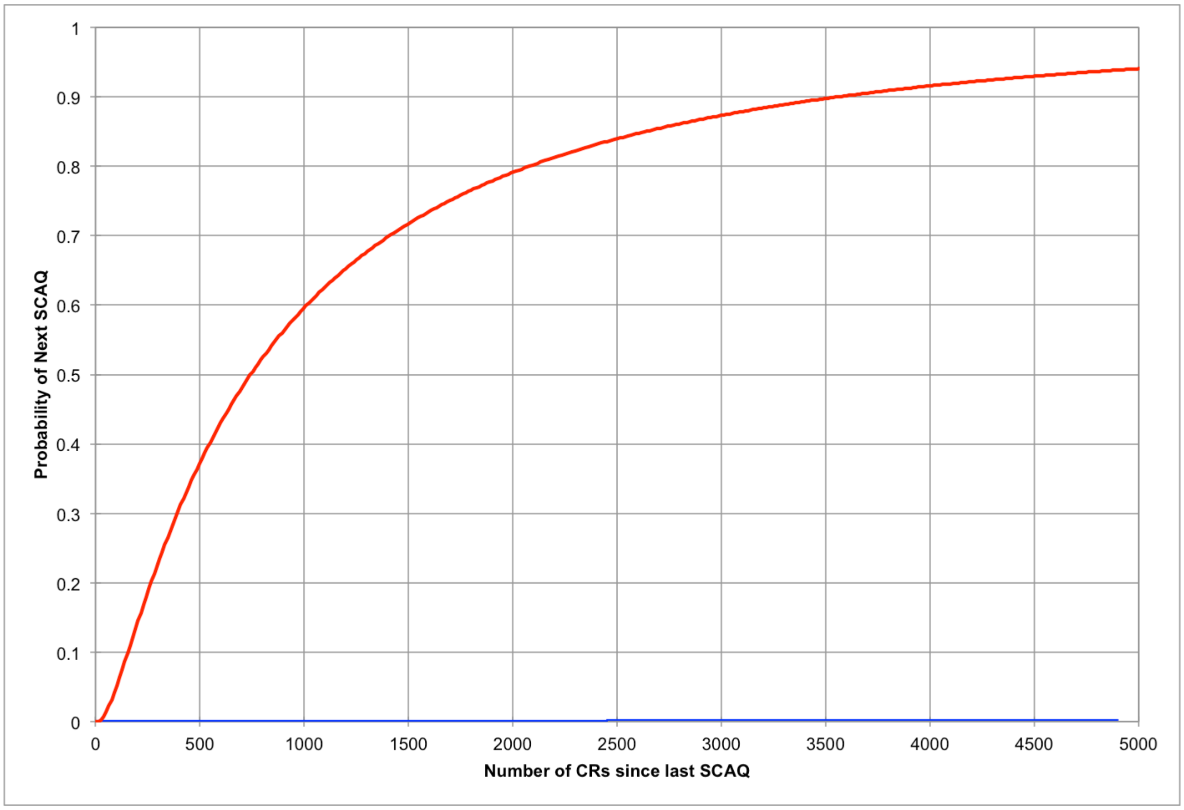

Figure 2-5 Probability of an SCAQ given number of events since last occurrence.

Although this curve is a simplified approach for developing an indicator (a more complete approach is presented later), this curve can be used to determine the position of the station relative to overall workload, which has been shown to be correlated with the likelihood of occurrence of an SCAQ. In fact, performance indicator thresholds could be established to indicate when a management barrier or other compensatory action may be implemented in order to reduce the likelihood of conditions meeting SCAQ criteria. In the case of this particular plant, for example, before there have been 5000 CRs since the last SCAQ, an organizational barrier or other actions (e.g., increased equipment performance reviews and monitoring) should be implemented in order to reduce the probability of the next SCAQ occurrence. While this can be helpful, the plant requires more insight into how the organizational factors influence the failures in human performance, in order to select the proper barrier to implement. An analysis of causal factors of the events and methods for choosing effective barriers is discussed in Nelson & Martín-del-Campo (2014). These failures in human performance are not only human errors, but also process and procedural complexities, as well as management decisions that impact plant performance. These organizational processes and decisions can have both direct and latent affects on plant equipment and can encompass all types of engineering, maintenance, and operations programs. For example, testing and maintenance frequency decisions should be based on historical data and the significance of the equipment to nuclear safety and reliable plant operations. So, a surveillance testing interval of every six months may be too infrequent to detect the onset of corrosion, and should be modified given the historical data. Also, in order to comprehend how plant processes and activities affect organizational factors and the resultant stress and strain they impose on station personnel, the inter- and intradepartmental factors are discussed in the following section.

2.1.3 Interdepartmental Factors

As part of the effort to better understand the organizational factors and human performance events that cause station level events, a detailed review of the CAP database was performed. It is the best source of empirical data for records of events at all levels and across all organizations, and through an analysis, enables one to understand which plant organization identified the problem and the organizations responsible for correcting problem. The number of CRs generated with an organization being identified as responsible, either as the identifier (i.e., generating a CR) or as the owner of an action within the CR, gives important insights into station procedural and process functions that result in specific plant organizations being more at risk for causing or responding to station events. Also, through the analysis of the actions that are generated after the occurrence of an event, we gain more insight into how the organizations communicate and work together or, at times, do not work together.

Figure 2-6 presents the distribution of Condition Reports among the station departments for all the severity levels. In ten years, more than 121,000 CRs were created by 169 organizational functions (it is recognized that some organizational functions may be shared among different station departments). In this data survey, the Procedures development function (labeled "Procedures" in the figures) is the leading generator of CRs. Procedures are recognized as being part of the cause, as well as the resolution. Since the procedure-writing function affects all activities at a station, it does not seem unreasonable that this function produces and receives the maximum number of actions (Figure 2-7). During this ten-year period there was a total of more than 400,000 actions generated, 106 significant conditions adverse to quality (SCAQs), and 7 plant trips.

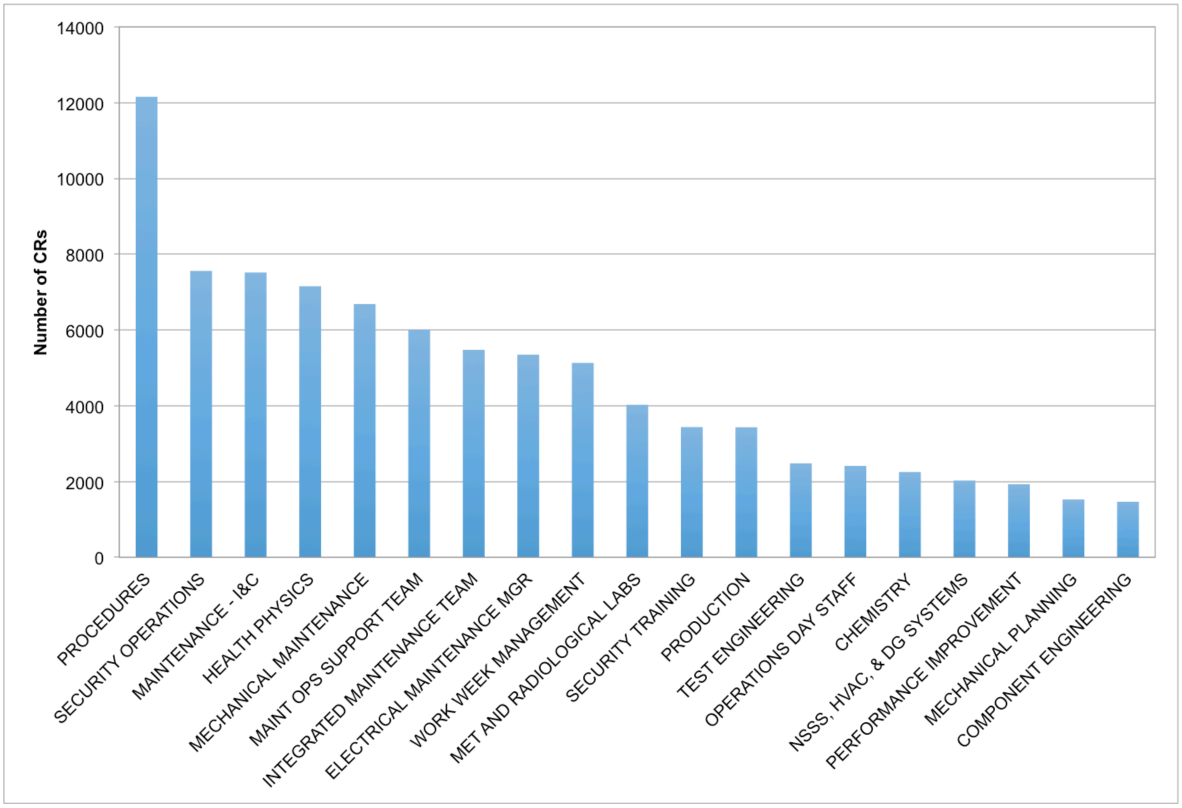

Figure 2-6 Departments creating CRs 2005-2014.

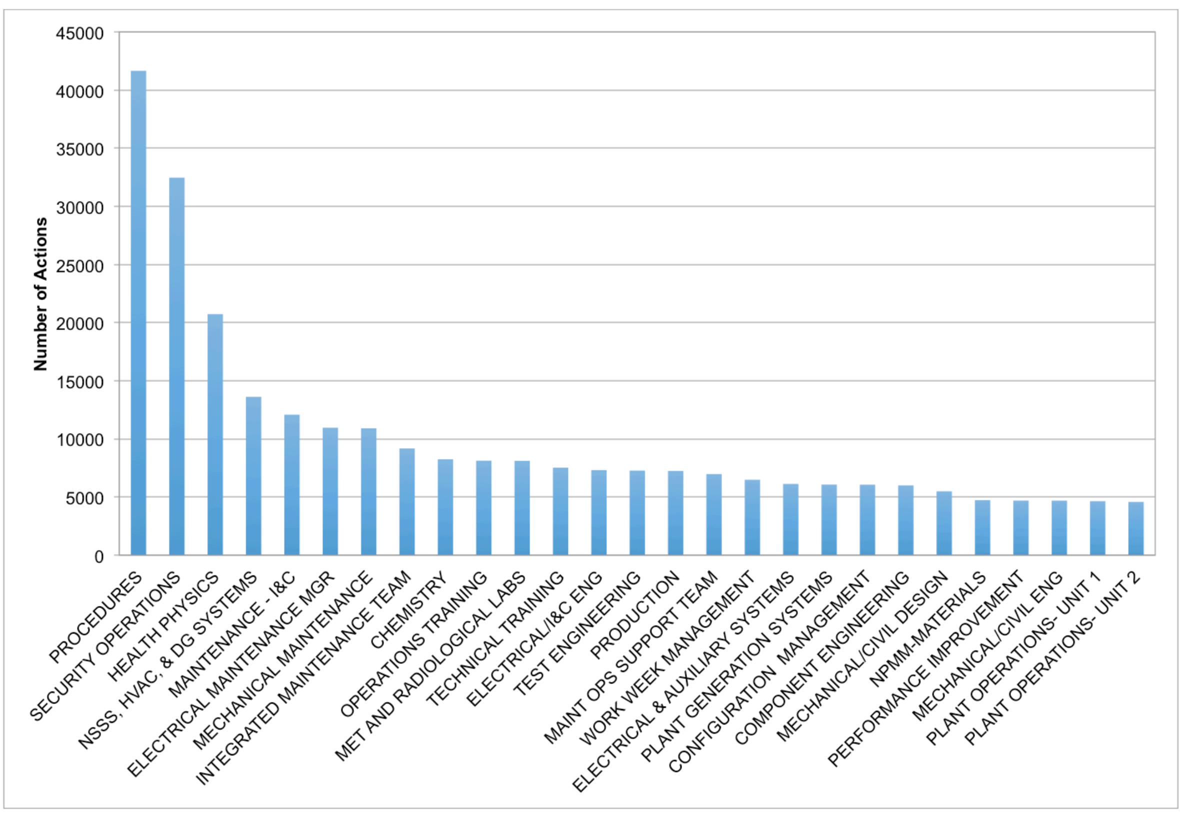

Figure 2-7 Departments receiving actions 2005-2014.

However, the Procedures function does not play a role in generating the most significant events in the ten-year period. As shown in Figure 2-8, the organizational functions that have caused two or more SCAQ events are organizational functions fall under the responsibility of the Engineering, Operations, and Maintenance departments. That is, procedures are responsible for the majority of the CRs, but not the SCAQs.

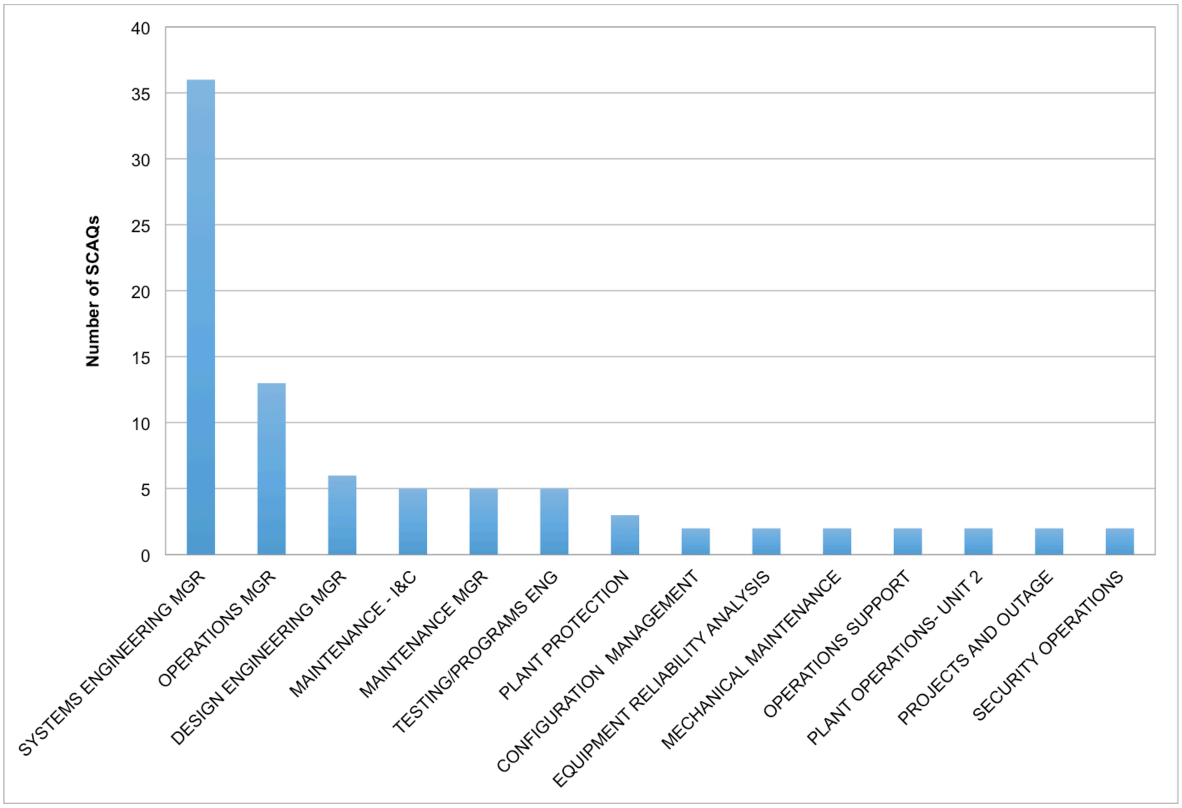

Figure 2-8 Number of SCAQs for departments responsible for more than one SCAQ.

The actions for other organizational functions received after an SCAQ was generated are shown in Figure 2-9 and the number of actions for the SCAQ owner in Figure 2-10. The observation is that the CR owners assigning actions add considerable strain on the individual departments, which in turn can increase workloads. In the next section this is shown to increase organizational stress.

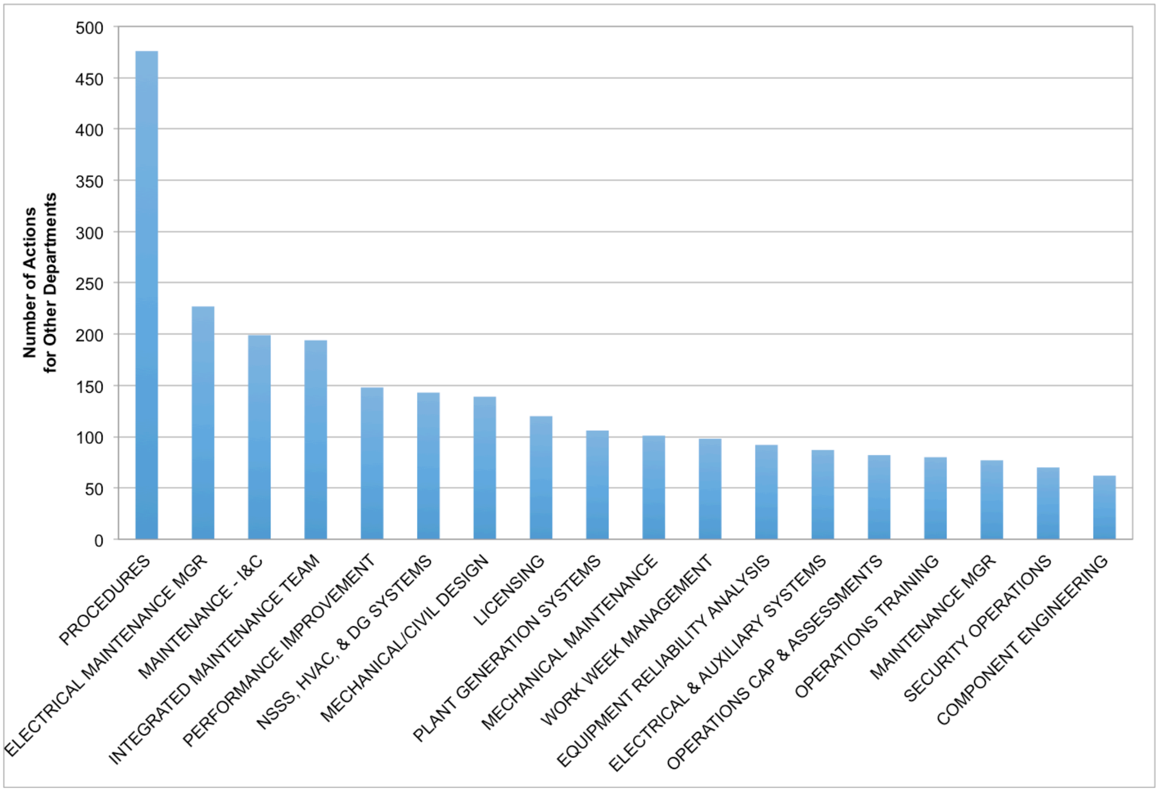

Figure 2-9 Actions for others from SCAQs.

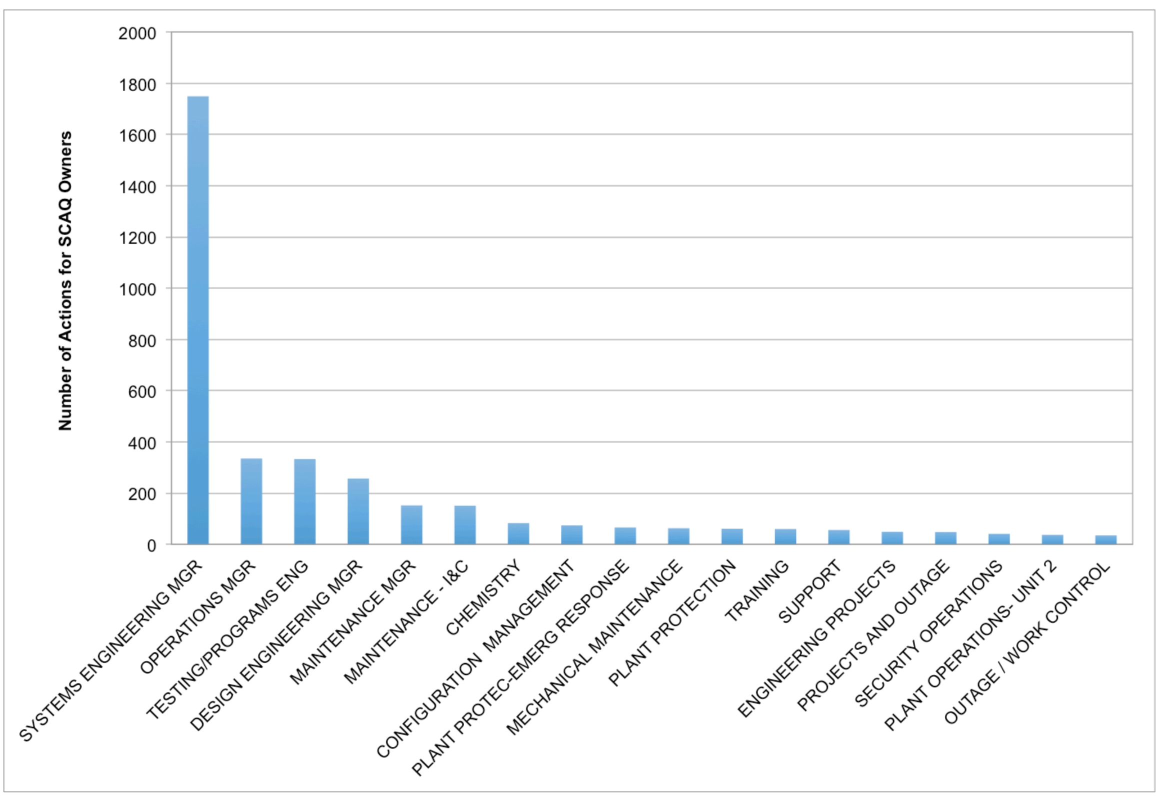

Figure 2-10 Actions generated for SCAQ owner.

It is difficult to describe organizational responsibilities and authority relations in simple statements. Plant organizations have specific functions and associated products (e.g., create procedures, perform maintenance), but they must also perform a variety of administrative activities. These activities include job-specific qualification and certification training, access authorization, emergency response organization participation, outage related assignments, etc. It is possible, through interviews and an extended set of observations over many different organizational activities, to begin to understand the amount and complexity of interdepartmental relationships, as done by Schulman (1993). We have found, as Schulman found in his qualitative study at Diablo Canyon, "Where error, oversight, or failure had foreseeable consequences that threatened individual or environmental safety, the administrative procedures were likely to be most elaborate and the interdepartmental interactions most intense". The process in this study is to determine the responsibilities, interactions, successes and failures through analysis of the reports included in the Corrective Action Program (CAP). The methodology developed to create the resilience model is discussed in the next section.

2.2 Methodology for Resilience Model

Due to a similarity between Cognitive Systems Engineering (CSE) and how organizations adapt and engineer resilience into their organizations, we propose a new method that provides organizational stress and resilience insights with respect to their relationship to plant performance. Using the ten years of CAP data, the correlation between increasing organizational demands and the likelihood of consequential events (i.e., plant trips, equipment clearance order error, component trips, inadvertent actuation of safety injection, etc.) is examined.

In this regard, it is anticipated that new and different insights into how organizational activities that support or facilitate work processes (i.e., soft processes) can and do result in both direct and indirect changes to equipment performance and reliability (i.e., hard impacts). A correlation was observed between the demand on an organization and the level of risk at the plant. This concept, which relates the resilience to the demands over time, is presented in Figure 2-11. In this figure, we can observe that the demand on the plant can be thought of as the stress placed on the organizational capacity, and this is related to the risk that exists at the plant due to all the ongoing activities. The resilience can be thought of as the organization's ability to cope with the risk and bounce back from increased risk (i.e., strength). However, if the stress reaches a resilience threshold, the plant will become brittle and not be able to adapt. In this case, the failure point is reached when an SCAQ occurs.

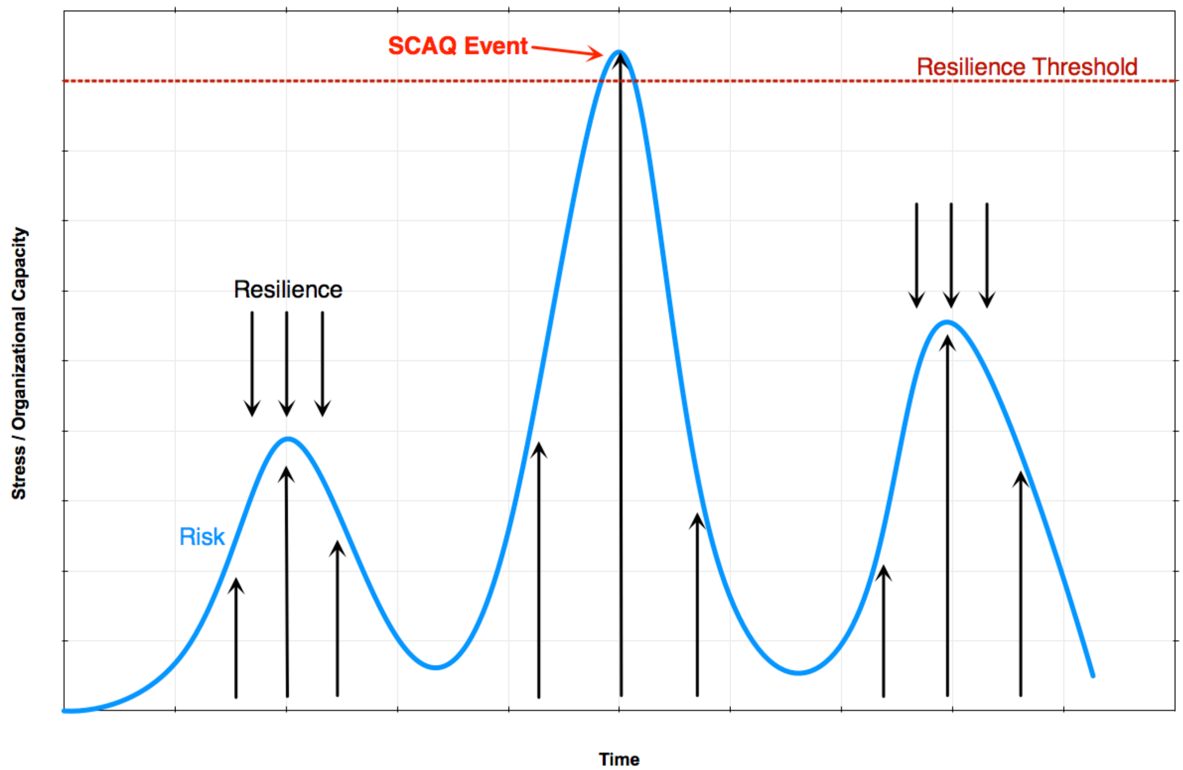

Figure 2-11 Organizational stress over time, adapted from (Pariès, Wreathall, Woods, & Hollnagel, 2012).

2.2.1 Organizational Stress and Strain Curve

One way to characterize and measure an organization's resilience can be based on an analogy from the field of materials engineering, the stress-strain curve (Figure 2-12). A stress-strain curve is created by stretching (straining) a material and measuring the resulting load (stress). The area under the linear (uniform) portion of the curve is called the resilience, the energy the material is able to absorb before deforming permanently. Materials that are brittle break along this linear region, without any yielding (permanent deformation). These terms and concepts correlate well with the basic finding in Cognitive Systems Engineering that demand factors are critical (Woods & Wreathall, 2008). Thus, the hypothesis is that to characterize a cognitive system of people and machines, one should examine how that joint system responds to different amounts of work activities. It is interesting that the two fields use similar language, resilience and brittleness, to characterize how an organization "stretches" as demands increase.

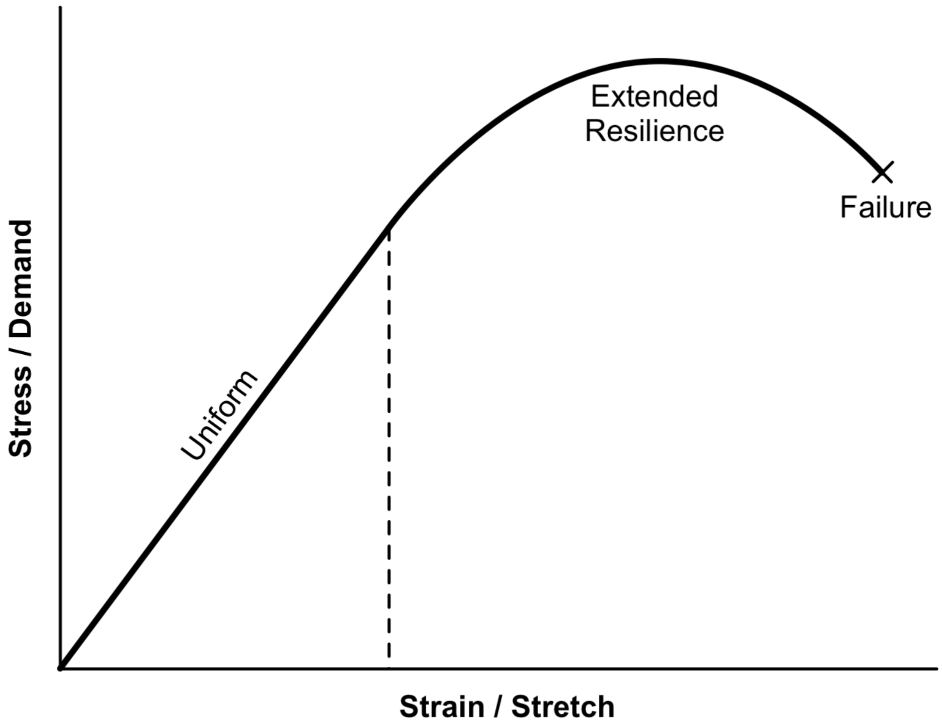

Figure 2-12 Basic demand-stretch or stress-strain curve.

2.2.2 Organizational Resilience Curve Methodology

The methodology is data-based and includes consideration of human error and organizational factors because of their large contribution to consequential events.

- Step 1. Gather CRs, and work activities (i.e., actions, preventive maintenance activities (PMs) and work orders (WOs)) per month from the CAP database, covering for a period of ten years. The outage history is needed for the same period of time. Within the category of severre events (SCAQs), the consequential events (main turbine trips and reactor trips) should be highlighted.
- Step 2. A scatter plot is developed with stress on the y-axis and strain on the x-axis, to develop the resilience curve. The stress is represented by the number of open CRs. The strain is the number of activities (i.e., WOs, PMs and open actions).

- Step 3. Develop the equation for the resilience curve, with a breakpoint defined as the resilience threshold. The Resilience Threshold is the point where main turbine and reactor trips begin to appear.

Finally, this equation can be used to calculate where the plant is on the resilience curve at any time as well as to predict where it will be in the next months, if no changes are made in the organization. When the stress factor (the number of CRs and the sum of the different work activities that are in a process) approaches the resilience threshold, a barrier should be installed, that is, some additional compensating actions should be implemented by the station organization to reduce the likelihood of failures in human performance and potentially avoid a consequential event. These failures in human performance are not only due to human errors, but also process and procedural complexities, as well as management decisions that impact plant performance. These organizational processes and decisions can have both direct and latent effects on plant equipment and can encompass all types of engineering, maintenance, and operations programs. For example, testing and maintenance frequency decisions should be based on historical data and the significance of the equipment to nuclear safety and reliable plant operations. So, a surveillance testing interval of every six months may be too infrequent to detect the onset of corrosion, and should be modified given the historical data.

2.3 Resilience Curve in an NPP

We can plot the strain as the number of preventive maintenance activities, CAP actions, and other work orders completed per month, which corresponds to an ever-present base level activity load for the plant organizations. The open actions are summed since these increase the organizational strain level of the station. The stress is related to the number of CRs opened or remaining open in the month. . Figure 2-13 presents the resulting organizational resilience curve for the plant used for this pilot study. The red squares represent plant trips, the point of exceeding the resilience threshold – the ability to absorb malfunctions in performance and maintain performance to some standard of performance (e.g., online power generation). The shaded area indicates the area where an increased likelihood of a plant trip is found and the base of this trapezoid is the perpendicular line that indicates where this increase in likelihood begins and is defined as the resilience threshold. At this point, it is assumed that the organizational elements and their interactions with plant equipment through planned and unplanned work result in more failures that cause consequential events (e.g., plant trips).

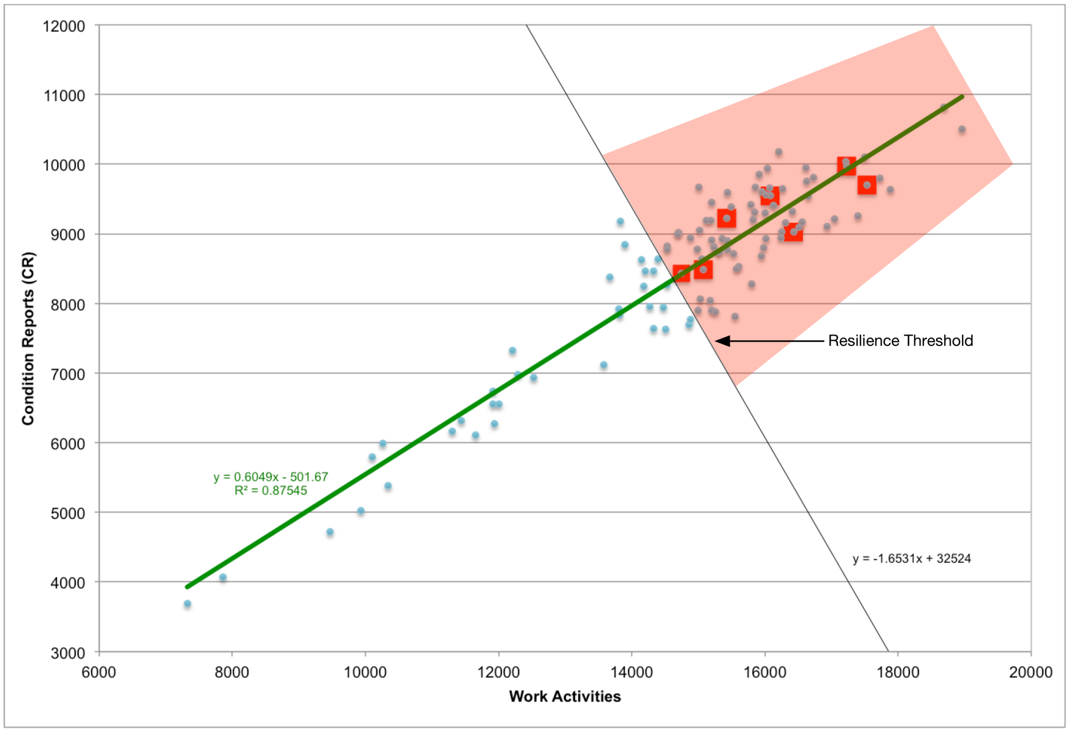

Figure 2-13 Organizational Resilience. The shaded area contains the plant trips and majority of consequential events.

2.4 Application

Based on this resilience curve, a method of anticipating consequential events was developed in the form of a leading performance indicator, using fuzzy logic. It provides the ability to monitor organizational demands against the increasing probability of a consequential event over time. Performance indicating alerts and thresholds are then proposed to provide awareness and recognition of "challenges" to organizational stress levels and resilience limits. This is shown as an increase in the probability of consequential events versus work activities, with thresholds associated with specific levels of risk (i.e., likelihood of plant trips). As noted earlier, the key premise is that increasing organizational demands, as recorded in the CAP database, reflect equipment or process problems that, in turn, increase the likelihood of a consequential event. As organizational demands increase, the organizational resilience limit is approached and the likelihood of the occurrence of a consequential event increases up to the point that a probabilistic prediction of the next consequential event can be made. This approach bases itself on plant-specific operating experience and history, specifically the number of consequential events and the demand on the organization. Thus, this indicator can predict the need to take action in order to avoid causing significant events; in this case, implement a barrier to protect the plant from such an event.

2.4.1 Performance Indicators

There are three types of performance indicators used in the nuclear industry: lagging, current, and leading. Lagging performance indicators provide information about a selected parameter (e.g., human performance) as reflected in events that have occurred in the past. For example, review of the NRC Human Factors Information System database (NRC, 2015b) for a randomly selected nuclear plant lists the Licensee Event Reports, Examinations Reports, and Inspection Reports associated with human factors that were reported during each year. Analysis of these events can determine human performance cause and contributing factor error categories. Counting the number of occurrences in each error category provides the basis for a lagging indicator of human performance. According to Reason's model (Reason, 1997) the lagging indicators are measures associated with the unwanted consequences of unsafe acts, such as those described in licensee event reports and significant event reports.

Current performance indicators provide information on selected parameters based on current conditions. For example, most nuclear plants have the voluntary Problem Identification and Resolution Program reporting system that is part of the CAP, as described earlier. Those items flagged as involving human performance can be placed in error categories and counted. The current performance indicator in this example is the number of items in each error category. According to Reason's model, current performance indicators are measures associated with the occurrence of unsafe acts, such as acts that are self-reported by workers, whether or not there was a significant consequential event

Leading performance indicators provide information about developing or changing conditions and factors that tend to influence future human performance. This same concept holds true for plant performance as well, since equipment or component events can provide information about developing or changing conditions that influence future plant performance. According to Reason's model, the leading indicators would be associated with the causes of unsafe actions, particularly the workplace and organizational factors. There have been efforts to develop leading performance indicators in the nuclear industry, such as EPRI's Human Performance Assistance Package (EPRI, 1999). The EPRI systems were piloted at three nuclear plants in the US and a concern that was presented in the final report of the pilot study (EPRI, 2003) was the inability to create a mapping from a leading indicator to an outcome, which is one of the intentions of the model in this chapter.

The development and use of leading performance indicators of human performance is a reasonable expectation given the volume of data being collected on a continuous basis in the nuclear industry. A structured approach to analyzing the data is presented here in order to establish a useful focus on available proactive, or leading, information and intelligence. Ready access to these ideas is fundamental for any organization in order to avoid consequential events. While the lagging and current performance indicators are fairly well understood and used, the leading performance indicators have been more challenging and, thus, have not yet been used to their fullest potential. The approach to developing leading indicators in this chapter is to establish a resilience threshold and monitor when the stress factor approaches this threshold, which will indicate when measures should be taken to reduce the likelihood of occurrence of a consequential event.

2.4.2 Approach for Developing a Leading Performance Indicator

While Groen & Mosleh (2005) stay with the conventional notions of logic, and assume the representations r in a space of representation R can be considered mutually exclusive, theories such as Fuzzy Logic (Zadeh, 1965) attempt to account for this source of uncertainty by introducing the concept of graded membership, as discussed in detail in for instance [25]. In order to develop a leading performance indicator from the resilience curve (Figure 2-13), a fuzzy logic approach was chosen, because the data support approximation rather than precision; however, a mechanism is needed to convert this rather imprecise data to a crisp performance indicator. Several studies have introduced the Fuzzy Set Theory (FST) approach for performance assessment of health, safety and environment in organizations (Cheng, 2010; Grecco, Vidal, Cosenza, dos Santos & Carvalho, 2013). These studies show important reasons to use FST: reduction of human error, creation of expert knowledge, and interpretation of large amount of vague or highly varied data.

The CAP databases used in US nuclear plants prove to be appropriate for the use of FST for similar reasons. They have a preponderance of human error related events (most minor but some significant and consequential). They identify implemented CAP corrective actions and lessons learned, which are the primary plant mechanisms for authorizing changes to virtually all station processes to improve performance. They also function as the primary repository or data warehouse for identifying, assigning, and scheduling work related activities for most all station activities, whether or not they are a baseline function or an added CAP function. In this regard, CAP programs represent an excellent barometer of the time dependent ‘pressure’ an organization is exposed to relative to activities defined in normal (routine) job functions and those that represent additional scopes of work with due dates resulting from problems or issues captured by the CAP process.

In our case, the fuzzy inference system uses the amount of work activities and CRs as input and the if/then rules are applied to calculate the consequences of exceeding the resilience threshold, that is, the increased likelihood of plant trips. While the focus is on plant trips as being the consequential event of measure, it is important to mention that a large percentage of the other consequential events occurred above the resilience threshold value as well. These other non-plant trip consequential events include: 85% of the SCAQs, 80% of the significant component trips, and 80% of the Equipment Clearance Order (ECOs) problems.

MATLAB’s Fuzzy Logic Toolbox (MathWorks, 2004) was used to create and edit a fuzzy logic system. The required parameters are encoded in fuzzy representations, and the interrelationships between them take the form of well defined “if/then” rules following the following steps:

1. Membership functions are built for the two inputs (Condition Reports, work activities) and also for a single output called Plant Trip. The linguistic labels “low”, “medium” and “high” were used to “fuzzify” the functions, based on normalized distribution of the values up to 50%, 50% to 75% and above 75% ; corresponding to up to 8480, from 8480 to 9400 and above 9400 CRs/month. Figure 2-14 shows the distribution of the CRs.

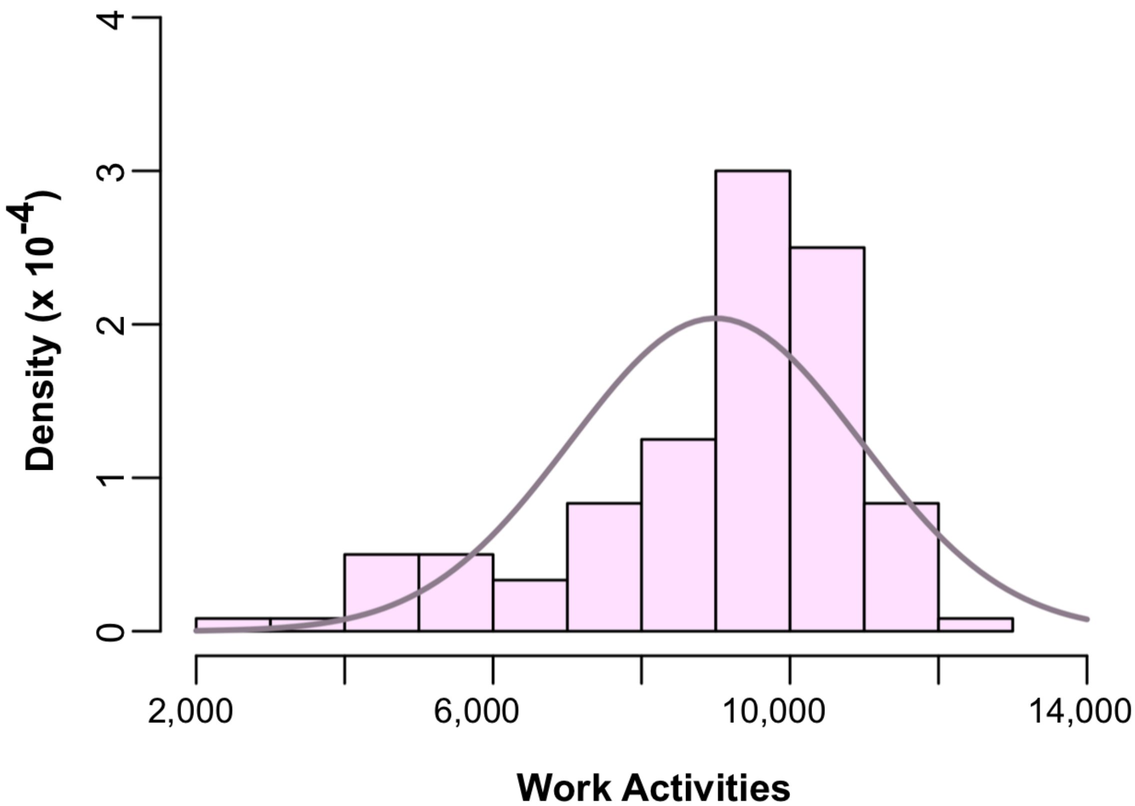

Figure 2-14 Normal Distribution of CRs.

2. Five fuzzy if/then rules are defined to determine the likelihood of a plant trip occurring in the short term, given the quantity of CRs and activities. These rules effectively define basically limit the area to the shaded area in Figure 2-11, although the last two include the area above and to the right of the point of the first plant trip, but with less weight since there is no evidence at this time.

   a. If "CRs" is <low> AND "Work Activities" is <low> Then "PlantTrip" is <low>.
   b. If "CRs" is <medium> AND "Activities" is <medium> Then "PlantTrip" is <medium>.
   c. If "CRs" is <high> AND "Activities" is <high> Then "PlantTrip" is <high>.
   d. If "CRs" is <high> Then "PlantTrip" is <high>, weight = 0.5.
   e. If "Activities" is <high> Then "PlantTrip" is <high>, weight = 0.5.

2. Apply implication method: to formulate the mapping from a given input to an output, the AND method with the *prod* (product) operator is utilized and the last two rules are assigned a weight of .5, due to less evidence obtained from the data.

3. The aggregation method *sum* is used to aggregate the output.

4. The output is "defuzzified" using the centroid calculation in order to obtain the likelihood of a plant trip given varying combinations of numbers of Condition Reports and work activities.

Figure 2-15 shows the surface graph of the likelihood of a Plant Trip occurring in the short term as a function of the number of Condition Reports and work activities obtained with this system. The general objective is to evaluate the conditions where the likelihood of a plant trip increases by varying the values for Condition Reports and work activities. The red squares again represent the plant trips.

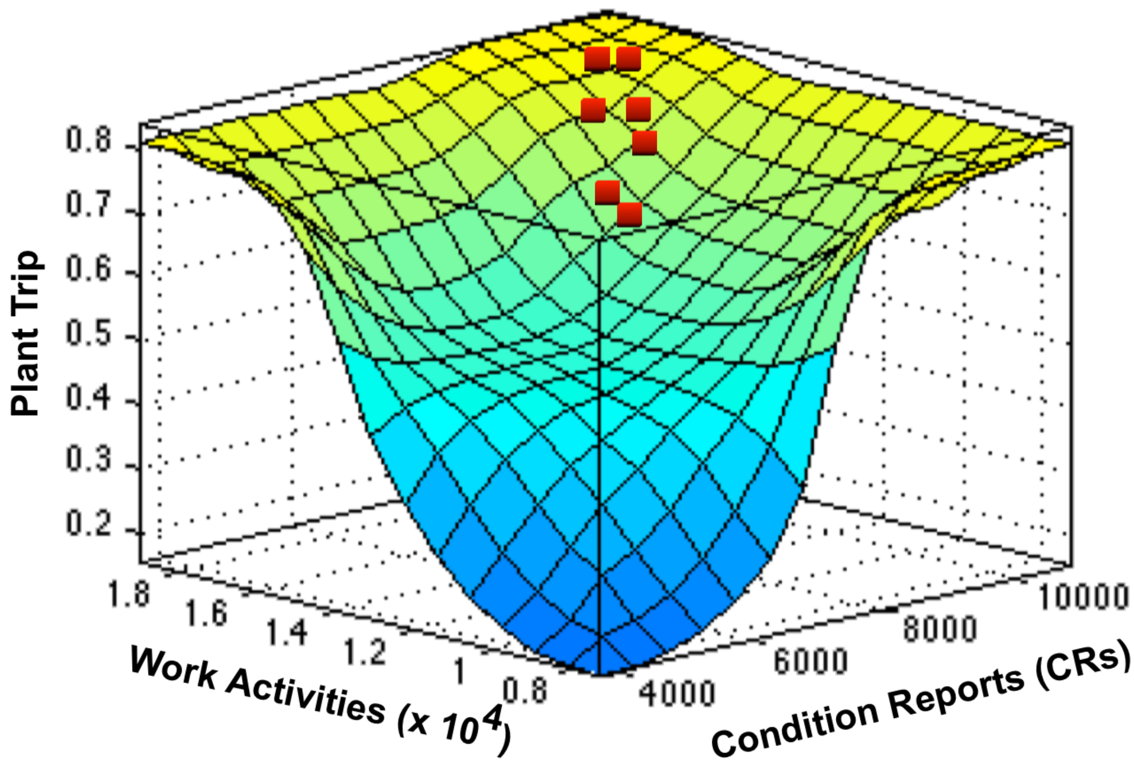

Figure 2-15 Likelihood of next plant trip as a function of Condition Reports and work activities in the surface viewer of the Fuzzy Logic Toolbox.

In order for a performance indicator to be useful, it should be uncomplicated (measurable) and straightforward. For this reason, the results acquired from the inference system are laid out in tabular form in Figure 2-16.

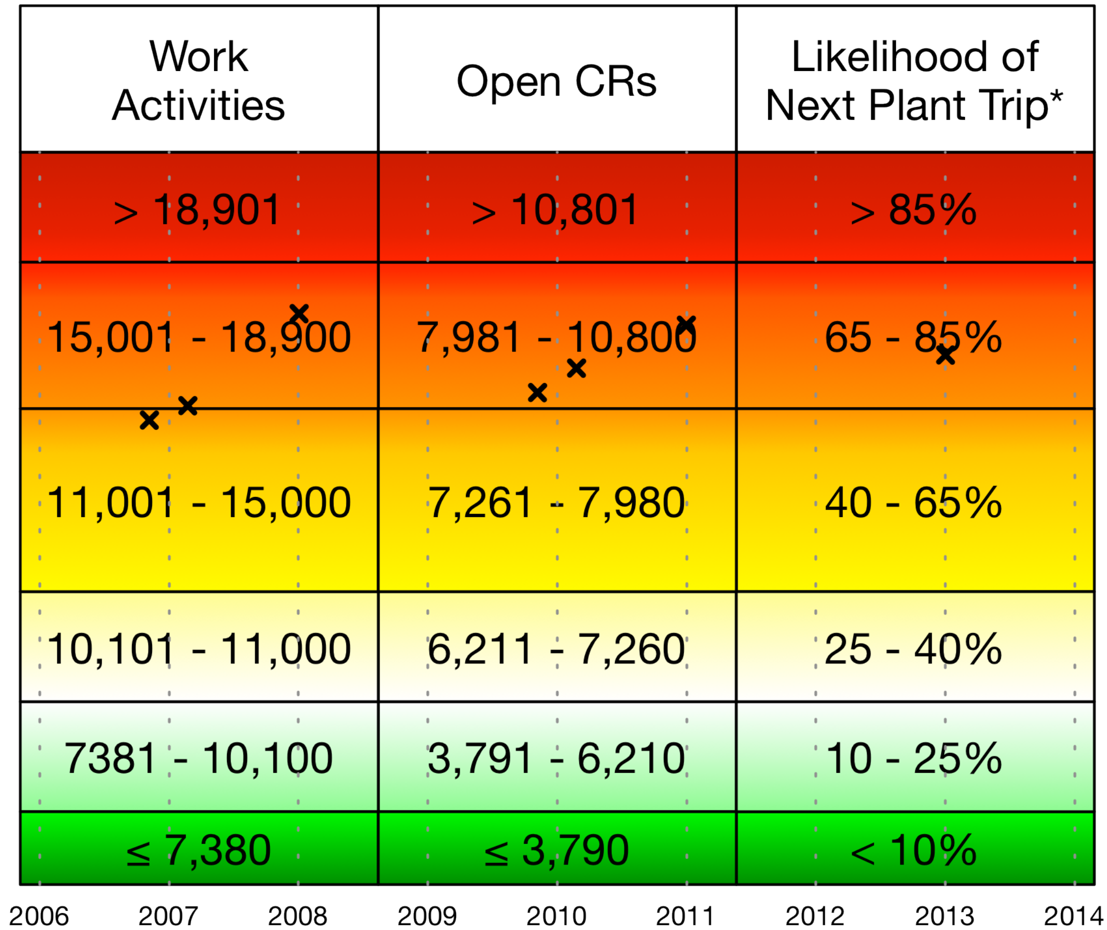

Figure 2-16 Conceptual performance indicator (*conditional likelihood in the short term).

If station personnel were to track their location on Figure 2-16 on a monthly basis, the indicator would notify station leaders when plant and organizational stresses are increasing beyond an "alert" level, developed through the inference process, described above. For stress levels in the white band, this would mean that at that level of stress, there is a 25 to 40% conditional likelihood that a plant trip would occur in the short term (within the next month). The X´s indicate the color band for each of the plant trips over nine years (2006-2014); the first plant trip occurred in the yellow band; however, the remaining six all occurred in the orange band (two in 2007, one in 2008, two in 2010, one in 2011 and 2013). There were no plant trips during 2009, 2012, and 2014. Those time periods are associated with the white and light green color bands, which further adds verification of the indicator's validity

This performance indicator could be further enhanced through identifying a "required action" level where management-directed compensatory measures could be taken based on examination of the current plant and organizational status or performance relative to significant plant and organizational functions. The leading indicator presented here is analogous to a thermometer-type or heat-index performance indicator. It should create awareness in management and station personnel leading to further internal examinations when stress levels are exceeding predefined limits. It should also lead the plant management and personnel to further examine current plant conditions for vulnerabilities of a plant trip. The use of the white band region could be assigned as the "alert" band and the yellow region could be designated as the "required action" band. The alert band indicates the appropriate time to begin to reduce strain on the plant (reduce work activities, such as PMs, WOs and/or CAP actions) or implementing measures or barriers to effectively address current station vulnerabilities and increase the resilience threshold. The required action band indicates the region where immediate development and implementation of identified stress reducing actions and/or actions to reduce current plant vulnerabilities are required to be implemented and monitored.

It is important to note that with this leading performance indicator that actions taken to return to lower regions may not be conducive to the intent of corrective action programs (i.e., to identify and correct problems). Thus, in that regard, this indicator's value is in its ability to predict when increased likelihoods of plant trips could occur, which in and of itself, represents the involuntary reduction in resilience since the plant trip changes both plant and organization to an outage frame of mind (i.e., other work stops and focus is solely on returning the plant back to at-power conditions). However, by predicting conditions whereby increased plant trips are more likely to occur based on plant specific operating history there is an opportunity for organizations to pause and make an assessment of current conditions and, in so doing, re-scope and re-prioritize activities to increase resilience (e.g., free up critical resources that may be currently committed to less significant activities or reschedule work activities to more appropriate time frames). This leading indicator is intended for this purpose and if implemented by nuclear plant organizations can provide an important cue to perform a "resilience examination".

2.5 Discussion

The resulting performance indicator implies that it is possible to monitor organizational stress levels and implement compensating actions before the plant organization and equipment reach the point where undesirable events (e.g., plant trips) occur. Organizational performance improvement is a generic concern at commercial nuclear power plants, and the approach described in this chapter provides a method to improve organizational performance, beyond that currently achievable with event reporting and Corrective Action Program monitoring, through the evaluation of organizational stress levels and associated resilience levels, leading to the development of a proposed leading performance indicator.

It has been shown here that the CAP database can be used for many purposes; the following four were covered briefly in this chapter:

1. Describe organizational factors between and within departments.
2. Calculate the probability of an SCAQ given the number of CRs reported since last occurrence.
3. Detect when the station is at risk of exceeding its resilience.
4. Develop an organizational performance indicator.

We have also shown that the CAP databases are proper candidates for the use of Fuzzy Set Theory, due to the scope and high variability (i.e., uncertainty) of items captured in CAP processes that, at some level, are all contributors to overall organizational and plant performance levels.

We have found, as did Hollnagel (2013), that resilience engineering provides a way to identify the capabilities that a complex socio-technical system must have to perform acceptably in everyday situations, as well as during accidents. Applying the cognitive system engineering analogy to organizational resilience, we were able to build a stress-strain curve to relate the station's stress (i.e., CRs) to the strain (i.e., work activities) that allows the station to continue to operate successfully. The station's CRs that are accounted for in this report are both "soft" CRs in terms of the process activities required to operate and maintain a nuclear power plant and the "hard" CRs in terms of the equipment and component issues that place further demand on organizational performance and that can also generate consequential plant events.

Thus, the organizational performance can be characterized by a strain and an associated stress, which indicate levels of organizational resilience. The strain is defined in this study as the sum of the preventive maintenance, work orders, and open CAP actions. Organizational strain is seen to increase before, during, and just after an outage but can also have peaks during at-power times. The stress is measured by the number of CRs, which is the plant's mechanism for identifying events, errors, and other failures across most all plant processes. An organization's resilience is its ability to withstand these stresses and strains and still satisfactorily perform activities. The point where the stress and strain result in consequential events, such as plant trip (i.e., the "breaking point"), is the resilience threshold.

This chapter provides a method for measuring and analyzing stresses in term of the likelihood of consequential events based on plant specific operating experience. These parameters form the technical basis for developing a leading organizational resilience performance indicator. Since the SCAQs represent times when demand on the organization (i.e., stress) exceeds its resilience limits, we use the occurrence of a plant trip as the consequential event of concern. Thus, when the stress factor exceeds the resilience threshold, it is more likely that a plant trip will occur. The performance indicator presents a conceptual color band arrangement representing the increased likelihood of a plant trip based on the stress factor. When the stress factor approaches the resilience threshold, additional barriers and other provisions should be considered for implementation.

When a particular problem is identified and resolved, the solution, represented by a corrective action or set of actions, does not always remain effective over long periods of time (i.e., years). The continual monitoring, application, and communication of the described process is necessary to assure that the resilience performance indicator continues to provide useful and timely information. Because change and adaptability increase resilience, the process will be improved by continual or periodic updating. Reductions in consequential events at the plant level over a period of time will be the key indication that either organizational stress has been reduced to more acceptable levels or that organizational resilience has been increased due to increased organizational capacities and capabilities.

Processing the operational data daily or at least weekly will provide a regular update of the stress factor and the resilience threshold and produce a more accurate leading performance indicator for preventing consequential events.

Chapter 3 Causal Analysis

Leading indicators can be used to identify the need for installing a barrier or defense to reduce human errors in a nuclear power plant. These indicators are developed from the Corrective Action Program data by detecting increases in events, as described in the previous chapter. Organizational barriers can then be identified to improve performance. The resulting identified barriers are evaluated to rank the value of each possible barrier, and determine the best barrier(s) to implement. The tool described in this chapter is designed to provide a systematic approach to identify areas where improvements in organizational effectiveness best reduce the likelihood of consequential events. Due to the uncertainty of many of the factors that influence the performance of humans in nuclear power plant activities, we propose using Bayesian networks to identify sources of organizational errors leading to consequential events. This researchy, using actual nuclear power plant data, includes a method for data processing and highlights some potential uses of Bayesian networks for improving organizational effectiveness in the nuclear power industry.

The process for determining compensatory measures is explained in this chapter. These measures take the form of barriers, defenses and safeguards, which are employed to protect High Reliability Organizations from accidents. High technology systems have many defensive layers: some are engineered (alarms, physical barriers, automatic shutdowns, etc), others rely on people (surgeons, anaesthetists, pilots, control room operators, etc), and yet others depend on procedures and administrative controls (Reason, 2000). Their function is to protect potential victims and assets from local hazards.

While there are already layers of barriers and defenses included in the design and operation of a NPP, there are times when an extra barrier is required, as shown in the previous chapter; the purpose of this chapter is to describe the process for determining which barrier to implement. This depends on the organizational resilience, the causes of the errors occurring at the times of lowered resilience and the near term program of activities, as will be shown in this chapter.

3.1 Corrective Actions and Barriers

Each CAQ requires action to correct the condition. Additionally, for SCAQ, corrective actions to preclude repetition are applied depending on the significance of the condition. Corrective actions should be completed in a timely manner commensurate with the condition's safety significance and complexity. In determining the actions to take, the following should be considered: (1) the consequence of malfunction or failure of the equipment; (2) the design and fabrication complexity or uniqueness of the equipment; (3) the need to apply special controls and/or surveillance over the processes and equipment; (4) the degree to which functional performance can be demonstrated by inspection or test of the equipment; (5) the quality history and degree of standardization of the equipment; (6) the difficulty of repair or replacement, especially after installation; and (7) the effect on ITAAC conclusions (refer to NEI 08-01). The actions taken to correct a condition should be documented to allow further review and evaluation.

Corrective actions implemented for SCAQ are to be promptly reported to appropriate levels of management. The appropriate management to be notified should be established within the implementing procedures. If CCAP are delegated to a supplier, the interface and requirements for reporting should be clearly documented.

Periodically, CAQ should be analyzed for adverse trends within and across the various work processes and the CAP. A trending process should be implemented that can identify adverse trends that are QAP deficiencies or significant to safety (such as repetitive failures or process weaknesses). This review is conducted to identify generic issues and vulnerabilities early in the work process before significant problems result. Management personnel responsible for the work activities are responsible for identification of thresholds for trending to determine the presence of adverse trends, repetitive failures, process weaknesses, or other indicators of extent of cause or condition beyond the immediate problem identified.

Construction or operating experience and NRC generic communications should be reviewed for applicability to conditions that exist at the facility and to assist in the identification of adverse trends.

To identify patterns that warrant broad corrective actions, trending can also be accomplished using detailed codes and data analysis techniques for certain work processes. One type of trending level or technique is not practical for all conditions; therefore, a structured approach to trending should be implemented by licensees and suppliers during nuclear construction.

Adverse trends should be reported to management responsible for the work process. Management should provide oversight of the trending process to assure the process is properly implemented.

An adverse trend may exist if:

- Deficiencies identified are of a repetitive nature and the number appears excessive or exceeds an established criteria or threshold, taking into consideration time frames and levels of associated line organization and QA/QC activities.

- Recurring deficiencies are of a significant or severe nature.

- Increases in the number of deficiencies that cannot be easily attributed to new or special work programs, or increased quality verification activities.

- Deficiencies are of a programmatic nature, apparently not limited to a specific organization.

- Previously identified corrective actions are apparently ineffective in reducing the number or severity of deficiencies.

- Recurring deficiencies appear to be related to a possible single root cause.

- Deficiencies of a like nature are being identified in multiple work activities.

The goal of the trending program is early recognition of trends so underlying causes can be investigated and actions taken before major issues or conditions occur, thus allowing for continual improvement.

3.2 Causal Data Mining

A causal model can be used as an estimating approach based on the assumption that the future value of a variable is a mathematical function of the values of other variable(s). It is used where sufficient historical data is available, and the relationship (correlation) between the dependent variable to be forecasted and associated independent variable(s) is well known (Woods & Wreathall, 2008).

Among the different data mining algorithms, probabilistic graphical models (in particular Bayesian networks) constitute a sound and powerful methodology grounded on probability and statistics, which allows building tractable joint probabilistic models that represent the relevant dependencies among a set of variables (hundreds of variables in real-life applications). The resulting models allow for efficient probabilistic inference. For example, a Bayesian network could represent the probabilistic relationships between large-scale synoptic fields and local observation records, providing a new methodology for probabilistic downscaling: i.e., allowing to compute the probability of a certain observation given la large-scale prediction, i.e., P(observation|large-scale prediction). For instance, the red dots in Figure 3-1 correspond to the grid nodes of a Global Climate Model (GCM), whereas the blue dots correspond to a network of stations with historical records (the links show the relevant dependencies, automatically discovered from data) (Santander Meteorology Group, 2013).

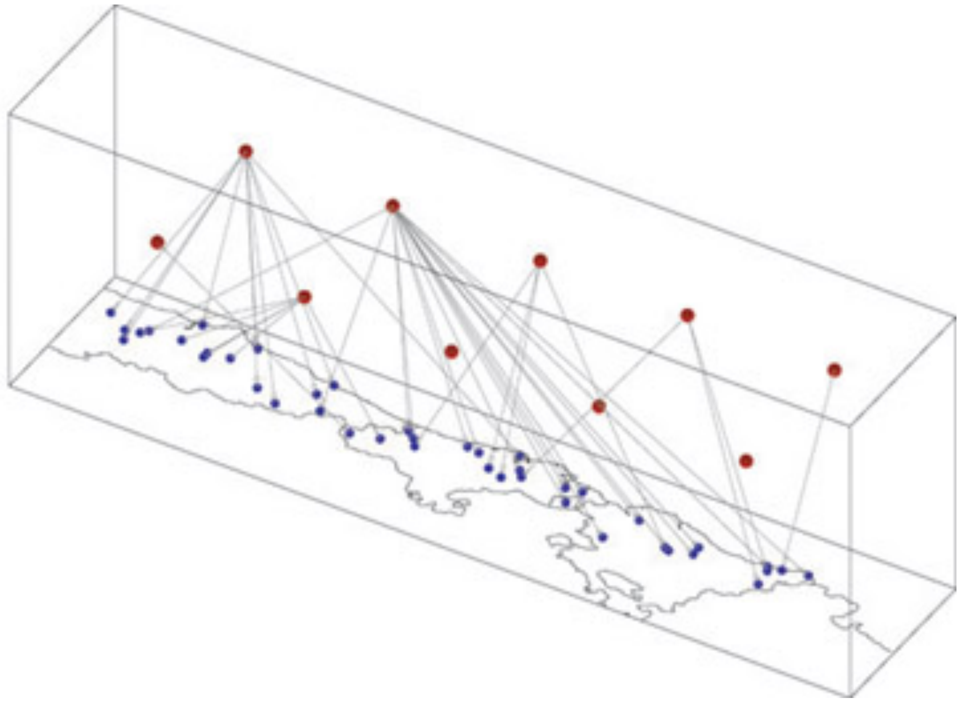

Figure 3-1 GCM modeling (from Santander Meteorology Group, 2013).

Among the different data mining algorithms, the use of probabilistic graphical models, in particular Bayesian Networks (BNs), is a sound and powerful methodology grounded on probability and statistics, which allows building tractable joint probabilistic models that represent the relevant dependencies among a set of variables. The resulting models allow for efficient probabilistic inference. More on the theory and use of BNs can be found in numerous books, such as those by Pearl, J. (1988), Jensen, F.V. (2001), Koller, D. & Friedman, N. (2009), Pourret, O., Naïm, P. & Marcot, B. (2008), and Darwiche, A. (2009).

In this work, a BN represents the probabilistic relationships between the causes and the incidents reported at a nuclear power plant during routine maintenance and surveillance activities. The BN provides a new methodology for probabilistic downscaling, i.e., allowing computation of the probability of a certain type of event, e.g., P (reactor trip|a certain cause combination), as well as the decreased probability given a decrease in the probability of a given cause. For instance, an improvement in procedures (in this case, we assume perfect procedures) can decrease the probability of reactor trip.

Formally, BNs are directed acyclic graphs (DAGs) whose nodes represent variable and whose arcs encode conditional independencies between the variables. The graph provides an intuitive description of the dependency model and defines a simple factorization of the joint probability distribution leading to a tractable model that is compatible with the encoded dependencies. Here we present a model derived from the set of data to predict the probability of undesired consequences in routine operation at a nuclear power plant. Efficient algorithms exist to learn both the graphical and the probabilistic models from data, thus allowing for the automated application of this methodology in complex problems.

Bayesian networks that model sequences of variables (such as, for example, time series of historical records) are called dynamic Bayesian networks. Generalizations of Bayesian networks that can represent and solve decision problems under uncertainty are called influence diagrams, an example of which is presented in Section 3.12.

The way a Bayesian network is used for quantification is presented through the simple example shown in Figure 3-2. If there are four variables that are used to describe cases, we can derive the graphical model from the data and quantify the probability, Pr, of the network using equation 3.1.

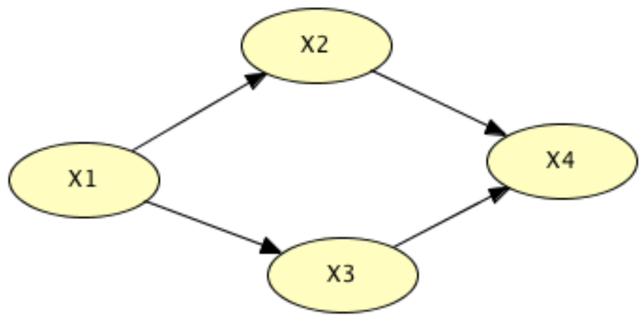

Figure 3-2 Example Bayesian Network.

$$\Pr(X1, X2, X3, X4) = \Pr(X1) * \Pr(X2|X1) * \Pr(X3|X1) * \Pr(X4|X2, X3) \qquad (3\text{-}1)$$

Although there is abundant information, these sources seldom provide all the numbers required for a real application; however, we propose that the use of the data available in the Corrective Action Programs, once put into ontological terms, could provide the statistics necessary for producing extremely accurate probabilistic models for human error, including organizational influence.

3.3 Probabilistic Networks

Building a probabilistic network for a domain of application involves three tasks:

- Identify variables that are of importance and their possible values.
- Identify the relationships between variables and express in a graphical structure.
- Obtain the probabilities that are required for the quantification.

With respect to the last task, the most common sources are statistical data, literature, and human experts. Although there is abundant information, these sources seldom provide all the numbers required for a real application.

For this reason, a BN can be used to generate the probabilities required. First the BN must be built. There are two approaches to learning (the computer programmer lexicon for building the structure) the graphical structure from data. The first is based on constraint-based search and the second on Bayesian search for graphs with highest posterior probability given the data. Once the graphical structure has been established, assessing required probabilities is straightforward and involves studying subsets of data that satisfy various conditions (Murphy, 1998).

Since the model is data-based, the next section describes the data and how it is used.

3.4 CAP Data

The variables considered in this study were taken from the causes indicated for each event reported in the Condition Reports and are listed in Table 3-1, these are the major categories from the GFTs in Table 1-3. We can observe the name and the description of the cause code variables, with some examples to demonstrate the types of causes considered in this research

Table 3-1 Variables Considered.

Variable	Description	Examples
DE	Equipment Design/Manufacture/Performance Monitoring	Predictive Maintenance Program Inadequacy, Preventive Maintenance Program Inadequacy
HF	Human Factors/Work Environment	Human Factors Not Properly Addressed in Work Area/Equipment
LS	Job Leader/Supervisory Methods	Pre-job Preparation or Briefing Inadequate, Prioritization of Work Activities Inadequate
MA	Management Assessment/Corrective Action	Organization Not Sufficiently Self-Critical, Cause Analysis for Known Problem Inadequate
MC	Change Management	Need for Change Not Recognized, Change Not Implemented in a Timely Manner
MP	Management Practices	Communication Within an Organization Inadequate/Untimely, Communication Between Organizations Inadequate/Untimely, Management Practices Promote/Allow Unacceptable Behaviors
MR	Management Resources	Prioritization/Scheduling of Activities Inadequate (Management level)
PA	Procedure Adherence	Procedure/Instruction/Step Implemented Incorrectly (Intent Not Met)
TR	Training	Necessary Initial/Refresher Training Not Provided
WI	Work Instructions	Document Contents Incorrect or Missing
WP	Work Practices	Slip or Lapse

3.5 Factor Analysis

The model development methodology uses factor analysis to collapse the causes into categories to structure the model. Factor analysis is a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors. This has been found useful for structuring a causal model and is employed in this chapter to build the Bayesian Network that is used for the barrier analysis, objective of the chapter.

Table 3-2 presents a sample of the first 10 lines of the data used for the factor analysis; there are 13900 entries in this example complete matrix; there are 121,000 entries in the complete matrix for the 10 year period. Sufficient data can be considered as having ten times the data samples as the number of variables that describe the samples. For example, in the case of the CAP database for the pilot plant, there are 11 variables as shown in Table 3-1, or if the subcategories are used, 62 variables, as shown in Table 1-3, and the number of data samples are the number of CRs, which are more than 120,000, when ten years of data are considered.

In order to conduct the data analysis and build the Bayesian networks, the data was converted to binary, which means that 0 represents "Adequate" for that variable, and 1 means "Less than Adequate" for that variable. In this way a binary matrix is formed for the data for the significant events reported in the Condition Reports.

Table 3-2 Binary Matrix for 11 Variables.

Condition Report	DE	HF	LS	MA	MC	MP	MR	PA	WI	WP	TR
1	0	0	1	0	0	1	0	0	0	0	0
2	1	0	0	0	1	0	0	0	0	0	0
3	1	0	0	1	0	0	0	0	1	1	0
4	0	0	0	0	0	0	0	0	1	0	0
5	0	0	0	0	1	0	0	0	1	0	0
6	1	0	0	0	0	0	0	0	0	0	0
7	0	0	1	1	0	1	0	0	0	1	0
8	1	0	0	0	0	0	0	0	0	0	0
9	1	0	1	1	0	1	0	0	1	0	1
10	0	0	1	1	0	1	1	0	0	1	0

The default in most statistical software packages is to retain all factors with eigenvalues greater than 1.0, corresponding to the first five factors in our analysis; however, there is broad consensus in the literature that this is among the least accurate methods for selecting the number of factors to retain (Murphy, 2008). Many experts describe that the best choice for researchers is the scree test. This method is described in every textbook discussion of factor analysis. The scree test involves examining the graph of the eigenvalues and looking for the natural bend or break point in the data where the curve flattens out. The number of data points above the "break" (i.e., not including the point at which the break occurs) is usually the number of factors to retain. Figure 3-3 presents the Scree plot where we can see that the first four factors explain almost 60% of the variance in the data.

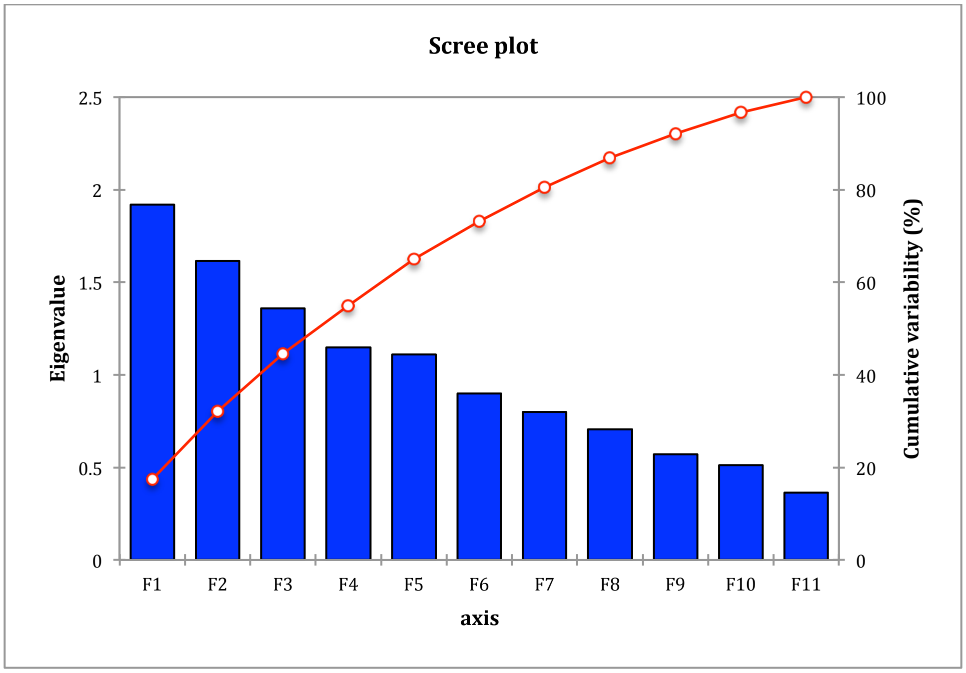

Figure 3-3 Scree Plot.

Table 3-3 illustrates the loading of the variables on the four extracted factors. As we can see, DE loads on F1, while the management variables load on F2 (MA, MC, MP, MR), etc. We can interpret this by observing that the factors divide the human performance difficulties into four categories: (1) maintenance programs, (2) management issues, (3) work practices and supervision, and (4) training, procedures and instructions.

Table 3-3 Factor Loadings.

Variables	F1	F2	F3	F4
DE	0.871			
HF		0.436		
LS			0.366	
MA		0.436		
MC		0.456		
MP		0.504		
MR		0.276		
PA				0.583
WI				0.415
WP			0.538	
TR				0.736

Table 3-4 shows the correlations among the variables, which in turn are used to define the links or arcs between the variables in the causal model that was developed and is presented in Section 3.3.

Table 3-4 Correlations among Variables.

Variables	DE	HF	LS	MA	MC	MP	MR	PA	WI	WP	TR
DE	**1**	0.216	-0.147	-0.001	-0.184	-0.248	-0.197	-0.137	-0.155	**-0.297**	0.008
HF	0.216	**1**	0.156	0.106	0.164	0.201	-0.062	-0.043	-0.131	-0.127	-0.072
LS	-0.147	0.156	**1**	0.075	0.044	**0.369**	0.243	0.086	0.034	0.254	-0.029
MA	-0.001	0.106	0.075	**1**	0.129	0.199	0.152	-0.143	0.030	-0.107	0.152
MC	-0.184	0.164	0.044	0.129	**1**	0.091	0.123	-0.115	-0.073	-0.228	-0.025
MP	-0.248	0.201	**0.369**	0.199	0.091	**1**	0.152	0.106	0.040	0.060	0.153
MR	-0.197	-0.062	**0.243**	0.152	0.123	0.152	**1**	-0.047	0.061	-0.036	0.076
PA	-0.137	-0.043	0.086	-0.143	-0.115	0.106	-0.047	**1**	0.043	-0.097	-0.055
WI	-0.155	-0.131	0.034	0.030	-0.073	0.040	0.061	0.043	**1**	0.126	**0.282**
WP	**-0.297**	-0.127	**0.254**	-0.107	-0.228	0.060	-0.036	-0.097	0.126	**1**	-0.072
TR	0.008	-0.072	-0.029	0.152	-0.025	0.153	0.076	-0.055	**0.282**	-0.072	**1**

Another result from the Factor Analysis is a biplot, as shown in Figure 3-4. As used in FA, the axes of a biplot are a pair of extracted factors. These axes are drawn in black and are labeled F1, F2 in this case. There is another plot for F2, F3, etc. which are not shown here.

A biplot uses vectors to represent the coefficients of the variables on the factors. Both the direction and length of the vectors can be interpreted. Vectors point away from the origin in some direction. A vector points in the direction that is most like the variable represented by the vector. This is the direction which has the highest squared multiple correlation with the factors. The length of the vector is proportional to the squared multiple correlation between the fitted values for the variable and the variable itself. For example, in Table 3-3 DE is loaded on Factor 1 with 0.871, thus the vector representing the DE variable has a value of 0.871 on the F1 axis in Figure 3-4.

The fitted values for a variable are the result of projecting the points in the space orthogonally onto the variable's vector (to do this, you must imagine extending the vector in both directions). The observations whose points project furthest in the direction in which the vector points are the observations that have the most of whatever the variable measures. Those points that project at the other end have the least. Those projecting in the middle have an average amount. For example, the perpendicular line from MP to the F2-axis intersects in 0.541, while MR in 0.276.

Thus, vectors that point in the same direction correspond to variables that have similar response profiles, and can be interpreted as having similar meaning in the context set by the data.

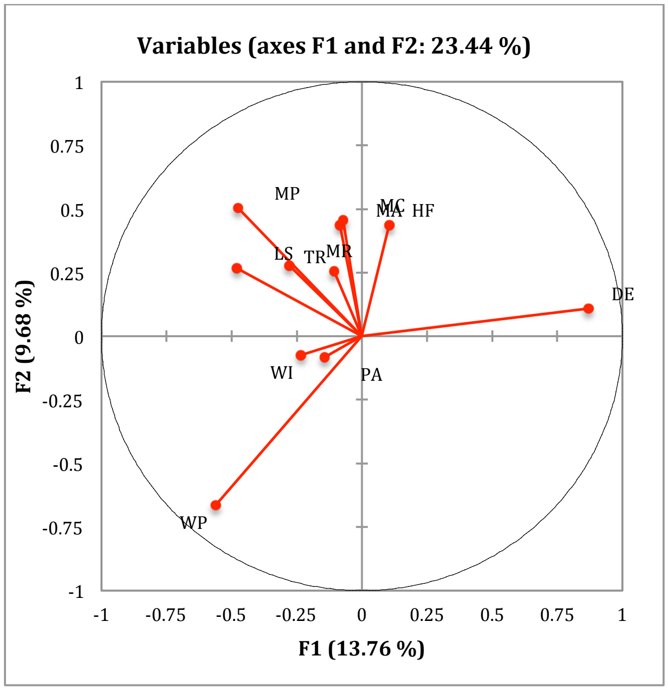

Figure 3-4 Biplot Vectors.

The biplot uses points to represent the scores of observations on the factors, and in Figure 3-5, each numbered point represents one of the condition reports, and the vectors represent the causes (variables). The relative location of the points can be interpreted in the following manner: points that are close together correspond to observations that have similar scores on the factors displayed in the plot. To the extent that these factors fit the data well, the points also correspond to observations that have similar values on the variables. In this example, events that are close together are ones that have similar profiles of causes.

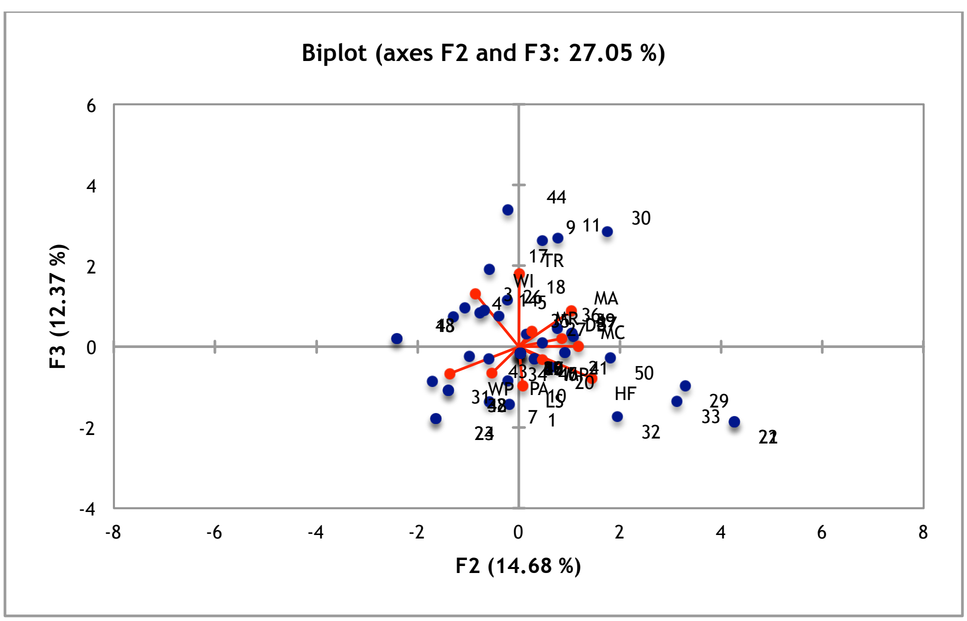

Figure 3-5 Scatter Plot of Significant CAP Events.

3.6 Slicing and Dicing the Data (Factor Analysis)

Since factor analysis is a statistical analysis procedure that has been used by researchers to analyze patterns of relationships among individual variables to produce a smaller set of ‘factors’ that summarize the unique relationships among the variables, it serves as a composite measures of the variables (Hallbert, 2005). In the case of the analysis performed using causal data, the goal was to determine whether relationships among factors exist and are reducible to a stable set of factors through which their effects can be uniquely expressed on events that occur at the NPP during operation. Thus, rather than assessing the eleven causes (or 62 when all the sub-causes are considered) separately and treating their influences as independent of one another (an approach that was proven not viable in Chapter 1, due to the necessity of having a combination of supervision and worker errors or holes in the barriers), factor analysis may be used to identify a factor structure employing fewer causes that are tractable, predictive, and easier or more efficient to assess during analysis than the original factor set. The factor loadings are contained in Table 3-3. In this section, we will examine some of the specifics of the data in more detail.

In order to understand the data, results and robustness, we did several tests.

1: Examine consistency in the data over the years.
2: Examine sample size and ratio of samples to variables.
3. Examine correlation matrix.
4: Examine correlations between pairs of causes.
5. Examine factor loadings between each cause and factor.

1: Examine consistency in the data over the years.

Factor analysis is a statistical analysis procedure that has been used by researchers to analyze patterns of relationships among individual variables to produce a smaller set of ‘factors’ that summarize the unique relationships among the variables and are capable of serving as composite measures of the variables (Hallbert, 2007). In the case of the analysis performed using causal data from the CAP database, the goal was to determine whether relationships among factors exist and are reducible to a stable set of factors through which their effects can be uniquely expressed on events during operation normal at NPPs. Thus, rather than assessing the eleven causes separately and treating their influences as independent of one another (an approach that is already shown not to be viable in Chapter 1 where it is shown that causes tendo to include supervision or management and worker errors), factor analysis may be used to identify a factor structure employing fewer causes that are tractable, predictive, and easier or more efficient to assess during analysis than the original factor set. Since exploratory factor analysis was used to discover an underlying set of factors, the robustness of the data is desirable. Since the database consists of 10 years of CAP data, it was necessary to analyze the consistency of the data over time and over the level of reporting. There were some changes in coding; however, this was traced during the research conducted here. For example, in many of the analysis done, the cause code PA was eliminated since it was not labeled PA until after 2007. Also, as described in Chapter 2, the CAP reporting is a low level database and this was constant throughout the years studied.

In order to examine the consistency over time, different slices of data were used in the factor analysis. For example, Figure 3-6 shows the results of the correlations for slices of 120, 240 and 360 rows from the data. As can be seen, the results are similar in that the same combinations are the most important ones; that is, those relations having the highest correlations. In this NPP, they are DE-WI, DE-WP, HF-MA, MP-MA, MC-MR, and WI-WP.

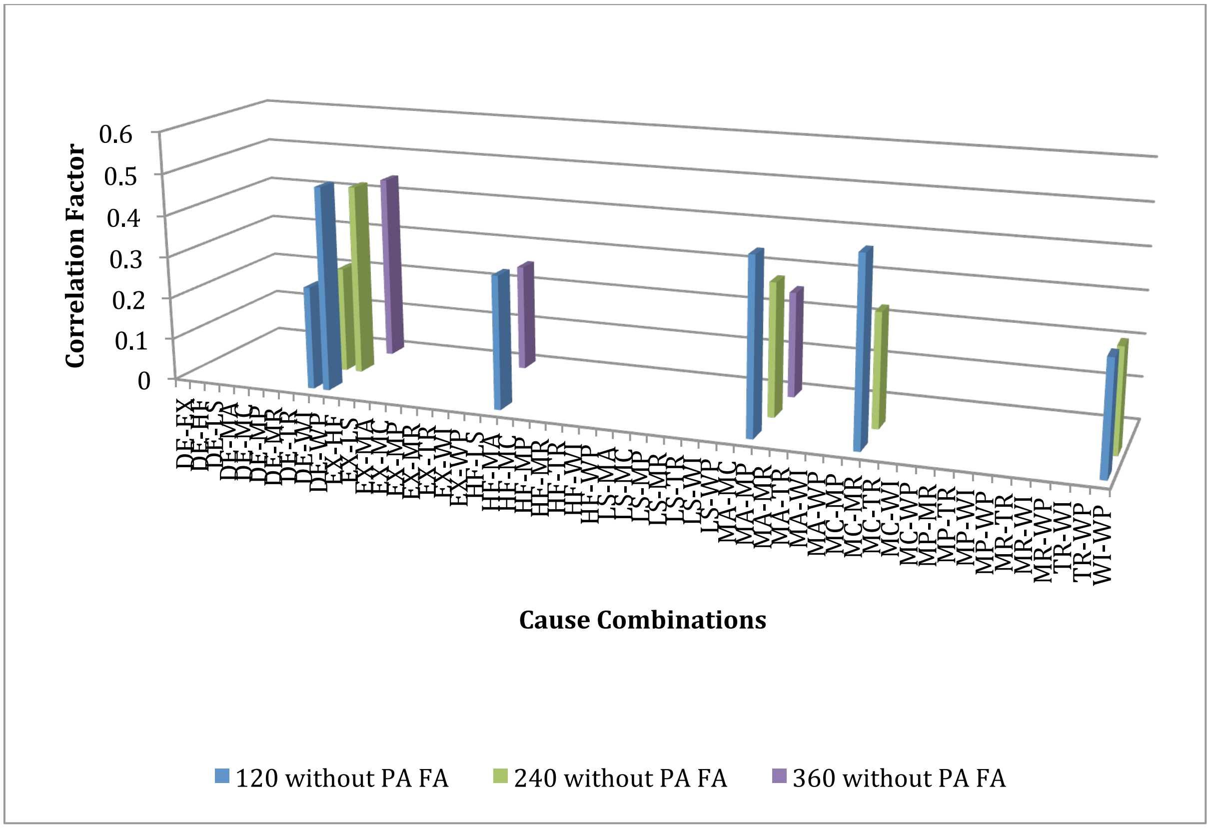

Figure 3-6 Results from Factor Analysis without PA, correlations are absolute values greater than 0.240.

2. Examine sample size and ratio of samples to variables.

Costello & Osborne observed many studies that had ratios of samples to variables much smaller than used in this dissertation, as can be seen in Table 3-5, where only 3% of the studies have a ratio of more than 100 samples to 1 variable.

Table 3-5 Current Practice in Factor Analysis.

Sample to variable ratio	% of studies	Cumulative %
2:1 or less	14.7%	14.7%
> 2:1 ≤ 5:1,	25.8%	40.5%
> 5:1, ≤ 10:1	22.7%	63.2%
> 10:1, ≤ 20:1	15.4%	78.6%
> 20:1, ≤ 100:1	18.4%	97.0%
> 100:1	3.0%	100.0%

In order to examine how sample size affects the likelihood of errors of inference regarding factor structure of this scale, an analysis of variance was performed, examining the number of samples producing correct factor structures as a function of the sample size. The results of this analysis are presented in Table 3-6. As expected, larger samples tended to produce solutions that were more accurate. Only 10% of samples in the smallest (2:1) sample produced correct solutions, while 70% in the largest (20:1) produced correct solutions. Further, the number of misclassified items was also significantly affected by sample size. Almost two of thirteen items on average were misclassified on the wrong factor in the smallest samples, whereas just over one item in every two analyses was misclassified in the largest samples. Given that the CAP database has a ratio much greater than 20:1, in fact it is closer to 10000 samples for each variable; thus making it a solid database for factor analysis in this regard.

Table 3-6 The Effects of Sample to Variable Ratio on Exploratory Factor Analysis (Costello & Osborne, 2005).

Variable:	2:1	5:1	10:1	20:1
% samples with correct factor structure	10%	40%	60%	70%
Average number of items classified on wrong factor	1.93	1.20	0.70	0.60

3. Examine correlation matrix.

Cause selection begins with correlation analysis on the full cause set. Ways to reduce the cause set include merging similar causes (e.g., MR1: Manpower Inadequate, MR2: Budget/Funding Inadequate and MR3: Prioritization/Scheduling of Activities Inadequate become Resource Management, MR), dividing a cause into more categories (e.g., Procedures becomes Procedure Availability and Procedure Quality) or eliminating the cause entirely. The analyst should examine the correlation results to identify correlations that are erroneous (e.g., correlations exceeding 0.95 tend to be erroneous in the data). Depending on the correlation technique used, the correlation set can be further reduced based on significance values or sensitivity analysis. Once the set of causes has been analyzed, the initial structure of the model is developed by using correlation analysis. Each cause becomes a node in the model and arcs are drawn between variables with correlations > |m|. For analysis of the current data, m = 0.24; this correlation cut-off value may be adjusted for different data sets (Groth, 2009). First we observe the correlation matrix in Figure 3-7, which shows the correlations among the causes of the CRs, in 3-D to get an overall picture of the relationships. Inspection of the correlation matrix shows the interrelationships among the variables. This correlation matrix shows that there are positive relationships among the variables, and the relationships within some subsets of the variables are higher than others. For example, MP and MA have a stronger positive relationship than MP and LS.

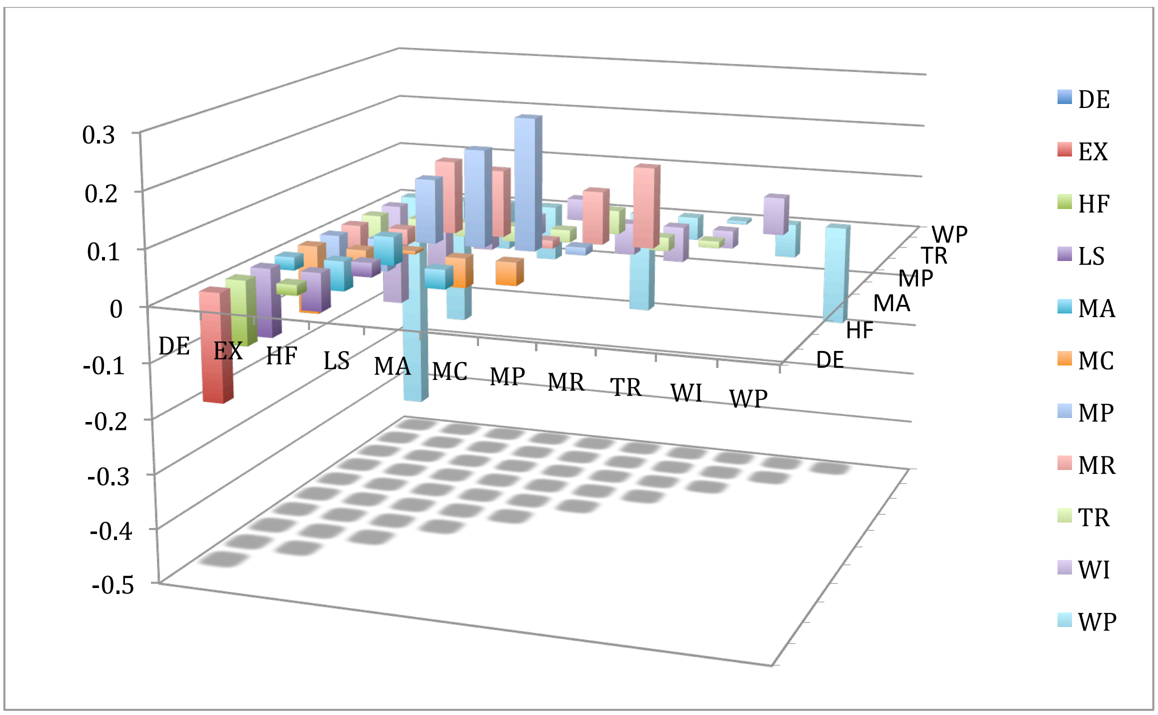

Figure 3-7 Results from Factor Analysis without PA, all CRs (correlation matrix results).

Perhaps, the 2-D graph presents the correlations more clearly as can be seen in Figure 3-8.

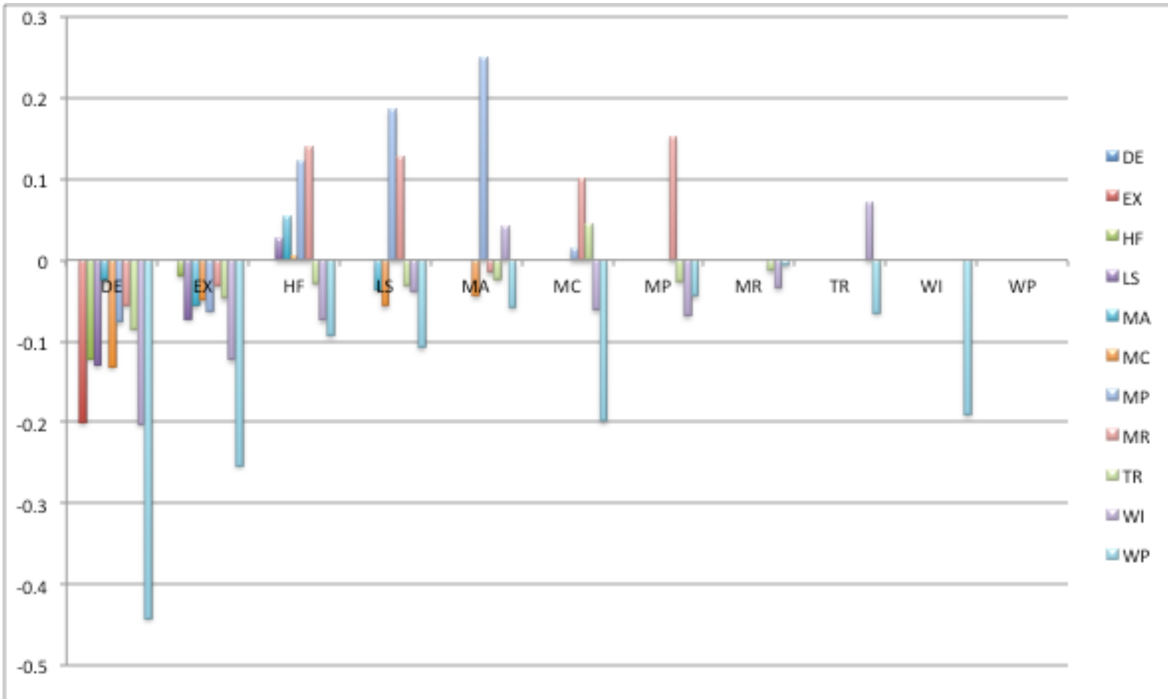

Figure 3-8 2-D Representation of the correlations.

4: Examine correlations between pairs of causes.

Another test is to observe the correlation factors for the cause combinations by severity evel. Figure 3-9 shows the cause combinations for the different severity levels. It can be observed that the DE-WP has the strongest correlation at all levels of severity.

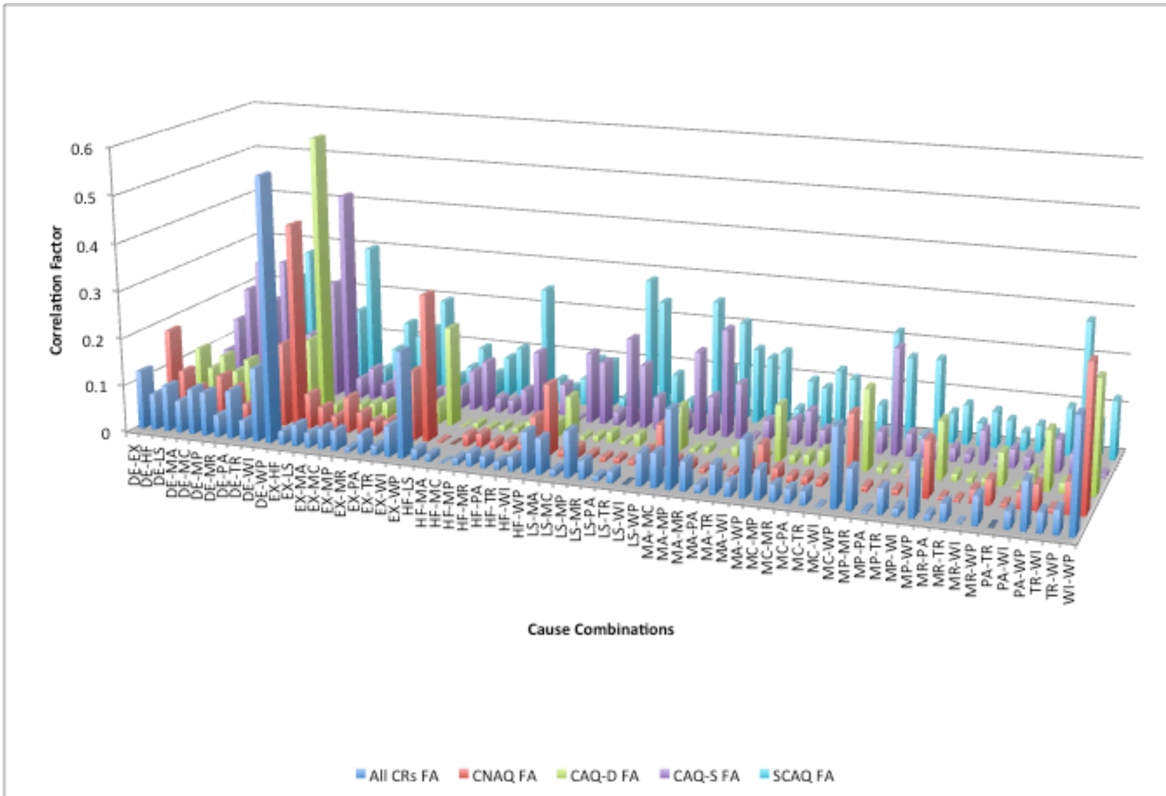

Figure 3-9 Cause combinations.

And, Figure 3-10 shows only those combinations with an absolute value greater than 0.24. These results are used to determine the arcs in the Bayesian Networks, as described in section 3.5.

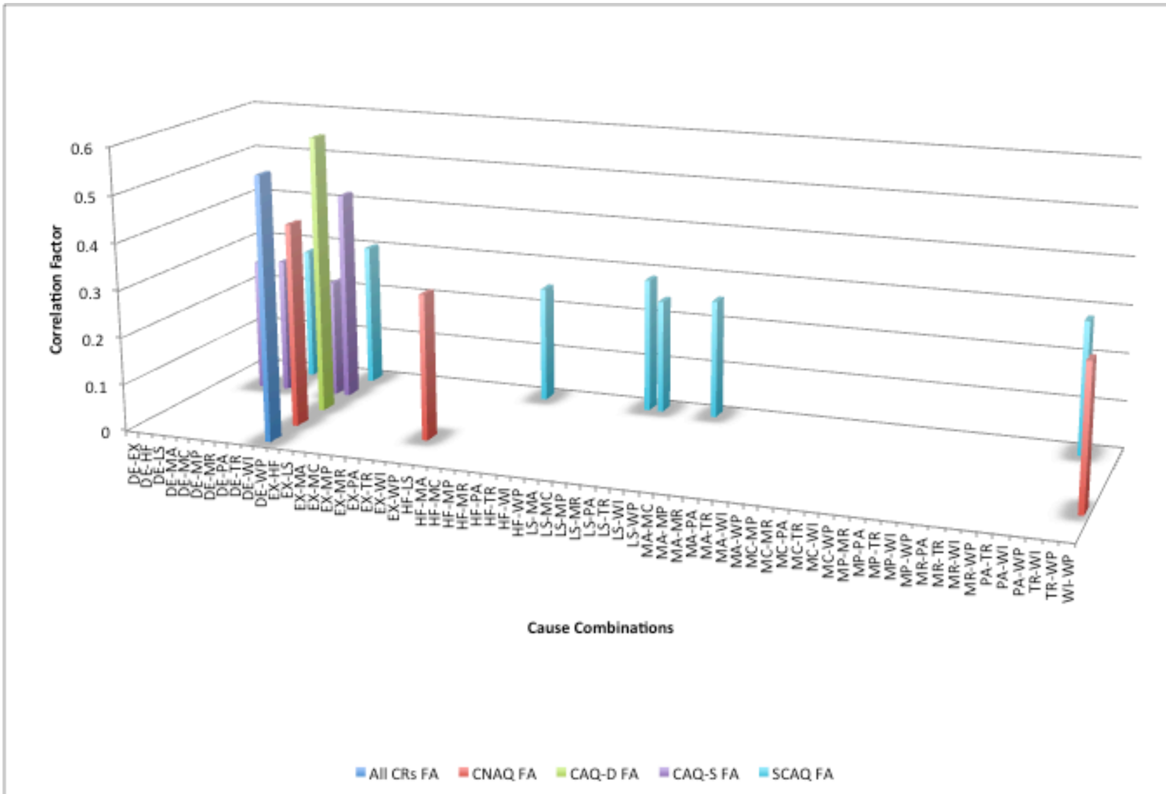

Figure 3-10 Results from Factor Analysis without PA, correlations are absolute values greater than 0.240, for different severity levels.

3.7 Factor Loadings

5. Examine factor loadings between each cause and factor.

The examination of the factor loading of each cause can enrich our knowledge about the areas or error forcing contexts existing at the NPP. Figure 3-11 shows the factor loadings for severity level CAQ-L1, which corresponds to the station level or CAQ-S. Figure 3-12 shows largest factor loadings for each of the severity levels. Thus, we are able to direct arcs from each of the causes to the Factor on which its load is highest. Figure 3-13 shows the factor loadings for CAQ-S and SCAQ severity levels, and Figure 3-14 shows the largest factor loadings for these two severity levels.

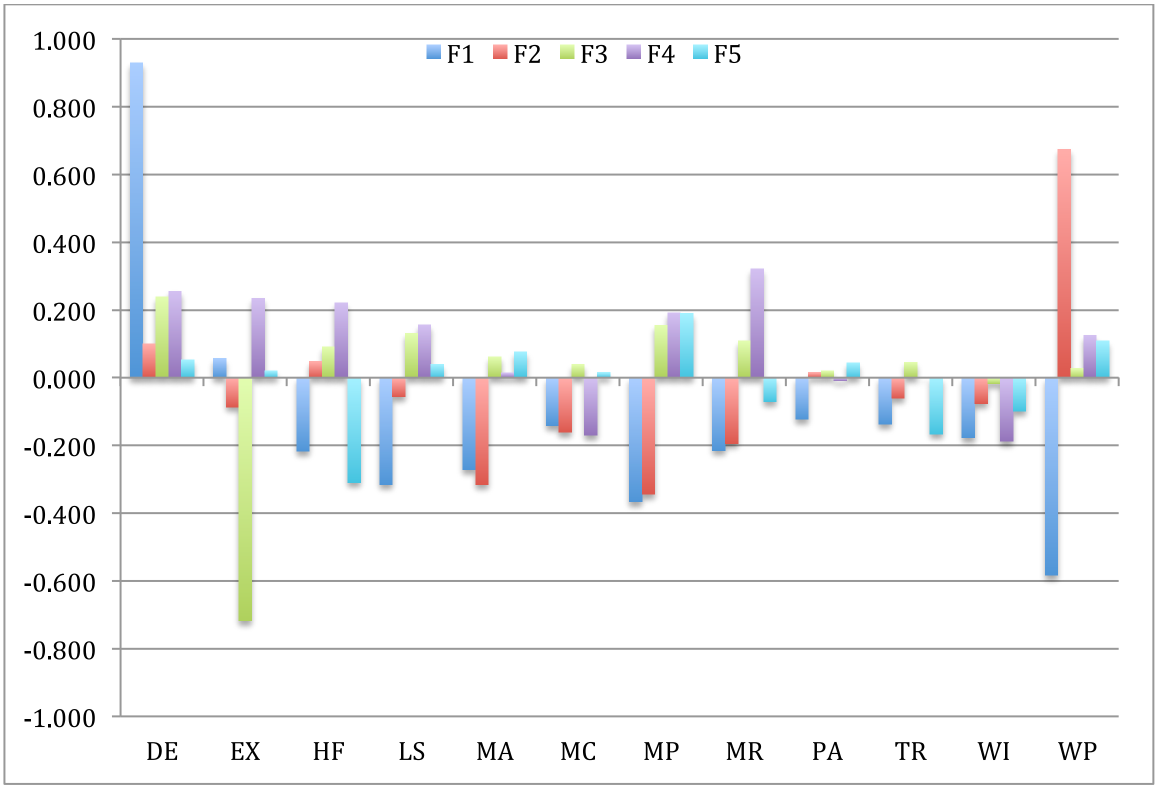

Figure 3-11 Factor loading on first 5 factors for CAQ-L1.

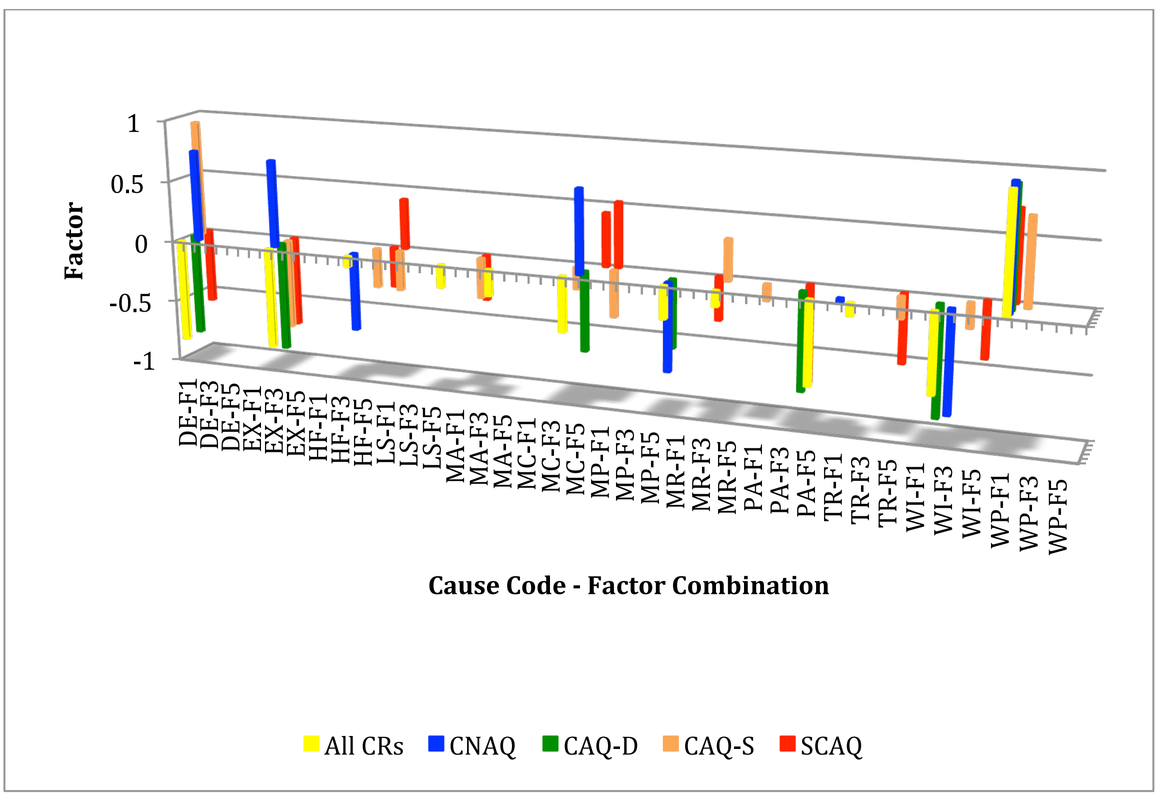

Figure 3-12 Largest Factor loadings for severity levels.

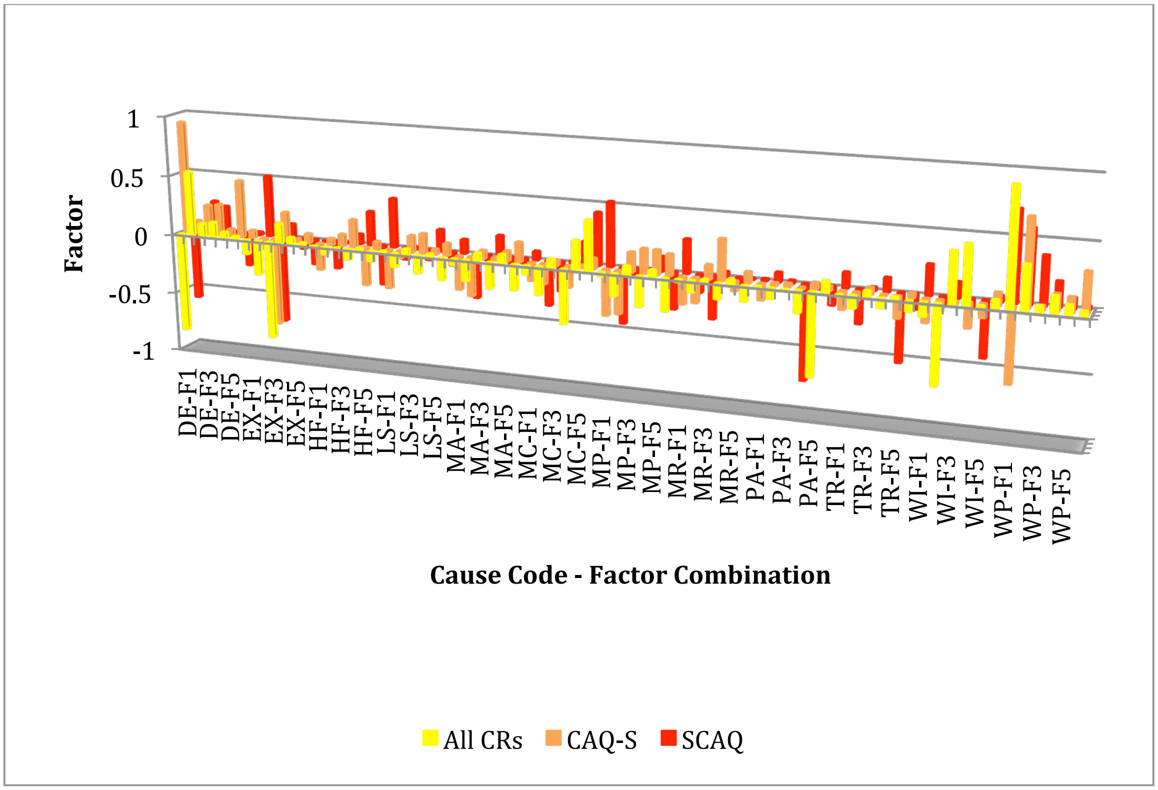

Figure 3-13 Factor loadings for highest severity levels.

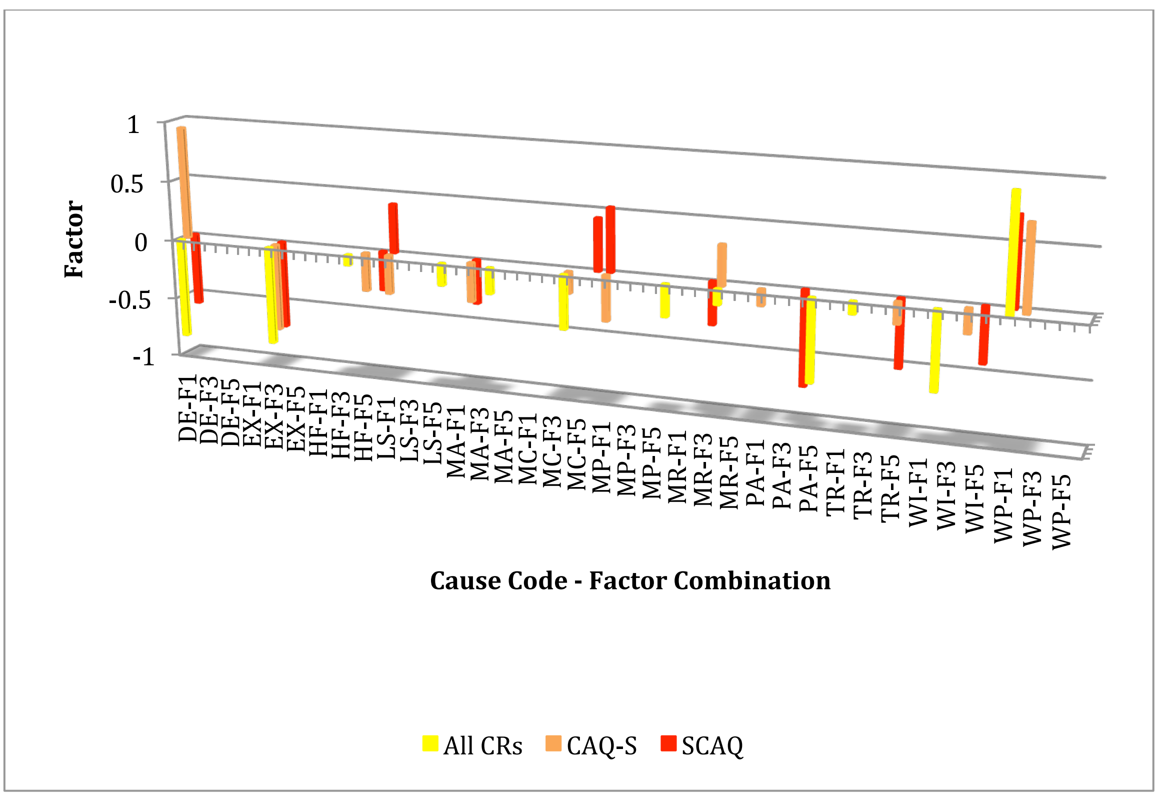

Figure 3-14 Largest Factor Loading for highest severity levels.

3.8 Building a Bayesian Network

Figure 3-15 illustrates the model learned from the data analyzed from the pilot plant's CAP database. From the factor analysis, four factors were maintained. Thus, we can observe that the model identifies the variables loaded on the four factors, previously identified in the factor loadings and the arcs represent the correlations between the variables.

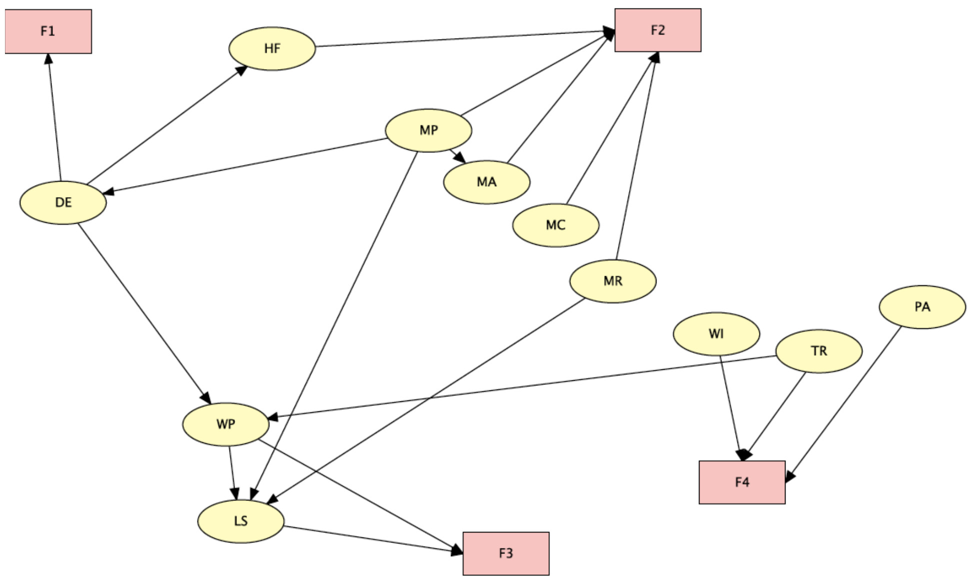

Figure 3-15 Causal model for 11 variables.

Figure 3-16 is the same model, just ordered in such a manner as to be able to observe more clearly the parent and child nodes. The quantification of the model is discussed in the next section.

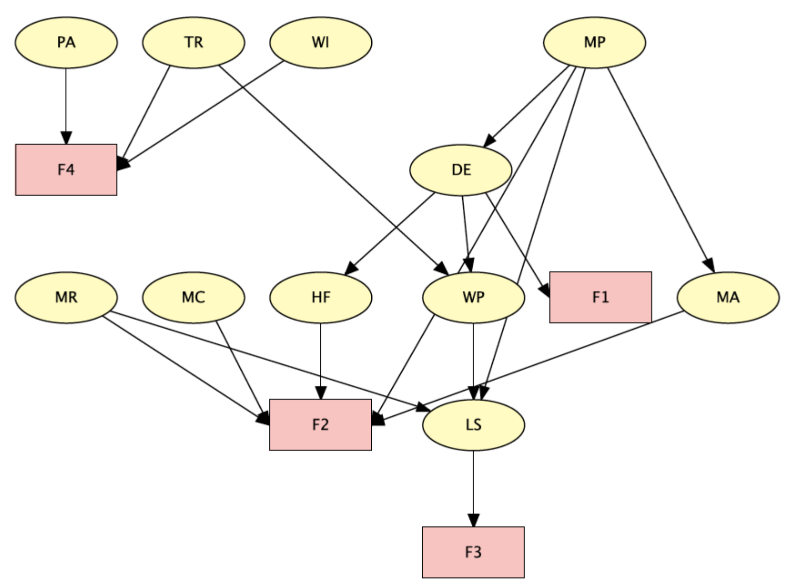

Figure 3-16 Final bayesian network for 11 variables.

3.9 Quantification of Consequential Events

The retained factors are patterns of variance identified by factor analysis and represented by the pink boxes, labeled F1 to F4 in Figure 3-16; each eigenvector retained forms one factor, summarized as the factors that divide the human performance difficulties into four categories: (F1) maintenance programs, (F2) management issues, (F3) work practices and supervision, and (F4) training, procedures and instructions.

Patterns of variance identified through factor analysis are traditionally labeled latent variables. However, in this dissertation, these patterns are interpreted as did Groth (2009) in a recent research project. That is, these factors represent different areas in the plant that affect the events that occur; for example, given problems in maintence programs cause more errors occurring during maintenance. While this is straight forward, given we are analyzing human failure events from the CAP database, the observed patterns can be viewed as visible manifestations of the context underlying the error. This interpretation is justified for factors with eigenvalues greater than 1.0. An eigenvalue greater than 1.0 indicates that its eigenvector accounts for more than its proportional share of variance. Each factor is a group of causes that contributes more to human performance errors than would each cause if acting alone; the whole factor is greater than the sum of its parts.

Model quantification requires populating a full probability table for each node. The methods used to convert correlation into conditional probability are discussed in this section. Given a well-populated database, conditional probability tables for the BN can be developed automatically. In fact, both the network structure and the conditional probabilities can be automatically “learned" given sufficient data (Cowell, R., Dawid, A., Lauritzen, S., & Spiegelhalter, D., 1999). CAP data is a candidate, given that automatic quantification requires a large sample size.

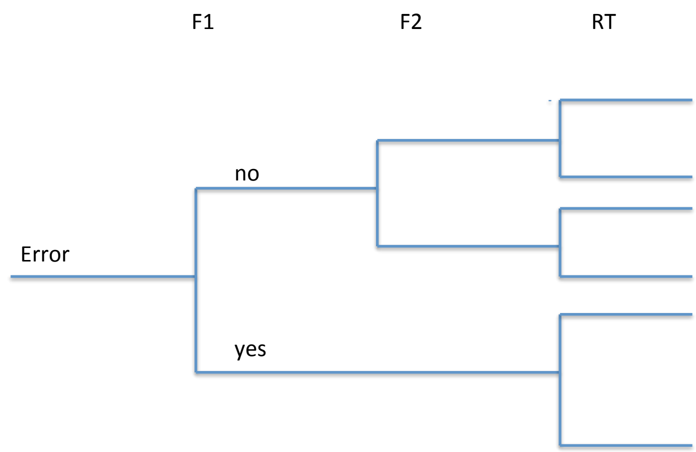

Figure 3-17 Possible effects of factors on occurrence of reactor trip given an error.

From Figure 3-17 we can see that the tree explains how the Factors (Fi) can be used to describe whether a reactor trip (RT) is likely to occur given a particular error or event (E). Given an error (E), and considering only two of the four factors for simplicity, the presence of F1 or F2 may have an impact on occurrence and likelihood of a Raector Trip (RT). Equation 3-2 expresses the manner to calculate the probability of a reactor trip given a factor and an error or event, using Bayes theorem. It is the way the Bayesian network works for the causal model presented in Figure 3-16.

$$P(RT \mid Fi,E) = \frac{P(Fi \mid RT,E)xP(RT \mid E)}{P(Fi \mid RT,E)xP(RT \mid E)+P(Fi \mid \overline{RT},E)xP(\overline{RT} \mid E)} \qquad (3\text{-}2)$$

Where:

P (RT | F_i, E) = the probability of Reactor Trip given a specific Factor, with event type E;

P(F_i | RT, E) = the likelihood of observing Factor F_i given a Reactor Trip and event type E;

P(RT|E) = probability of Reactor Trip with event E.

The values can be obtained through CAP data, using analysis and the database specifications indicated in Chapter 1.

3.10 Procedure for Barrier Analysis

It is proposed here that when the resilience threshold is approached, some action should be taken to reduce the chance of a significant event; this can be considered as filling the hole in the defenses against accidents, according to James Reason's Swiss cheese analogy (Figure 3-18) (Reason, 1995). This plug is considered as a barrier in this thesis. A barrier can be anything from placing a hold point or warning in the corresponding procedure, requiring a photograph of the aligned system after the test or maintenance activity to hiring another person to conduct an independent verification. Obviously each of these has its costs and savings. For example, the cost of the photograph includes the cost of the camera, the software, placing adequate administrative controls, etc.

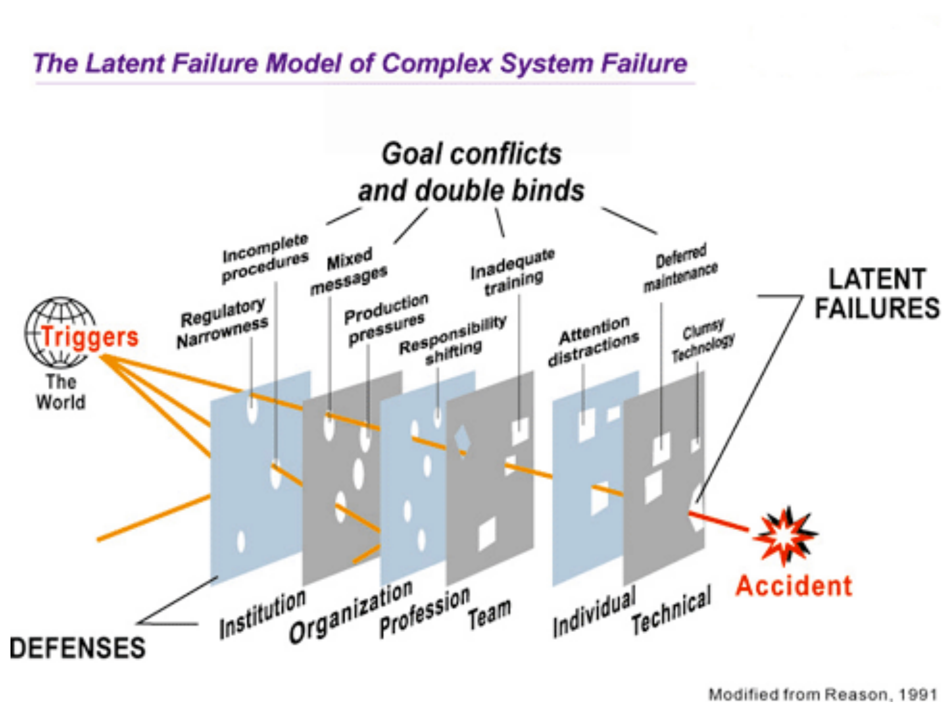

Figure 3-18 James Reason's Swiss Cheese Model (Reason, 1995).

There are several types of barriers to analyze, for example, there are those that can be placed in the procedures. As shown in Figure 3-19, it is necessary to first identify the risk significant procedures, or identify them from the CAP procedures (there is a field for Procedures in the CAP database utilized for this study). Next, in order to risk rank the procedure, the following questions must be asked:

A. Can the procedure cause reactor trip?
B. Can another safety system or train actuate?
C. Can risk significant equipment be actuated?
D. Is there equipment in the procedure that is used to change modes or for shutdown?
E. Is the equipment in the procedure necessary for emergency operations?

Once the risk-significant steps are identified, the applicable barriers are evaluated in order to rank the cost/benefit (utility) obtained, including the costs and savings incurred for each barrier.

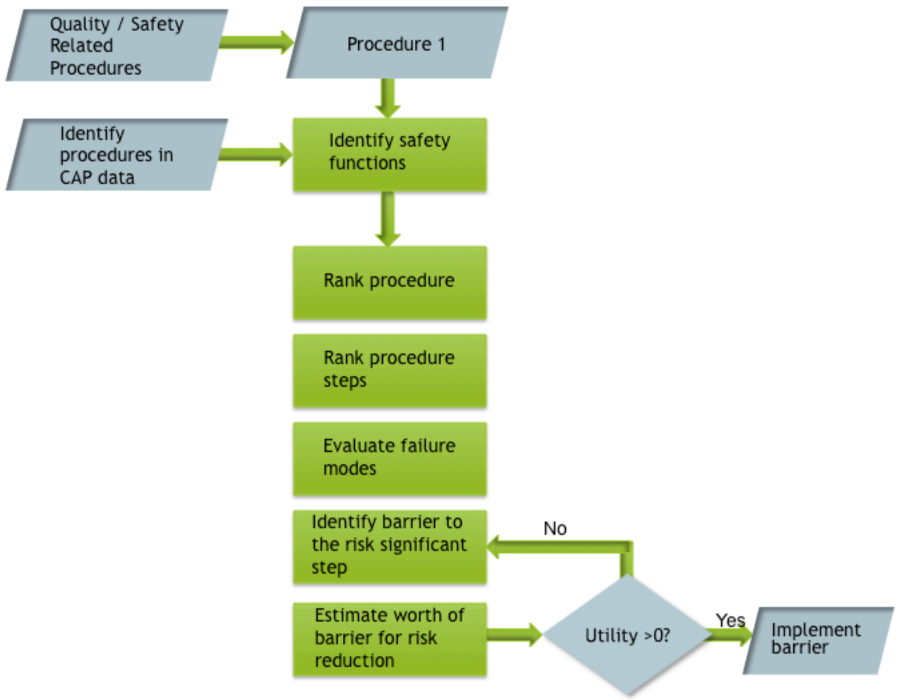

Figure 3-19 Process for Barrier Evaluation.

3.11 Predicting Reactor Trip

In order to predict an undesirable consequence, this node is added to the model. For this example, we add only one node, in this case, reactor trip (RT), to simplify the explanation of the method. However, before we can create the influence diagram, we must develop the graphical representation of the relationship between the variables and the consequences (defined as reactor trip in this example). The structure of the resulting model (Figure 3-20) was learned directly from the data. While the sample size is sufficient for the statistical analysis, it may be too scarce to directly determine all the arcs (correlations between the variables) directly using the HUGIN program (Anderson, Olesen, Jensen & Jensen, 1989). The node HF, which is not linked, evidences this effect. A column was added to the data table for reactor trip, and its value was set to 1 in rows where the event caused or terminated in reactor trip. Thus, this BN can be used to determine the effect of reducing adverse impacts of the underlying causes of the events on the probability of reactor trip. One result shows that the elimination of procedural adherence errors would decrease the probability of reactor trip by one third. This particular BN was learned using the Greedy algorithm (Chickering, 2002).

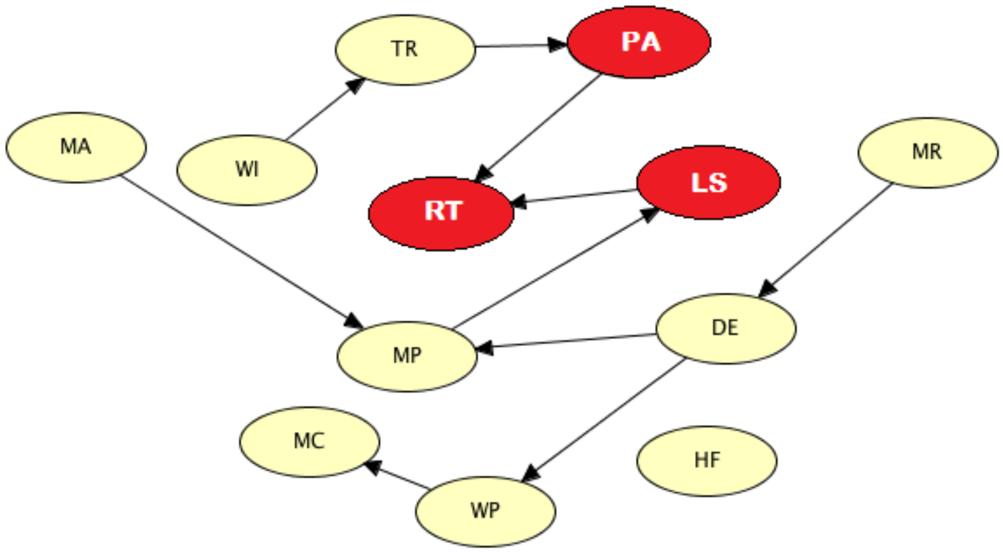

Figure 3-20 Learned Structure of the BN.

3.12 Results

Now we have the information necessary to be able to construct the influence diagram. For the purposes of this chapter, we shall concentrate on only one part of the diagram. As shown by the red nodes in Figure 9, the probability of reactor trip (RT) depends directly on procedure adherence (PA) and on Supervision (LS). Thus, we will amplify this section of the network. We add a decision node, that is, should plant management implement a human performance barrier to reduce the probability of reactor trip from these cause sources. In this case we are referring to human factors/organizational barriers, not actual physical barriers. The pink box is used to represent “treatments” or “aids” that can be employed to reduce the probability of the undesired event. These treatments are considered barriers or defenses used in nuclear power plants to aid human performance, such as 3-way communication, process and procedure approvals, pre-job briefings, etc. The pink box in Figure 10 represents this decision node. We can also add a cost node and a utility node (savings acquired from avoiding a reactor trip), represented by the green diamonds in this same figure.

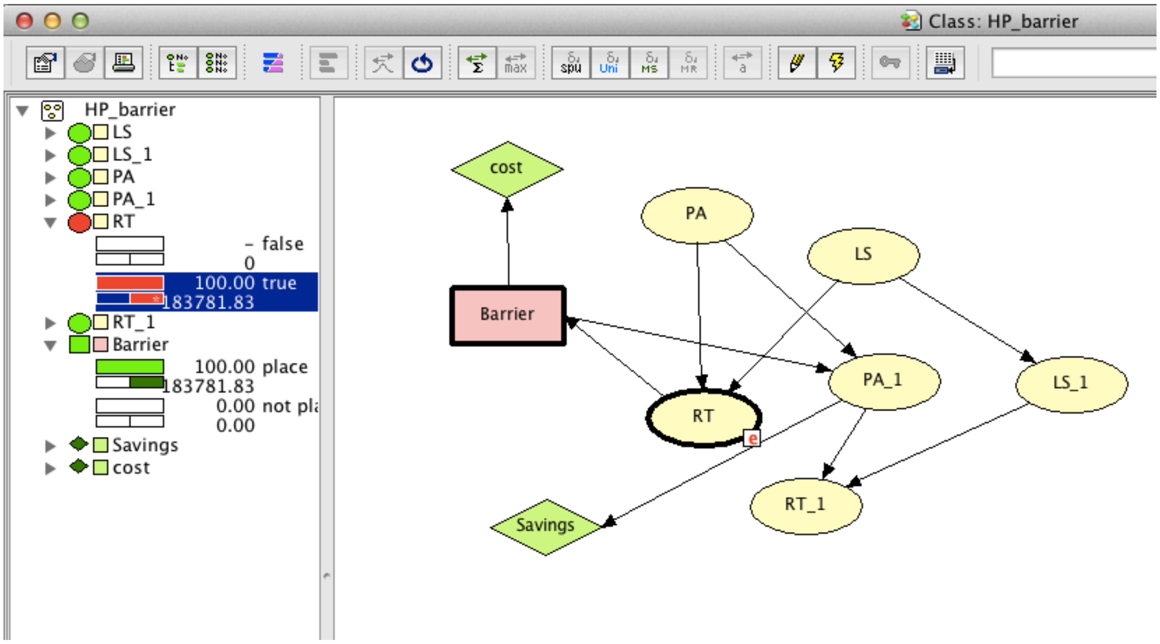

Figure 3-21 The Influence Diagram propagated with RT=100%.

For this example, it is postulated that operating experience has shown an unacceptable level in the occurrence of reactor trips due to human errors. The model indicates that procedure adherence (PA) and supervision (LS) are key contributors to reactor trip, thus we propose implementing a barrier to PA (e.g., additional procedure approvals). The implementation of the barrier influences the PA variable to become PA_1, and so forth for each node. The process shown in this chapter allows the proposed barrier to be evaluated in terms of cost and savings to determine its viability, and it provides a means to determine its effectiveness in reducing the occurrence of future reactor trips. In this hypothetical situation, the barrier was determined to have an expected utility of nearly $184,000 (i.e., benefit) and an associated organizational implementation cost of $40,000. Thus, due to high value relative to cost, the analysis indicates it is advantageous to implement the barrier. A similar analysis could be performed for the LS node. This demonstrates a key benefit of this approach in that focused barriers can be specifically structured to target and improve specific human and organizational performance activities relative to consequential historical plant events such as reactor trips.

It is recognized that important improvement initiatives fail because underlying problem causes are not well understood or because corrective actions specified do not align well with the analyzed causes (INPO, 2005). Therefore, it is necessary that as improvements are being made, the problems that continue to occur are documented in new condition reports. Because these new condition reports address a problem already being pursued, the emergent condition reports are reviewed by management and then closed without specifying additional action beyond the immediate corrective action, using as justification for the closure the active adverse trend condition report and ongoing improvement project. Such condition reports are then included in trending efforts to gauge the degree of improvement being achieved. The tools developed here should be continually applied, in this way the probabilities are updated and become valuable as trending tools, in a ddition to decision tools.

3.13 Discussion

Since a small set of data was used for the development of the models relative to the amount of data contained in Corrective Action Programs, the quantitative results are preliminary. However, the models developed in this chapter are functional and the results are promising for several reasons: (1) the methodology enables the incorporation of operational experience into the model by using information from the Corrective Action Programs at nuclear power plants; (2) the models make it possible to identify and incorporate organizational factors into the probabilities of human error in a meaningful way; and (3) the influence diagrams, developed from the Bayesian networks, enable the user to evaluate the utility of adding human performance barriers or other organizational effectiveness initiatives and calculate their effect on undesirable consequences in a nuclear power plant caused by human error during routine operation and maintenance activities.

It is important to emphasize that the purpose of these models is to illustrate the type of insights that can be gathered through the model development process and data collection effort, as well as to provide a road map for future model development and data collection process improvements. Despite the many limitations of the data, the models are useful and the uncertainty in the results will be reduced by additional data collection and associated screening. Additional work will be performed to develop a more comprehensive model and data screening process to support the development of a database and associated data processing specification (e.g., to define the necessary data to be collected, etc.) for Corrective Action Programs that would further support and facilitate an analysis such as described in this chapter. This work will contribute to the development of a method for trending and tracking human and organizational performance events, as well as associated causes to support efforts in improving knowledge management and organizational effectiveness.

Conclusions

This final chapter reviews and summarizes the dissertation research, identifies the main tools and methodologies developed, and discusses the implications of this study to the nuclear industry. This study originated at the request of a supervisor at a nuclear power plant who had received the assignment to determine apparent and root causes of human performance errors that occurred during a periodic surveillance test that resulted in a plant trip. The idea was to examine the feasibility of risk-informing nuclear power plant operations and maintenance procedures. The NPP, as well as the nuclear industry, was interested in this particular topic due to increasing number of human errors and the resulting increased number of consequential events, in particular undesired SCRAMs. Rather than concentrate soley on improving procedures, this dissertation sought to develop processes to identify precursors to human error and provide methods to reduce human error once identified.

Given that maintenance and testing (e.g., surveillance testing) of reactor systems are required activities and are also important causes of consequential events (e.g., unplanned reactor trips, turbine trips, down-power events, inadvertent system actuations, damage to the plant equipment, plant personnel safety, public health and safety), it is essential to find ways to identify situations where human errors are more likely to occur and then provide the tools and mechanisms to reduce the likelihood of these undesired events.

Due to regulatory requirements for Problem Identification and Resolution (PI&R) programs, nuclear power plants have large amounts of plant event data. This data exists in several areas but is comprehensively contained in plant specific Corrective Action Programs (CAP). The regulatory requirement for PI&R programs was due to the overarching importance of reactor safety. PI&R Programs (i.e., CAP) provide the base experiential data to enable NPPs to design methods to track and trend events in order to provide a process to facilitate quantitative trending analyses, provide feedback and lessons learned from plant experience, and to provide a tracking mechanism for correct actions and enhancements. NPP programs relative to specific and industry operating experience are based on actual plant events as documented in Corrective Action Programs and other industry programs (e.g., NRC Licensee Event Reports, INPO Significant Operating Experience Reports). For this dissertation, it was determined that once developed such a database and associated computational algorithms could be maintained and updated by plant personnel to provide insightful trends into human performance and also provide formal data sources for Human Reliability Analyses (HRA) not only on a plant specific basis but also on a fleet or industry basis. The methods described in this report will allow human error precursors to consequential operational events to be better identified and allow implementation of risk management methods to reduce the likelihood of events.

During the development of this study, it became apparent in the analysis of the CAP database that the majority of the condition reports were caused by combinations of organizational and individual errors. When a condition report ítem identified the occurrence of an equipment failure, the level of the condition report severity level assignment increased. While many databases exist for reporting equipment failures, the human performance aspect is often times not emphasized, except when it is directly related to the equipment failure. The analysis of the CAP database provided many insights into the importance of considering all organizational and individual errors as precursors to more severe outcomes and the necessity to include these factors in any quantitative tool developed for plant use.

This dissertation has introduced a methodology for using a Corrective Action Program database to develop tools to identify the emergence of plant and organizational factors adversely affecting the resilience of the plant-human interface with the intent to inform decision-makers so that appropriate risk management compensatory measures or other management directed actions can be implemented to reduce the likelihood of consequential events. The data was used to develop an organization specific resilience curve and convert it into a leading performance indicator with the purpose of identifying (quantifying) conditions of reduced organizational resilience and thereby reduce the likelihood of consequential events. A methodology was described to build Bayesian Networks of the causes of events and errors during normal operation. This latter model was converted into an influence tool for determining the best barrier to incorporate at the station to reduce the likelihood of a consequential plant event (i.e., plant trip). In summary, the conclusions and products of this research are the following:

1. Empirical proof that organizational stress and strain levels (i.e., resilience) and consequential plant events are related.

2. That organizational resilience is a factor in the likelihood of a consequential event.

3. A model of organizational resilience.

4. A leading performance indicator to track and trend resilience.

5. A tool to help in determining effective barriers to reduce the likelihood of a consequential event.

The model of organizational resilience is a linear relation developed from the complete CAP data. While an analogy to the material's science stress-strain graph has been presented by Woods & Wreathall (2008), the contribution of this work is in the development of a quantitative resilience model. Since the resilience identifies the ability of the organization to adapt to changes and withstand the increasing levels of work that are required to not only maintain the plant operating at a safe level, but also to perform the activities necessary to satisfy regulatory requirements, administrative requirements, and other activities necessary to maintain all plant programs, procedures, and processes currrent and acceptable. The model developed in this study also identifies the number of plant specific activities and events that have caused consequential events in the operational history of the plant.

Based on this model, a leading performance indicator was developed, and presented in a simple, easy-to-use manner for the plant personnel. This leading indicator should alert plant personnel to reduced organizational resilience margins and the associated risk of a consequential event (i.e., a plant trip). The intent of this performance indicator is to allow the station to initiate the evaluation of the plant situation and determine compensating actions to either increase organizational resilience margins or focus attention on activities associated with plant equipment whose failure or malfunction would result in a plant trip. These compensatory actions could include actions such as install a physical barrier or implement a new procedure, delay some work on trip sensitive equipment, provide additional management or supervisory oversight, etc.). This finding contributes to the decision making tasks for planning that should be directed by risk- informing the planning of activities in the near term, always considering the organizational resilience in order to continue with successful, safe operation of the plant.

The tool for determining the most adequate barrier to install or defense to implement was built from the Bayesian Network that was informed from the identified causes of the events contained in the CAP data. This influence tool can be used to conduct a cost/benefit analysis in a quick, informed manner and aid in the decision making process for plant personnel.

While the final presentations of these models and tools have been presented in a simple way, in order to make them practical tools for NPP personnel, the identification, research and development of these is the major contribution of this research. The fact that almost all the NPPs in the world have similar CAP databases opens the possibility of two areas for further research. The first, that similar organizational resilience tools could be built for each NPP. The second, that a large database could be created for human performance data and incorporate empirical data in the HRA methods for evaluating human error probabilities, as well as incorporate organizational factors in the calculations.

The limitations of this study is that the research was conducted on one NPP's operating experience and research will be required to determine how it can be applied to other NPPs and define the modifications, if any. Also, additional work will be needed to identify organization-specific consequential events and how those events are linked to consequential plant events. This area will be important in order to determine the significance applied to organization-specific activities and their true relation to consequential events. Additionally, organization-specific resilience level performance indicators could provide important leading information into those conditions where the resilience margin of a plant organization can be measured and subsequent organization-specific compensatory measures developed.

Overcoming this issue should bring valuable results to the nuclear industry in the treatment and incorporation of human and organizational performance into risk and safety studies. In this way, this development of a human performance monitoring and tracking methodology and tool can be deployed to nuclear plant organizations for the purposes of quantitatively measuring and monitoring human performance events and trends for the purposes of reducing the occurrence of consequential human errors. Other potential uses are for Procedure/Process changes, new organizational performance indicators, identification of important communication hold points, identification of conditions or situations where adverse organizational interactions are most likely to occur based on plant history.

.

References

Andersen, S.K., Olesen, K.G., Jensen, F.V., & Jensen, F. (1989). HUGIN – a Shell for Building Bayesian Belief Universes for Expert Systems. In Proc. of the 11th IJCAI, pp. 1080–1085 (Available at http:www.HUGIN.com).

Boring, R.L. (2012). Fifty Years of THERP and Human Reliability Analysis, PSAM11.

Boring, R.L., Gertman, D.I., Whaley, A.M., Joe, J.C., Tran, T.Q., Keller, M.D., Thompson, R.P., & Gerszewski, I.R. (2006). A taxonomy and database for capturing human reliability and human performance data. Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society, pp. 2217-2221.

Boring, R.L., Roth, E., Straeter, O., Laumann, K., Blackman, H.S., Oxstrand, J., & Persensky, J.J. (2009). Is human reliability relevant to human factors? Proceedings of the 53rd Annual Meeting of the Human Factors and Ergonomics Society, pp. 610-614.

Canavan, K., & Hannaman, G. (2004). Introduction to simplified generation risk assessment modeling. EPRI Report, EPRI 1007386.

Cheng, P.D. (2010). Bayesian Belief Network and Fuzzy Logic Adaptive Modeling of Dynamic System: Extension and Comparison, Professor Mohammad Modarres, Thesis UMD.

Chickering, D. (2002). Optimal structure identification with greedy search. Journal of Machine Learning Research, Vol. 3, No. 3, pp. 507-554.

Costello, A.B., & Osborne, J.W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis, Vol. 10 No.7, Practical Assessment Research & Evaluation, Available online: http://pareonline.net/getvn.asp?v=10&n=7.

Cowell, R., Dawid, A., Lauritzen, S., & Spiegelhalter, D. (1999). Probabilistic networks and expert systems. Springer Verlag.

Darwiche, A. (2009) "Modeling and Reasoning with Bayesian Networks". Cambridge University Press. United States.

Dhillon B.S., & Liu, Y (2006). Human error in maintenance: a review, Journal of Quality in Maintenance Engineering ; 12(1): 21-36.

EPRI (1999, October). EPRI Guidelines for Leading Indicators of Human Performance Preliminary Guidance for Use of Workplace and Analytical Indicators of Human Performance, TR-107315.

EPRI (2003). In-Service Use of the Human Performance Assistance Package Experience Using PAOWF and Leading Indicators of Organizational Health, 1007491.

Evans, M. (2011, April 25). Nuclear (http://www.earthtimes.org/encyclopaedia/environmental-issues/nuclear/).

Grecco, C.H.S., Vidal, M.C.R., Cosenza, C.A.N., dos Santos, I.J.A.L., & Carvalho, P.V.R. (2013, November 24-29). A Fuzzy Logic Based Method to Monitor Organizational

Resilience: Application In a Brazilian Radioactive Facility, 2013 International Nuclear Atlantic Conference - INAC 2013, Recife, PE, Brazil.

Groen, F. & Mosleh, A. (2005). Foundations of probabilistic inference with uncertain evidence. International Journal of Approximate Reasoning (39) pp. 49–83.

Groth, K. (2009). A Data-informed model of Performance Shaping Factors for use in Human Reliability Analysis, Ph.D. dissertation, University of Maryland, College Park, MD.

Hallbert, (2007). The Employment of Empirical Data and Bayesian Methods in Human Reliability Analysis: A Feasibility Study. NUREG/CR-6949. Washington, DC: US Nuclear Regulatory Commission.

Hallbert, B. (2006, March 7). Developing Human Performance Measures, Project Milestone Deliverable: Letter Report, JCN Y6843, INL.

Hallbert, B., Whaley, A., Boring, R., McCabe, P., & Lois, E. (2007). Human Event Repository and Analysis (HERA):The HERA Coding Manual and Quality Assurance, NUREG/CR-6903, Vol. 2. Washington, DC: US Nuclear Regulatory Commission.

Herrera, I.A. & Hovden, J. (2008, October 28-30). Leading indicators applied to maintenance in the framework of resilience engineering: A conceptual approach, The 3rd Resilience Engineering Symposium, Antibes-Juan Les Pins, France.

Hirschberg, S. (2004). Human reliability analysis in probabilistic safety assessment for nuclear power plants. CSNI technical opinion paper, OECD/NEA.

Hollnagel, E. (2010, May). How Resilient Is Your Organisation? An Introduction to the Resilience Analysis Grid (RAG). Sustainable Transformation: Building a Resilient Organization, Toronto, Canada.

Hollnagel, E. (2011). How Resilient Is Your Organisation?, https://hal-mines-paristech.archives-ouvertes.fr/hal-00613986.

Hollnagel, E., & Fujita, Y. (2013). The Fukushima Disaster – Systemic Failures as the Lack of Resilience, http://dx.doi.org/10.5516/NET.03.2011.078, Nuclear Engineering and Technology, Vol. 45, No. 1, pp. 13-20.

Hollnagel, E., Woods, D. & Leveson, N. (2006). Resilience Engineering: Concepts and Precepts, Ashgate.

IAEA (1988). Safety Series 75-INSAG-1.

IAEA (1992). Procedures for conducting probabilistic safety assessments of nuclear power plants (Level 1). IAEA Report, Safety Series No. 50-P-4.

IAEA (1995). Human reliability analysis in probabilistic safety assessment for nuclear power plants. IAEA Report, Safety Series No. 50-P-10.

IAEA (2012). NS Tutorial. Safety case study: the Three Mile Island accident (http://www.iaea.org/ns/tutorials/regcontrol/assess/assess3233.htm).

INPO (2005). Improvement at Nuclear Power Stations, INPO 05-005.

INPO (2010). Engaged, Thinking Organizations (SOER 10-02).

INPO (2013). INPO Consolidated Event System (ICES) Query for Human Performance-related Causes using Cube (database retrieved April 23, 2013).

Jensen, F.V. (2001). "Bayesian Networks and Decision Graphs". Springer.

Johnson, C.W. (1999). Improving the Presentation of Accident Reports over the World *Wide Web*. Proceedings of the 17th International Systems Safety Conference, J. Dixon(Eds.), The Systems Safety Society, Unionville, Virginia, United States of America, 396-405.

Kim, J., & Park, J. (2008, December). Task Types and Error Types Involved in the Human-Related Unplanned Reactor Trip Events, Nuclear Engineering and Technology, Vol. 40, No. 7.

Kim, M. C. & Park, J. (2011), Development of a path model for human-induced unplanned reactor trips in nuclear power plants. Quality and Reliability Engineering International, 27: 141–147. doi: 10.1002/qre.1097.

KINS (2013). Korea Institute of Nuclear Safety's Operational Performance Information System (OPIS, http://opis.kins.re.kr).

Kirwan, B., Basra, G., & Taylor-Adams, S.E. (1997). CORE-DATA: A computerized human error database for human reliability support. Proceedings of the 1997 IEEE Sixth Conference on Human Factors in Power Plants, pp. 9-7 – 9-12.

Koller, D. & Friedman, N. (2009), "Probabilistic graphical models: principles and techniques", MIT Press.

Kosko, B. (1993) Fuzzy Thinking, The New Science of Fuzzy Logic, Hyperion, New York, NY.

Laakso, K., Pyy, P., & Reiman L. (1998). Human errors related to maintenance and modifications. STUK-YTO-TR 139, STUK.

Magwood, W. (2015, July 14). The human side of nuclear safety and the future of nuclear energy, World Nuclear News.

MathWorks (2004). Fuzzy Logic Toolbox for MATLAB, Version 2.1.1.

Murphy, K. (1998). A Brief Introduction to Graphical Models and Bayesian Networks (http://www.cs.ubc.ca/~murphyk/Bayes/bnintro.html#appl).

NEI (2010). Corrective Action Processes for New Nuclear Power Plants During Construction. NEI-08-02 rev 3.

Nelson, P., & Martín-del-Campo, C. (2014, June 22-27). Use of Corrective Action Programs at Nuclear Plants for Knowledge Management. PSAM 12.

Nelson, P., & Martín-del-Campo, C. (2014, September). Evaluation of Barriers and Resilience to Improve Organizational Performance in Nuclear Power Plants, Proceedings from Annual Mexican Nuclear Society conference, Boca del Río, Veracruz.

NRC (1989, January 24). Nuclear Regulatory Commission's "Policy Statement on the Conduct of Nuclear Power Plant Operations" (54 FR 3424).

NRC (2015a, February 26). Inspection Manual, Inspection Procedure 71152, Problem Identification and Resolution.

NRC (2015b, April). Human Factors Information System, http://www.nrc.gov/reading-rm/doc-collections/human-factors/.

Pariès, J., Wreathall, J., Woods, D., & Hollnagel, E., eds. (2012, October 1). Resilience Engineering in Practice: A Guide Book, Ashgate Publishing, Ltd.

Pearl, J. (1988) "Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference". Morgan Kaufman Publishers, Inc. San Francisco, California.

Pence, J., Mohaghegha, Z., Kee, E., Hubenak, R., Billings, M.A., & Dang, V. (2015, April 26-30). Modeling Organizational Factors in Human Reliability Analysis, International Topical Meeting on Probabilistic Safety Assessment and Analysis, PSA 2015, Sun Valley, Idaho.

Perrow, C. (1984). Normal Accidents: Living with High-Risk Technologies, p5, Basic Books, New York 1984.

Pourret, O., Naïm, P. & Marcot, B. (2008) "Bayesian Networks. A practical guide to applications". John Wiley & Sons. England.

Pyy, P. (2001). An analysis of maintenance failures at a nuclear power plant. Reliability Engineering and System Safety 72, 293–302.

Reason, J (1990). Human error. Cambridge University Press.

Reason, J. (1995). A System Approach to Organizational Error. Ergonomics, Vol. 38, pp. 1708–1721 (doi:10.1080/00140139508925221).

Reason, J. (1997). *Managing the Risks of Organizational Accidents*, Ashgate Publishing Company, Brookfield, VT.

Reason, J. (2000, March 18). Human error: models and management. British Medical Journal, Vol. 320(7237): pp. 768–770.

Santander Meteorology Group (2013). Data mining and artificial intelligence: Bayesian and Neural networks (http://www.meteo.unican.es/research/datamining).

Schulman, P.R. (1993, November). The negotiated order of organizational reliability, Administration & Society, Vol. 25, No. 3, p. 353.

Šimić, Z., Zerger, B., & Banov, R. (2015). Development and first application of an operating events ranking, Nuclear Engineering and Design, Vol. 282, pp. 36–43.

Sutcliffe, K.M. & Vogus, T.J. (2003). Organizing for Resilience. In Cameron, K., Dutton, J.E., & Quinn, R.E. (Eds.), Positive Organizational Scholarship. San Francisco: Berrett-Koehler. Chapter 7 pp. 94-110.

Swain, D.A. (1987). *Accident Sequence Evaluation Program*, NUREG/CR-4772, Sandia Laboratories.

Swain, D.A., & Guttman, H.E. (1983). *Handbook for Human Reliability with Emphasis on Nuclear Power Plant,* NUREG/CR-1278, Sandia Laboratories.

Trepess, D. (2003). *A Classification Model for Human Error in Collaborative Systems*, School of Computing, Staffordshire University, UK.

Wegner, M.S. (1999). Survey of SCRAMS during Maintenance and Surveillance Activities (Docket ML993350496).

Weick, K., Sutcliffe, K., & Obstfeld, D. (2008). "Organizing for high reliability: Processes of Collective Mindfulness", Crisis Management, Vol. 3, pp. 81–123.

Woods, D., & Wreathall, J. (2008). Stress-Strain Plots as a Basis for Assessing System Resilience, Resilience Engineering Perspectives, Vol. 1, Ashgate Publishing.

Woods, D., Chan, Y. & Wreathall, J. (2013, June 24-27). The Stress ‐ Strain Model of Resilience Operationalizes the Four Cornerstones of Resilience Engineering, 5th REA Symposium: Managing Trade-Offs, Soesterberg, Netherlands.

Zadeh, L.A. (1965). “Fuzzy Sets,” Information and Control, Vol. 8, pp. 338-353.

Articles Written

- P.F. Nelson, T. Ruiz-Sánchez, C.R. Grantom, C. Martín del Campo, "*HRA pre-initiator analysis gone statistic",* ANS PSA 2013 International Topical Meeting on Probabilistic Safety Assessment and Analysis Columbia, SC, September 22-26, 2013.
- Nelson, P., & Martín-del-Campo, C., *Evaluation of Barriers and Resilience to Improve Organizational Performance in Nuclear Power Plants*, Proceedings from Annual Mexican Nuclear Society conference, Boca del Río, Veracruz, September 2014.
- P.F. Nelson, T. Ruiz-Sánchez, C. Martín del Campo, "*Use of Corrective Action Programs at Nuclear Plants for Knowledge Management",* Probabilistic Safety Assessment and Management , PSAM 12 2014, Honolulu, HI, June 22-27, 2014.
- P.F. Nelson, C. Martín del Campo, "Determining Resilience Thresholds For Nuclear Power Plants*", International Topical meeting on Probabilistic Safety Assessment and Analysis*, PSA 2015, Sun Valley, Idaho, April 26-30, 2015.
- Pamela F. Nelson, Cecilia Martin del Campo, Bruce Hallbert, Ali Mosleh, "Development of a Leading Performance Indicator from Operational Experience and Resilience in a Nuclear Power Plant", *Nuclear Engineering and Technology*, Elsevier, accepted 1 octubre, 2015. http://dx.doi.org/10.1016/j.net.2015.10.010.