A Bayesian Network Approach to Early Reliability Assessment of Complex Systems

by

Petek Yontay

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2016 by the
Graduate Supervisory Committee:

Rong Pan, Chair
Douglas C. Montgomery
Dan L. Shunk
Xiaoping Du

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

Bayesian networks are powerful tools in system reliability assessment due to their flexibility in modeling the reliability structure of complex systems. This dissertation develops Bayesian network models for system reliability analysis through the use of Bayesian inference techniques.

Bayesian networks generalize fault trees by allowing components and subsystems to be related by conditional probabilities instead of deterministic relationships; thus, they provide analytical advantages to the situation when the failure structure is not well understood, especially during the product design stage. In order to tackle this problem, one needs to utilize auxiliary information such as the reliability information from similar products and domain expertise. For this purpose, a Bayesian network approach is proposed to incorporate data from functional analysis and parent products. The functions with low reliability and their impact on other functions in the network are identified, so that design changes can be suggested for system reliability improvement.

A complex system does not necessarily have all components being monitored at the same time, causing another challenge in the reliability assessment problem. Sometimes there are a limited number of sensors deployed in the system to monitor the states of some components or subsystems, but not all of them. Data simultaneously collected from multiple sensors on the same system are analyzed using a Bayesian network approach, and the conditional probabilities of the network are estimated by combining failure information and expert opinions at both system and component levels. Several data scenarios with discrete, continuous and hybrid data (both discrete and continuous data) are analyzed. Posterior distributions of the reliability parameters of the system and components are assessed using simultaneous data.

Finally, a Bayesian framework is proposed to incorporate different sources of prior information and reconcile these different sources, including expert opinions and component information, in order to form a prior distribution for the system. Incorporating expert opinion in the form of pseudo-observations substantially simplifies statistical modeling, as opposed to the pooling techniques and supra Bayesian methods used for combining prior distributions in the literature. The methods proposed are demonstrated with several case studies.

DEDICATION

To my parents and Mr. Goksan Aytekin who believed in me and my dreams...

ACKNOWLEDGMENTS

I want to thank my committee members for serving in my dissertation committee, sharing their expertise with me during my studies and for their valuable recommendations. I am so grateful to have worked with Dr. Pan during my PhD. He provided a supporting environment at work and has been a great mentor. His vision, knowledge, and encouragement always inspired me and helped me stay focused and pursue my research. Dr. Douglas C. Montgomery is also highly deserving of my gratitude for his extensive support during my PhD and for being on my side from the beginning. Being his Teaching Assistant was one of the greatest opportunities in my life. Also, I cannot deny the influence of the snacks he brought to my desk during long study nights. I would like to thank Dr. Dan L. Shunk and Dr. Xiaoping Du for partaking their time and effort as my committee members to help me fulfill the degree requirements.

I would like to also recognize Dr. Murat Caner Testik, who was my advisor during my master studies in Turkey. He was a great mentor to me, and he was my inspiration for an academic career. He taught me all about quality and reliability, and encouraged me to pursue my PhD in the US. Working with him was one of the milestones of my life, and I will be always so thankful for his support.

I have been so lucky to have so many sources of support during my education life. I am so grateful for my family, who supported my dreams and allowed me to follow my own path. They were always there for me, during my worst times, even just on the other side of the phone line. They believed in me, even when I did not and was about to give up. They always encouraged me to keep going and finish what I have started. I also would like to say special thanks to Mr. Goksan Aytekin, who has been a close family friend and has treated me like a daughter. Not only he contributed to my education at school, but he also taught me all about music, arts, and literature. He has been a great milestone for my personal development.

Special recognition is also due to the people whom I call friends and whom stayed with me during this stressful era of PhD. They have been a great family for me here, when my family was thousands of miles away. Among these special friends are Siddhartha Sampath, who has been the most positive and supportive person and has brought the best in me as a friend; Dr. Nicholas Mason, who would drop everything and come for help anytime; Mickey Mancenido, who has been my “power-pact” and enlightened me with her wisdom; Gina Dumkrieger, who has been great company during our short coffee breaks; Dr. Edgar Hassler, who was of great help when I got stuck with my research; and Derya Kilinc, who flattered me with her “hang in there” gifts when I was so stressed out with my research. Last but not least, Ayca Koseoglu and Ayse Bike Baykara were my greatest support and helped me stay strong and proud even during hard times.

There are many people who helped me get to where I am now. Unfortunately, I have to fail to mention all of the names, but I still would like to thank them all with my heart for touching my life and helping a small-town girl get the biggest achievement in her life.

Last, I gratefully acknowledge the financial support I received from the National Science Foundation through grant CMMI-1301075.

TABLE OF CONTENTS

Page

LIST OF TABLES x

LIST OF FIGURES xii

CHAPTER

1 INTRODUCTION 1

1.1 Problem Statement 1

1.2 Motivation 3

1.3 Overview of Dissertation 6

2 BAYESIAN NETWORKS FOR RELIABILITY PREDICTION IN FUNCTIONAL DESIGN STAGE 9

2.1 Introduction 9

2.2 Literature Review 11

2.3 Background and Framework 13

2.3.1 Conceptual Design 13

2.3.2 Functional Analysis 14

2.3.3 Bayesian Networks 15

2.4 Methodology 17

2.4.1 BNs from Functional Analysis 17

2.4.2 Conditional Probability Table 18

2.4.3 With Complete Function Log Data 18

2.4.4 With Function Failure Records 20

2.5 A Case Study 21

2.6 Conclusion 24

CHAPTER Page

3 A COMPUTATIONAL BAYESIAN APPROACH TO DEPENDENCY ASSESSMENT IN SYSTEM RELIABILITY 26

3.1 Introduction 26

3.2 Background 29

3.2.1 Models for Multilevel System Reliability Assessment 29

3.2.2 Computational Methods in Bayesian Inference 32

3.3 Methodology 33

3.3.1 Bayesian Networks 33

3.3.2 Bayesian Parameter Estimation in Bayesian Networks ... 36

3.3.2.1 Complete System Log Data 36

3.3.2.2 Summarized System Failure Data 43

3.3.2.3 Incomplete Lower-Level Data 45

3.4 Case Study 53

3.4.1 Computational Complexity 60

3.5 Conclusion and Future Research 62

4 A BAYESIAN APPROACH TO SYSTEM RELIABILITY ASSESSMENT WITH INCOMPLETE HETEROGENEOUS DATA 65

4.1 Introduction 65

4.2 Background 68

4.3 Methodology 73

4.3.1 Bayesian Networks in Reliability Assessment 73

4.3.2 Integrating Incomplete Lifetime Data Using Bayesian Inference 75

4.3.3 D-Separation in Bayesian Networks 81

4.3.4 Formulation of the Likelihood Function for Incomplete Lifetime Data . . . 83
4.3.5 Integrating Incomplete Hybrid Data Structures by Bayesian Inference . . . 90
4.4 Illustrative Example . . . 92
4.4.1 The Guided Missile System Structure . . . 92
4.4.2 Bayesian Network Model for the Guided Missile with Incomplete Data . . . 93
4.4.3 Reliability Assessment and Prediction . . . 100
4.5 Conclusion and Future Research . . . 101
5 A BAYESIAN FRAMEWORK FOR INCORPORATING DIFFERENT SOURCES OF PRIOR KNOWLEDGE IN RELIABILITY ASSESSMENT . . . 104
5.1 Introduction and Background . . . 104
5.1.1 Elicitation Techniques . . . 106
5.1.2 Verification and Validation of Experts . . . 111
5.1.3 Combination of Several Prior Distributions . . . 113
5.2 Methodology . . . 118
5.2.1 Incorporating Priors From Experts . . . 118
5.3 An Application to an Anti-Aircraft Missile System . . . 123
5.4 Conclusion . . . 127
6 CONCLUSIONS AND FUTURE WORK . . . 130
6.1 Summary of Methods and Contributions . . . 130
6.2 Suggestions for Future Research . . . 132

CHAPTER Page

REFERENCES 135

LIST OF TABLES

Table Page

1 Conditional Probability Table. 19

2 Function Failure Records.... 20

3 Marginal Probability Tables for the BN Example. 35

4 Conditional Probability Table (CPT) for the BN Example. 35

5 Marginal Probability Tables for the Multi-State BN Example 40

6 Conditional Probability Table (CPT) for the Multi-State BN Example.... 40

7 System Failure Records.... 43

8 Data from a System Failure Record.... 45

9 Empirical Mean, Standard Deviation, and Quantiles for Posterior Failure Probabilities.... 45

10 State Vectors of System in Figure 7.... 51

11 Possible State Vector Combinations of System in Figure 7. 52

12 Empirical Mean, Standard Deviation, and Quantiles for Parameters. 53

13 State Vectors of System in Figure 10. 58

14 Empirical Mean, Standard Deviation, and Quantiles for Parameters. 59

15 Missile Guidance System's Basic Component Reliability Characteristics. 95

16 Simulated Evidence Data. (Data with Superscript (P for Pass) Are Right-Censored Observations Where the Unit Worked at the Specific Time. Data with Superscript (F for Fail) Are Left-Censored Observations Where the Unit Has Failed.).... 98

17 Summary Statistics of the Posterior Samples for the Parameters.... 100

18 Summary Statistics for Reliability Assessment of the System. 101

19 Test Data. 123

Table Page

20 Expert Opinions for the Weapon System. 124

LIST OF FIGURES

Figure Page

1 A Simple BN from Functional Analysis 17

2 The Relationship between the Main Functions 22

3 Conditional Probability Tables 23

4 Evidence Analysis of Filter Function Failure 24

5 A Simple 2-Component System BN Example 35

6 Box Plots of Conditional Probabilities with Different prior Distributions 42

7 Basic Two Component BN System with Sensors on the System and Component 1. 50

8 The Fault Tree of an Active Vehicle Suspension. 54

9 The Corresponding BN Model of the Fault Tree Model in Figure 8. 55

10 The AVS Model with Sensors. 56

11 Prior (Dashed Lines) and Posterior (Solid Lines) Distributions of Conditional Probabilities. 60

12 A Sample Bayesian Network 74

13 BN Representation of a Hierarchical System. 78

14 Two Conditionally Independent Sub-Systems Given S_1. 83

15 Descriptive Flowchart of the Proposed Approach. 84

16 Reliability Block Diagram of a Simplified Missile Guidance System. 93

17 BN Representation of the Missile Guidance System. 94

18 Three Conditionally Independent Subsystems Given S_0, S_1 and C_5. 96

19 The Reliability Distribution of the New System with Respect to Mission Time. 101

20 Reliability Block Diagram for a Weapon System. 123

Figure Page

21 Posterior Distributions of the Reliability Parameters. 125

22 Posterior Reliability Distributions of the System and Components. The Solid Line Is the Posterior Mean and the Dashed Lines Are the 90% Credible Interval. 125

23 Posterior Distributions of the Experts' Weight Parameters. The Solid Line Is the Posterior Distribution for the First Expert and the Dashed Line Is the Posterior Distribution for the Second Expert. 126

24 Posterior Distributions of the Failure Rate Given Different Priors. 127

Chapter 1

INTRODUCTION

1.1 Problem Statement

Due to the increasing rate of introduction of new products in today's marketplace, it is becoming more and more important to satisfy the consumers' demands, which requires that the products be highly reliable. As the demand of reliability is strictly increasing, achieving high quality and reliability has become a default requirement during a product's life cycle. The complexity of engineered products have also increased vastly over the last decades; therefore, the need to develop efficient methods for reliability assessment and building tools to incorporate these methods into the product's life cycle is undeniable and a lot of researchers and engineers have worked on reliability assessment of complex systems to achieve good reliable products.

Therefore this dissertation aims to address the reliability assessment problem and presents Bayesian network approaches for two research problems: early system reliability during functional design stage and system reliability assessment with incomplete and overlapping data. A third problem this research addresses is how to combine prior information from diverse sources for reliability assessment in a Bayesian framework. Systems are getting more and more complex due to added functionalities; therefore, traditional methods like fault trees and reliability diagrams are not capable of analyzing these complex systems properly. As a result, there is a need for a method to model and analyze complex systems and evaluate the system reliability by incorporating information from different sources.

The first challenge for the research problems mentioned above is the lack of complete compatible system reliability information. A complex system is usually composed of sub-systems and components, structured in a hierarchy. In addition, information comes from multiple levels of the system in different forms. In most of the cases in real life, we do not have complete information coming from all levels of the system. We propose a Bayesian network methodology to incorporate available information into the system and component reliability assessment processes. Especially, during the design stage of a system, we may not have the detailed knowledge of all possible failure mechanisms of the system, and the scarcity and poor quality of reliability data during the design phase might be very problematic. In these cases, we would like to learn more about the interactions between components and how they work together and the effects of these interactions on system reliability.

Bayesian networks (BNs) have significant advantages over traditional reliability assessment methods due to their efficiency in evaluating associations and simplicity in providing a system assessment. They are very efficient at propagating the uncertainty and updating the system with new data in the network. They are also applicable when system structures are too complex to be represented by fault trees or reliability block diagrams. As fault trees and reliability block diagrams model the system's reliability-wise structure in a deterministic way, they are in general ill-suited for a conceptual design where even the components of system and their configurations have not been determined. Bayesian network, on the other hand, can model the uncertainties in various system functions and the generating processes of system functions, thus it is a viable tool for studying product reliability at its early design stages. Therefore, the purpose of this research is to gain reliability insight starting from early stages of the design of a new product using different sources of information using a Bayesian network framework.

1.2 Motivation

Reliability prediction at a product's very early design stages has been gaining attention over the last decade. Build-in-reliability (BIR) and design-for-reliability (DFR) philosophies have been a great influence on the necessity to estimate the reliability of a product during its conceptual design phase. However, predicting reliability during the conceptual design stage is challenging, as the available knowledge is very limited and it is descriptive and qualitative in nature.

Bayesian network models have been proved to be powerful tools that provide important advantages over traditional techniques in early reliability assessment. Traditional methods, such as fault trees or reliability block diagrams, do not show enough flexibility to capture the uncertainties in the dependencies among components and the system. Bayesian networks are modeled by conditional probabilities instead of deterministic "AND" and "OR" relationships, providing a probabilistic measure of dependencies between components and the system. They are especially useful during the early stage of product design process when we are not sure about the reliability structure of a complex system. When we use Bayesian inference techniques for parameter estimation, BNs provide a very efficient framework for combining information from multiple sources and multiple levels for system reliability assessment. As a result, we aim to use BN models and Bayesian inference together for dependency assessment in system reliability. BNs can effectively address the uncertainty in all stages of the product life cycle due to their probabilistic structure and they can solve complex problems due to advancements in simulation-based computing techniques, making them very favorable to work with.

Reliability assessment techniques in the early stages of the product development process have been studied extensively in the past few decades. Most of these approaches are centered on component-specific failures (Kurtoglu and Tumer, 2008; Stone and Wood, 2000; Derelöv, 2008). These studies mostly focused on the functional design stage, and they were descriptive and qualitative in nature. Sanchez and Pan (2011) provided statistical inference on the failure rate of a new design, emphasizing the value of reliability prediction at a product's very early design stage. However, their study also analyzed the failure causes of components individually. With the advent of highly complex systems that derive functionalities from multiple domains, more emphasis is required on identifying failures arising due to various interactions among components, which is largely absent in existing failure analysis approaches.

There are many mechanisms through which failures occur in any given system. One typical example of a complex failure mechanism is carburetor icing in internal combustion engines (ICE), which results due to the freezing of air moisture during the suction of highly humid air through the carburetor (Augustine *et al.*, 2012). An ICE has many components and these components all interact with each other. We need to understand how these interactions affect the working mechanism of the system in order to gain an understanding of the reliability structure. However, assessment of these interactions in early stages of product development is limited due to the general non-availability of hard numerical data and representative mathematical relationships. There exist very few techniques that support effective identification of failure mechanisms at the design stage and help generate an understanding of the early reliability.

Many advantages can be gained by beginning the reliability analysis of a new design at the conceptual design stage. The main advantage comes from arriving at a more reliable product without the need for multiple redesigns in order to eliminate failure modes in advanced stages of the design process, as happens in the traditional approaches such as FMEA. Reliability for any product or service is crucial. It becomes even more crucial for those complex systems that cannot fail, such as military weapon systems, aerospace systems, automotive systems and nuclear systems. For new products in these applications, reliability must be considered in the design phase to meet all the requirements given the high risks in case of failure.

In the early stages of product design, traditional reliability information is scarce. Many studies in literature assess complex system reliability with complete independent data. Therefore, it is of utmost importance to develop methods to incorporate available information to assess system reliability. There might be different sources of data that provide reliability information while designing a new product and these data might be available from different components or different levels of the system, as complex systems are usually structured in hierarchical levels. As another example, we can think of a contaminant reduction device (CRD) used in automotive industry (Sanchez, 2014; Yontay *et al.*, 2015). If we would like to propose improvements on the existing design to comply with some regulations, we will have to evaluate several design options. Since the development of the CRD is in the conceptual design phase, the data for the new model is scarce. In this scenario, using a Bayesian network to create a graphical model of the design parameters (functions of the system) and combine whatever information is available from the previous designs (parent products) is crucial in assessing the early reliability of the device because of the uncertainty involved in the design. We can then compare different design options using the early reliability analysis using a Bayesian network framework and choose the best design.

Motivated by the above-mentioned facts, this dissertation presents Bayesian network methods for system reliability assessment of complex systems. The main motivation for this research is to address the gap in the area of addressing the dependencies in a system using incomplete and simultaneous data due to the fact that recent research on BNs has mostly focused on using complete and independent data for system reliability assessment.

1.3 Overview of Dissertation

The remainder of this dissertation is organized as follows. Chapter 2 proposes a Bayesian network approach to incorporate data from functional analysis and parent products in order to analyze the relationships among the functions of a system during design stage. Chapter 3 and Chapter 4 look into the system in more detail and these chapters are devoted to learning the parameters of a Bayesian network with incomplete simultaneous data. Chapter 5 focuses on incorporating different sources of prior information using a Bayesian model. Finally, in Chapter 6, we summarize the contributions of our research and discuss further research directions in this area.

More specifically, in Chapter 2, we focus on the concept of integrating the product design information from functional analysis with the product failure information derived from other sources. A product failure is defined as when one or more of its designed functions cannot be executed as expected. Failure modes can be stated in terms of deviation of functions. Thus, we use functional analysis to reveal a preliminary reliability structure for the product and to create a BN. The nodes of BN are the designed functions and their corresponding failure modes. The conditional dependencies among these nodes are extracted from engineering experience, expert opinions, and the failure data from historical failure occurrence of the same function in similar (parent) products.

In Chapter 3, we present a Bayesian network approach for evaluating the conditional probability of failure within a complex system, using a multilevel system configuration. The novel feature of this model is that Bayesian network (BN) is used to represent the probabilistic relationship between system and component reliability, which is a generalization of the deterministic relationship usually modeled by fault trees and reliability block diagrams. The model allows incorporating simultaneous discrete data coming from several sensors in the system and can provide an initial analysis of the dependency structure in system reliability especially when the failure structure is not well known. The methodology is illustrated with three different scenarios, each scenario demonstrating our Bayesian methodology by using data coming from different system levels.

In Chapter 4, we extend the main ideas in Chapter 3 to the incomplete and continuous failure time data, in which case the Bayesian inference becomes much more challenging. In this case, we propose a Bayesian network approach for assessing the time-to-failure distribution parameters of the components and for predicting early reliability of the system and components over time. Our model allows us to incorporate incomplete and simultaneous life time data from several sensors in the system and it is applicable to any lifetime distribution. We also extend the case to the hybrid data structures, where we have both discrete and continuous data. We illustrate the methodology through a demonstrative example.

Chapter 5 is devoted on combining multiple sources of prior information for the system. The aim of our research in this chapter is to obtain prior data from the system and components, in addition to using expert opinion effectively and combining these different streams of information to derive prior distributions for the parameters of the Bayesian model. Specifying prior distributions in a Bayesian network is an important part of the modeling process. We plan to develop a method that allows us to incorporate non-observed, subjective and legacy information, such as expert opinions, historical data and specifications from similar products, into the model efficiently.

Chapter 2

BAYESIAN NETWORKS FOR RELIABILITY PREDICTION IN FUNCTIONAL DESIGN STAGE

2.1 Introduction

Reliability prediction at a product's very early design stages has been gaining attention over the last decades. Build-in-reliability (BIR) and design-for-reliability (DFR) philosophies have been a great influence on the necessity to estimate the reliability of a product during its conceptual design phase. However, predicting reliability during the conceptual design stage is challenging, as the available knowledge is very limited and it is descriptive and qualitative in nature.

Probabilistic methods for the system reliability assessment of a product design have been used extensively by reliability engineers. These modeling techniques mostly utilize measures like mean time to failure (MTTF), failure rate and failure distributions obtained by some life tests conducted in the detailed design stage. However, reliability should be incorporated into the product life cycle as early as possible and maintained throughout the cycle to ensure good quality of a product (Pahl and Beitz, 2013). The acceleration of product development speed and the reduction of product's life cycle cost are the major benefits that can be gained by beginning the failure analysis of a new product at its conceptual design stage, in particular during its functional analysis. Traditional approaches like FMEA, fault trees and reliability block diagrams (RBD) could only be implemented after a detailed design of the product has been carried out; therefore, they are not well suited for product reliability predict at early design stages.

Bayesian networks (BNs) have significant advantages over traditional reliability assessment methods due to their efficiency in evaluating associations and simplicity in providing a system assessment. They are very efficient at propagating the uncertainty and updating the system with new data in the network. They are also applicable when system structures are too complex to be represented by fault trees or reliability block diagrams. As fault trees and reliability block diagrams model the system's reliability-wise structure in a deterministic way, they are in general ill-suited for a conceptual design where even the components of system and their configurations have not been determined. Bayesian network, on the other hand, can model the uncertainties in various system functions and the generating processes of system functions, thus it is a viable tool for studying product reliability at its early design stages.

This chapter focuses on the concept of integrating the product design information from functional analysis with the product failure information derived from other sources. A product failure is defined as when one or more of its designed functions cannot be executed as expected. Failure modes can be stated in terms of deviation of functions. Thus, we use functional analysis to reveal a preliminary reliability structure for the product and to create a BN. A typical BN model consists of two parts: a direct acyclic graph (DAG) modeling presentation and conditional probability tables between parent and child nodes. The nodes of BN are the designed functions and their corresponding failure modes. The conditional dependencies among these nodes can be extracted from engineering experience, expert opinions, and the failure data from historical failure occurrence of the same function in similar (parent) products. The chapter is organized as follows: A literature review is provided in

Section 2.2. Section 2.3 presents specific descriptions of the framework introducing functional analysis and Bayesian networks. Our proposed methodology is described in Section 2.4, followed by a case study in Section 2.5. Finally, Section 2.6 draws the conclusion.

2.2 Literature Review

System reliability can be defined as the probability that a system will perform its intended function for a specified period of time under stated conditions. Analytical methods, with the assistance of graphical tools such as fault trees, reliability block diagrams and network graphs, are frequently used to estimate system reliability.

In literature, the idea of using BNs for system reliability assessment was discussed by several studies (Langseth and Portinale, 2007; Wilson and Huzurbazar, 2007; Mahadevan *et al.*, 2001). Mahadevan *et al.* (2001) proposed the methodologies of applying BNs to structural system reliability assessment with multiple failure sequences. Bobbio *et al.* (2001) and Boudali and Dugan (2006) also proposed BNs as the alternatives to traditional reliability estimation approaches. Doguc and Ramirez-Marquez (2009) presented a holistic method for constructing a BN model for estimating system reliability. They introduced a method that uses historical data and provided efficient techniques for construction of the BN model.

The aforementioned studies were conducted at existing products with the availability of product failure data. There are very few studies implementing reliability-based design at the very early product design stage. Clark and Paasch (1996) described a diagnostic modeling methodology in the conceptual design phase. Their method was based on the relationship between a system's functions and the failure modes of components. Eubanks *et al.* (1997) proposed a method to address reliability during the early stages of design. They utilized behavior modeling to identify failures with the help of function-structure relationships and then analyzed the effects of these failures. Derelöv (2008) proposed a qualitative framework of potential failure identification in a conceptual design. He modeled the system in a qualitative and deterministic way. Huang and Jin (2008) addressed the gap between reliability and design, and developed a conceptual strength interference theory by parameterizing the conceptual design space via introducing reliability-related parameters into functional design. Due to the lack of direct reliability information in the early design stage, some unconventional sources of reliability information need to be explored. Even so, how to integrate objective and subjective information from disparate sources in a systematic way is still a challenging task. Sanchez and Pan (2011) presented an enhanced parenting process for predicting reliability of a new product by using the reliability information of parent products. They relied on expert elicitation for assessing the effects on design changes on individual failure causes.

Product functional analysis is a critical step in the product conceptual design. Qian and Gero (1996) presented an approach of using the associations between function, behavior and structure to build a formal structure. Stone and Wood (2000) introduced a consistent design language, called a functional basis, in which they provided clear definitions for each function and flow. Otto and Wood (1998) discussed various techniques in product design and development that address conceptual formulation, and functional design issues. Hirtz *et al.* (2002) provided a set of function bases in order to standardize and formalize function structure design, modeling and evaluation. Sridharan and Campbell (2005) presented an approach to developing the graph grammar for function structures. In addition, Chandrasekaran *et al.* (1993) used functional representation (FR) to define the design space, describing the overall function first, and then the behavior of each component with respect to that function. They presented FR as a good framework for capturing the casual components in performing the product's functions. Wang and Jin (2002) proposed an analytical approach to functional design by introducing a new concept, called function-behavior, and developing a BN based analysis method. The function-failure design method, developed by Tumer and Stone (2001), relates failure modes to product functions. It can be utilized for the conceptual design of new products or the redesign of existing products.

In general, the existing methods are largely qualitative and the function-failure relationships are often represented by a matrix, which is inadequate for modeling failure-cause dynamics and for representing the intricate connections among multiple functional failure modes and their causes. We propose a methodology of transferring functional analysis to BNs such that the quantitative analysis of a new product's reliability could be performed even at its early design stage.

2.3 Background and Framework

2.3.1 Conceptual Design

Conceptual design is the first phase of design, providing a description of the proposed system through a set of concepts about its functionalities. A conceptual design utilizes concept and function structure formulations corresponding to functional requirements for the product. It does not address the detailed information about physical components.

The tasks of conceptual design are defined differently in various sources in literature. But according to the definition of the design process by Pahl and Beitz (2013), the stages of a conceptual design are:

- Identify customer requirements.
- Decompose the customer requirements into design requirements.
- Establish functional structures.
- Generate candidate conceptual design solutions.
- Evaluate the design concepts and the functional structures for the detailed design stage.

As a result, the conceptual design phase generates the concepts that will be implemented during the next stages of the product design. Function structures are used during conceptual design to transform the customer requirements into specific functional tasks.

2.3.2 Functional Analysis

Functional design is an important step in the product design process. The lack of analysis for functional design is a factor that can cause inefficient and unreliable designs. The problems might not be detected until the embodiment design, which might be costly and time consuming.

In early stages of design, system failures are identified as failure to achieve one or more predefined functions, and a functional model of a system is simply a graphical representation of the system functionality, without any details of the structure (Otto and Wood, 1998). In the initial stages of design, based on the customer requirements, an overall function for the design can be identified, which includes the flows of energy, material and signal of the function. This overall function is then broken down into sub-functions with less complexity but more details (Tumer and Stone, 2001). In order to effectively represent functions and sub-functions, a standardized modeling language is required. Various studies have been conducted on a generic functional basis for functional modeling (Hirtz *et al.*, 2002).

Failure of a system is defined as the termination of the ability of the system to complete its intended function. Thus, a system failure mode can be correlated to functions of components. If this correlation can be established, then failure modes can be eliminated or significantly reduced by improving component quality or reconfiguring system reliability structure. This is the fundamental logic behind the system reliability improvement using FTA or FMEA. Following the same logic, in functional analysis a function failure is caused by the interruption of material, energy and signal flows. Different design concepts may cause different types of interruption, which are the failure causes that designer should be aware of. When a new product is being designed, its intended functionalities will be matched to the functionalities of existing products, so the designer can generate several design options to materialize the intended function. Therefore, the new design will inherit the failure mode from its parent products and the reliability prediction for the intended function will become possible by combining the failure information from parent products and from expert opinions.

2.3.3 Bayesian Networks

A Bayesian network (BN) consists of two main parts:

- Qualitative part: consists of a directed acyclic graph (DAG) where the nodes represent random variables (continuous or discrete) and directed arcs representing causal relationships between the random variables.

- Quantitative part: conditional probability tables between parent and child nodes.

In a BN, the nodes without any arrows directed into them are called root nodes and they are described according to their marginal probability distributions. The nodes that have arrows directed into them are called child nodes and the nodes that have arrows directed from them are called parent nodes. Each child has a conditional probability table associated with it, given the values of parent nodes.

Consider a BN over variables $X_1, X_2, \ldots, X_n$. By the chain rule of probability theory, the joint probability $P(X_1, X_2, \ldots, X_n)$ is

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid pa(X_i)) \tag{2.1}$$

where $pa(X_i)$ is the set of parents of node X_i.

Certain nodes in a BN may become uncorrelated if there is no link between these nodes. This situation is called conditional independence. These conditional independences allow us to decrease the number of terms in the chain rule, providing a simpler structure.

BNs can be utilized to model function structures where the nodes are represented by the designed functions and their corresponding failure modes. The conditional dependencies among these nodes can be extracted from engineering experience, expert opinions, and the failure data from historical failure occurrence from existing products with similar functions.

2.4 Methodology

2.4.1 BNs from Functional Analysis

A product's functions are typically determined based on customer requirements, as well as marketing analysis. A rigorous functional analysis provides the possible failure causes from material, energy and signal aspects that can be utilized for constructing Bayesian networks. It needs to be emphasized that in early design stages it is the product function, instead of component, to be analyzed, as individual components will only be materialized in a later design stage. For example, assuming that a functional failure is caused by four possible direct causes as shown in Figure 1, a designer can select different design options (e.g., choosing different function generating mechanisms or different components) during the embodiment design process so that some failure causes can be avoided. In functional analysis, all possible causes for a functional failure are elicited. They are, in general, structured hierarchically, extended to multiple levels. For simplicity, Figure 1 only shows one level, i.e., the direct causes to a functional failure.

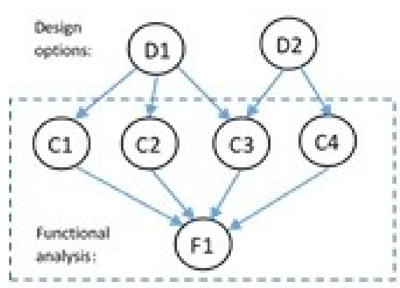

Figure 1. A simple BN from functional analysis

2.4.2 Conditional Probability Table

In a discrete BN each node may take values from several states. For example, the function node in Figure 1, F1, have two states, 1 or 0, corresponding to failure or success. However, a node, in general, can have more than two states. For example, let the direct parent node, C1, in Figure 1 be the material strength, then its states can be assigned as Strong, Medium or Weak. Conditional Probability Tables (CPTs) are needed to quantify the probabilistic relationships between nodes; i.e., to specify $Pr(F1 \mid C1, C2, C3, C4)$ in Figure 1. This is not addressed in a typical functional analysis. We propose two approaches to the quantification problem by integrating available data about a function and subjective assessment from experts.

2.4.3 With Complete Function Log Data

Consider a single function and two direct causes that govern the successful execution of this function. Given two states to each cause and the CPT as shown in Table 1, there are four parameters, p_1, p_2, p_3, p_4, that need to be specified. Although it is uncommon in practice, we start our discussion with this naïve scenario – a complete historical dataset of the states of the function and its direct causes is available. This is possible if this function and its associated causes are continuously monitored by sensors and the log data from existing products that perform the same function can be obtained.

Using all observed instances of function states and cause states it is straightforward to obtain the estimation of the conditional failure probability given a combination of cause states. For the previous example,

$$p_i = Pr(F = 1 \mid C1_i, C2_i) = \frac{\sum_k I_k(F = 1, C1_i, C2_i)}{\sum_k I_k(C1_i, C2_i)} \tag{2.2}$$

where the denominator is the total number of instances of the specific combination of C1 and C2 and the numerator is the number of instances of function failure at this combination.

Table 1. Conditional probability table.

$C1$	$C2$	$Pr(F = 1 \mid C1, C2)$	$Pr(F = 0 \mid C1, C2)$
0	0	p_1	$1 - p_1$
0	1	p_2	$1 - p_2$
1	0	p_3	$1 - p_3$
1	1	p_4	$1 - p_4$

However, even this simple formula could become troublesome in practice when there are many states for each cause node. In such case, the number of combinations grows large, thus the log file could be highly fragmented. There might be no observation for a particular combination. Therefore, it is better to combine Eq. 2.2 with the expert's opinion on how many function failure may happen for a given parent nodes combination. This is equivalent to assign a prior distribution to the function failure probability. Assume a Beta prior distribution, $Beta(a_i, b_i)$, for p_i, then the posterior estimation of p_i is given by

$$p_i = \frac{\sum_k I_k(F = 1, C1_i, C2_i) + a_i}{\sum_k I_k(C1_i, C2_i) + (a_i + b_i)} \tag{2.3}$$

where $(a_i + b_i)$ is the equivalent sample size in the prior and a_i is the equivalent number of failures in prior samples.

Therefore, in the expert opinion elicitation process, two questions would be asked: In your experience, how frequent this type of combination of C1 and C2 may happen? And, in your experience, what is the chance of function failure given this type of combination of C1 and C2? The prior parameters, a_i and b_i, can be derived from the answers of these questions. By combining expert assessments and historical data, a robust conditional failure probability can be obtained.

2.4.4 With Function Failure Records

A function failure record is often maintained within an organization and it is the most common type of information that one can track for assessing the function failure probability. For example, given a checklist such as Table 2, we can see that function F1 failed once due to C1 and C2, and F2 failed once due to C3, etc.

Table 2. Function failure records.

Function	$C1$	$C2$	$C3$
$F1$	X	X	
$F2$			X
$\ldots$	$\ldots$	$\ldots$	$\ldots$

Notice that given these records, we can estimate the probability of failure causes given a failure mode; i.e., $Pr(C1, C2 \mid F1)$, but not the probability of a failure mode given failure causes. This is because, unlike the log data, Table 2 records only failure events. To obtain the conditional probability of failure given causes, we need to have the probability of occurrence of cause combination and the marginal probability of failure, because

$$Pr(F \mid C1, C2) = \frac{Pr(C1, C2 \mid F)Pr(F)}{Pr(C1, C2)} \tag{2.4}$$

Expert opinions on these marginal probabilities ($Pr(F)$ and $Pr(C1, C2)$) can be solicited. Experts are asked what the chance of a function failure is during the product's lifetime and what the chance of a cause state combination is. This can be obtained by directly estimating the occurrence rate of these events, then converting them to event probabilities based on exponential distribution. That is,

$$Pr(F) = 1 - e^{-\lambda_F t} \tag{2.5}$$

$$Pr(C1, C2) = 1 - e^{-\lambda_{C1,C2} t} \tag{2.6}$$

where λ_F and $\lambda_{C1,C2}$ are the occurrence rate of function failure event and cause combination event, respectively, and t is the product lifetime.

This approach is an extension of the parenting process presented by Sanchez and Pan (2011), in which only the probability of one failure mode given one failure cause was discussed. That is, they assumed that the effects of failure causes are independent to each other. Here, we generalize it to a general case without independence assumption.

2.5 A Case Study

A new contaminant reduction device (CRD) is being introduced for use in an automotive industry. A CRD is used to convert toxic exhaust emissions into less-toxic substances. A chemical reaction is stimulated through the exhaust flow and then contaminants are reduced in the system before the gas is released. Using functional analysis approach, some function structures are analyzed for the system. The functions to be represented by the Bayesian network are listed as follows:

- Flow of exhaust gas
- Injection of fluid
- Chemical reaction of catalysis
- Amount of contaminants

- Back pressure at outlet
- Filtering of the substances

It is assumed that the new CRD maintains the same failure structure as the previous designs; hence, information from the previous CRD products can be used to form the functional relationships. The function failure record from the parent products is analyzed and is combined with expert elicitation.

Our aim is to assess the product's reliability at the conceptual design stage. We use the methodology presented in this research to create a graphical model for capturing the relationships between the main functions of the system. The basic functional structure of the product is shown in Figure 2.

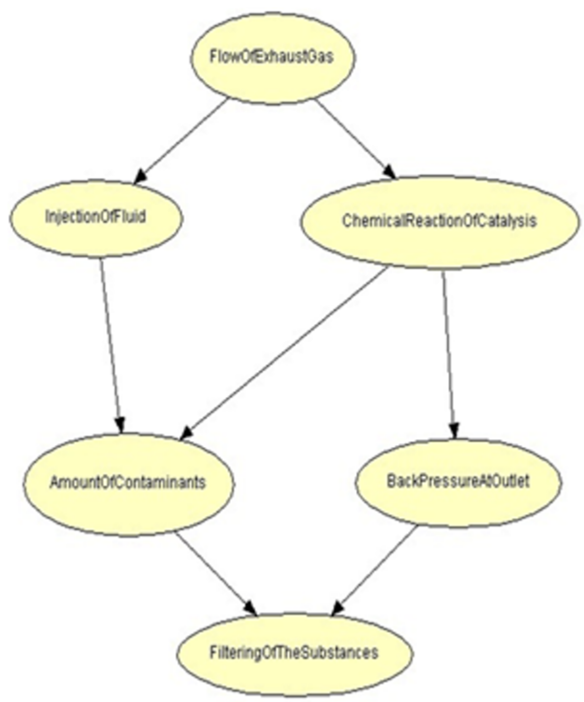

Figure 2. The relationship between the main functions

The next step is to obtain the conditional probability table for each node. The states of each node are expressed in binary variables: "1" for function failure and

"0" for function performing properly. Figure 3 shows the conditional probability tables for each node obtained using the failure records from parent products and then eliciting expert opinions to calculate the new failure rates for each function.

AmountOfContaminants

ChemicalReacti	1		0	
InjectionOfFlui	1	0	1	0
1	0.98	0.88	0.352	0.077
0	0.02	0.12	0.648	0.923

BackPressureAtOutlet

ChemicalReacti	1	0
1	0.99	0.141
0	0.01	0.859

ChemicalReactionOfCatalysis

FlowOfExhaust	1	0
1	0.995	0.018
0	0.005	0.982

FilteringOfTheSubstances

BackPressureA	1		0	
AmountOfCont	1	0	1	0
1	0.966	0.761	0.964	0.04
0	0.034	0.239	0.036	0.96

FlowOfExhaustGas

1	0.056
0	0.944

InjectionOfFluid

FlowOfExhaust	1	0
1	0.961	0.01
0	0.039	0.99

Figure 3. Conditional probability tables

Hugin Lite 8.0 was used for propagating the information through the network. The initial analysis of marginal distribution for each node shows that the key function, filtering of substances, is functional only 73% of the time. In order to find the probability distributions given that the filtering is not functioning, the evidence was propagated using the software and the back pressure node was found to be the function highly associated with the filter malfunction.

Figure 4 shows the impact of the state of back pressure on the distribution of filtering. It is obvious from the figure that it is very important that the back pressure at outlet must function properly so that the filtering could function properly too.

Figure 4. Evidence analysis of filter function failure

In this case study, the Bayesian network approach is able to provide the design team the information about which function parameters needed to be improved to meet the design specifications. Furthermore, sensitivity analysis is utilized for assisting an objective decision making process. As a result, the changes in the design are justified as they provide a more robust CRD.

2.6 Conclusion

In this chapter we propose to model system reliability using Bayesian network at the system's early design stage. The key idea is to utilize the reliability information of parent products that was stored as a function failure record. The relationships between failure modes and failure causes can be found from these historical records. Expert elicitation is also used in order to account for the changes from the parent products. Integrating both objective and subjective reliability information, we provide insights for the early reliability prediction problem. In our approach, the first step is the functional analysis of the system. It is necessary to identify and establish the relationships between the functions and a BN is constructed. Using belief propagation, the designer is able to evaluate the impact of different design scenarios on the system reliability of a conceptual design.

Chapter 3

A COMPUTATIONAL BAYESIAN APPROACH TO DEPENDENCY ASSESSMENT IN SYSTEM RELIABILITY

3.1 Introduction

Due to increasing demands of product functionality, engineered products have become more and more complex over time. The traditional reliability assessment methods for simple systems are often inadequate in analyzing more complex systems. Conducting full system tests is often too expensive to be implemented on such systems. This situation calls for a method to develop reliability models for complex systems and to integrate all available information for predicting system reliability.

There are situations that we do not have complete information of how a complex system would fail in its operating environment. We would like to learn more about the interaction between the system and its components and how they work together. In this chapter, we use Bayesian network (BN) to represent the probabilistic relationship between system and component reliability, which is a natural extension of the deterministic relationship typically modeled by block diagrams or fault trees when the failure structure is well understood.

The BN model has been proved to be a powerful tool that provides important methodological advantages over traditional techniques in reliability assessment. Traditional methods, such as fault tree or reliability block diagram, are still common representation in system reliability analysis; however, they are not flexible enough to capture the uncertainties in the dependencies among component, subsystem, and system (see Bobbio *et al.* (2001); Mahadevan *et al.* (2001); Boudali and Dugan (2006); Langseth and Portinale (2007); Wilson and Huzurbazar (2007)). BNs generalize fault trees by allowing components and subsystems to be related by conditional probabilities instead of deterministic "AND" and "OR" relationships; thus, they provide analytical advantages to the situation when we are not sure about the reliability structure of a complex system, especially during the early stage of product design process. Another important advantage of BN over the traditional approach is its ability of combining information from multiple sources at multiple levels for system reliability prediction, especially when the BN model is coupled with statistical Bayesian inference techniques. As a result, it is worthwhile to explore the use of BN model and Bayesian inference together for the dependency assessment in system reliability.

A BN model requires conditional probabilities to model the dependencies among components, subsystems, and systems. These conditional probabilities are capable of representing complex, probabilistic failure relationships in a multilevel system configuration. In a complex system, the failure relationship between system and component could be significantly more complicated than a typical series or parallel system, especially when the specific failure cause and failure mechanism has yet been understood, such as in a newly developed system (Sanchez and Pan, 2011). Therefore, investigating the conditional probability table of BN model can help engineers to sort out the unknown influential factors, if there are any.

The conditional probabilities in a BN model can be estimated by combining information from different sources. There are objective information sources, such as failures of older generation products, life test of component, and available field data, and there are subjective sources too, such as expert opinions. These data come with different types and different structures, causing difficulties in the estimation of conditional probability. Furthermore, a system evolves over time, so assigning fixed values to these probabilities limits the flexibility to account for the evolution process of system development. Therefore, we choose Bayesian inference for parameter estimation in the BN model. Bayesian inference is a statistical inference method that enables model parameter estimation by deriving the posterior distribution from a combination of prior distribution and likelihood function. It allows us to integrate both the prior information of model parameter and the data coming from different sources for model inference; therefore, we can obtain more precise estimation of BN model parameter.

The goal of this chapter is to develop the methodology of estimating conditional probabilities in a BN model using Bayesian inference so that the reliability-relevant information from different sources at different reliability structure levels of a complex system can be combined together. The next section presents a literature review of BN model and Bayesian inference. Our BN framework for system reliability and its inference method are discussed in Section 3.3. We start by discussing how to infer conditional probability using a conjugate model for a simple 2-state Bayesian network and then extend it to a multi-state model. We also briefly discuss the case where we have only system failure records. Finally, we develop a data analysis method for the scenario of having incomplete information from components. We illustrate the proposed method with a case study in Section 3.4 and conclude the chapter in Section 3.5.

3.2 Background

3.2.1 Models for Multilevel System Reliability Assessment

System reliability can be defined as the probability that a system will perform its intended function for a specified period of time under stated conditions. Analytical methods, with the assistance of graphical tools such as fault trees, reliability block diagrams and network graphs, are frequently used to estimate system reliability.

One of the primary goals in system design evaluation is to predict the reliability of the full system. A system is comprised of subsystems and components, or on functional wise, sub-functions and elementary functions, which can be represented by nodes in the system reliability topology. All nodes are potential source of failure. Consequently, reliability information may come from different levels of the system and it tends to be fragmented and heterogeneous. With data available at different system levels, the challenge becomes how to combine them to learn about the reliability of the system. The Bayesian method is very appealing for this challenging problem. Martz *et al.* (1988) and Martz and Wailer (1990) addressed the problem of integrating multilevel binary data from various levels of the system and expert guesses about the reliability of system components. These papers focused on series and parallel systems, whose component failure data were modeled using binomial distributions and beta distributions were used for the prior information at components, subsystem and system levels. They used approximations to provide a posterior distribution for system reliability. Several follow-up papers considered other computational Bayesian approaches to model inference and system reliability prediction. For example,Johnson *et al.* (2003) proposed a hierarchical Bayes model approach to system reliability prediction. Their approach utilized Markov chain Monte Carlo (MCMC) to infer model parameters, thus avoided analytical approximation. Hamada *et al.* (2004) applied the same approach on the non-overlapping, continuous failure time data of basic and higher-level failure events in a fault tree. Graves *et al.* (2007) further extended this line of research by considering multi-state fault trees. They used Dirichlet distribution to define the prior information about the probabilities of the states in the model. In addition, Graves *et al.* (2008) proposed a Bayesian approach to properly account for simultaneous multilevel data, i.e., use the simultaneous higher-level and partial lower-level data to determine the event of component failure. In a follow-up study, Reese *et al.* (2011) considered lifetime data throughout the system. They presented a Bayesian model that accommodates multiple lifetime information sources and provided a method to model the time evolution of a system's reliability. Wilson *et al.* (2006) proposed a methodology that allowed for the combination of different types of data at the component and system levels, and took a Bayesian approach to the estimation of reliability measure. Wilson *et al.* (2011) showed how to combine different types of reliability data with an example that had binomial data (modeled with a logistic regression) from the system and one component, lifetime data from another component, and degradation data from a third component. Guo (2011) discussed a unified Bayesian approach for simultaneously predicting system, subsystem, and component reliabilities when there are pass/fail, lifetime, degradation, or expert judgment data at any level of the system, which extended the work in Wilson *et al.* (2006). However, these studies were mostly based on fault trees and reliability block diagrams and did not cover the BN representation of system reliability.

In the system reliability literature, the idea of using BN model as the alternative to fault tree or block diagram for representing system reliability structure has been discussed by many authors (e.g., Bobbio *et al.* (2001); Mahadevan *et al.* (2001); Boudali and Dugan (2006); Langseth and Portinale (2007); Wilson and Huzurbazar (2007); Li *et al.* (2014)). However, previous studies do not address the problem of assessing reliability dependencies between system and its components. In this chapter, we will assess these dependencies using a computational Bayesian inference method; that is, given reliability information from multiple sources and at multiple levels of the system, we will provide the Bayesian estimation to the conditional probability parameter required in a BN model. The posterior distribution of conditional probability can be used to quantify of the variability of the dependency of system reliability to its components.

The aforementioned studies were conducted at existing products with the availability of product failure data. There are very few studies implementing reliability-based design at the very early product design stage. Furthermore, previous studies have not addressed the effect of simultaneous, yet incomplete, data, drawn from different system levels, on the BN model estimation. Since we aim to measure reliability dependencies within a system, datasets should be drawn simultaneously from the system and its components. Independent datasets will not be able to capture the dependencies within a system. However, getting simultaneous data from all components/subsystems may not always be possible due to lack of sensors or other observation limitations, especially during the design phase. Graves *et al.* (2008) and Jackson (2011) analyzed the effect of simultaneous data on system reliability prediction.

3.2.2 Computational Methods in Bayesian Inference

The posterior distribution resulting from a complex Bayesian model often cannot be written in a closed form. This results from the fact that the joint posterior distribution of multiple parameters in a complex model cannot be obtained analytically. This difficulty has hindered the adoption of Bayesian reliability assessment for many years. However, since the 1990s, advances in Bayesian computing through Markov chain Monte Carlo (MCMC) have facilitated inference based on samples from the targeted posterior distribution (Gelman *et al.*, 2014). MCMC is a simulation algorithm for performing Bayesian inference when conjugation is impossible (thus analytical result is impossible), which is particularly useful for high-dimensional Bayesian inference. MCMC algorithms draw samples from the joint posterior distribution of model parameters. Gibbs sampler, the most popular MCMC algorithm, relies on the fact that samples drawn sequentially from complete conditional distributions will converge to the joint posterior distribution as long as distribution parameters are constantly updated. So, after a certain number of preliminary iterations, the samples drawn from simulation chains can be viewed as from the targeted joint posterior distribution.

MCMC has also made the Bayesian models solvable when addressing the system reliability problem. In the reliability literature several authors used the MCMC technique for Bayesian inference (e.g., Johnson *et al.* (2003); Hamada *et al.* (2004); Reese *et al.* (2005a); Wilson *et al.* (2006); Graves *et al.* (2007); Wilson and Huzurbazar (2007); Graves *et al.* (2008); Pan and Rigdon (2009); Guo (2011)). To implement MCMC, we use WinBUGS, a statistical software for Bayesian inference (Spiegelhalter *et al.*, 2003).

3.3 Methodology

In this section Bayesian inference methods are discussed for simultaneously estimating conditional probabilities in a Bayesian network when data are collected from different levels of the system. We give a brief summary about Bayesian networks, and then present three different data scenarios with decreasing amount of available information along these scenarios. The first scenario involves a simple 2-state Bayesian network where all nodes and their states are recorded. We develop a conjugation model for inferring conditional probabilities and also extend it to a multi-state BN. In the second scenario, we discuss the case when we have only system failure records. Lastly, we consider a scenario where only the system and a subset of components are monitored by sensors, thus system health information is incomplete. We present a Bayesian inference method for estimating reliability dependency in such a system.

3.3.1 Bayesian Networks

Bayesian networks (BNs) are probabilistic graphical models depicting conditional independence relations and inducing a factorization into the joint probability mass/density function over the network variables (Koller and Friedman, 2009). The joint probabilities can be therefore expressed as a product of conditional probabilities, one for each variable given the corresponding values of the parent values. A Bayesian network consists of two main parts:

- Qualitative part: consists of a directed acyclic graph (DAG) where the nodes represent random variables (continuous or discrete) and directed arcs representing causal relationships between the random variables.

- Quantitative part: conditional probability tables between parent and child nodes.

In a BN, the nodes without any arrows directed into them are called root nodes and they are described according to their marginal probability distributions. The nodes that have arrows directed into them are called child nodes and the nodes that have arrows directed from them are called parent nodes. Each child has a conditional probability table associated with it, given the values of parent nodes.

Consider a BN over variables $X_1, X_2, \ldots, X_n$. By the chain rule of probability, the joint probability $P(X_1, X_2, \ldots, X_n)$ is

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid pa(X_i)) \tag{3.1}$$

where $pa(X_i)$ is the set of parents of node X_i.

Certain nodes in a BN may become uncorrelated if there is no link between these nodes. This situation is called conditional independence. These conditional independences allow us to decrease the number of terms in the chain rule, providing a simpler structure.

Figure 5 shows a BN with 3 nodes and 2 arcs. Each node C_i is a random variable. If there is a directed arc from C_i to C_j, C_i is called a "parent" of C_j. An arc characterizes the probabilistic dependency of a node on its parent nodes. That is, depending on the values a node's parents take on, the conditional probability distribution of the node may be different. In this example, node C_0 has 2 parents, C_1 and C_2. The marginal probabilities of these parent those are listed in Table 3. Assuming binary states for each node (functional state is 0 and dysfunctional state is 1), Table 4 shows the conditional probability table (CPT) for each combination of the parents of C_0. In this chapter, a BN is employed to represent the cause-and-effect failure relationship among elements of a multilevel system, in which the final child node represent the system and other nodes represent either components or subsystems of the system.

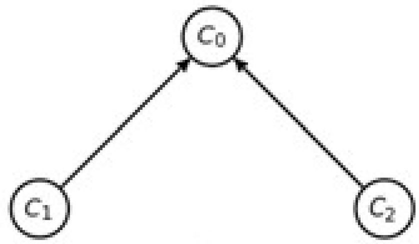

Figure 5. A simple 2-component system BN example

Table 3. Marginal probability tables for the BN example.

$C_1 = 0$	$C_1 = 1$
$1 - p_1$	p_1

$C_2 = 0$	$C_2 = 1$
$1 - p_2$	p_2

Table 4. Conditional probability table (CPT) for the BN example.

	$C_0 = 0$	$C_0 = 1$
$C_1 = 0, C_2 = 0$	$1 - p_{00}$	p_{00}
$C_1 = 0, C_2 = 1$	$1 - p_{01}$	p_{01}
$C_1 = 1, C_2 = 0$	$1 - p_{10}$	p_{10}
$C_1 = 1, C_2 = 1$	$1 - p_{11}$	p_{11}

The parameters, p_1 and p_2, listed in Table 3 are the distribution parameters of the marginal distributions (binomial) of the failure count variables of these 2 components, while Table 4 gives the parameters used in the conditional distribution of Eq. 3.1. A BN is fully defined if all of these parameters are specified, as the joint distribution of all nodes, Eq. 3.2, has become analytically available. In a fault tree (or reliability block diagram) representation of system reliability, the conditional probabilities in Table 4 are already pre-specified for a given logic gate (or the block diagram configuration). For example, for an "AND" gate, we have $p_{00} = p_{01} = p_{10} = 0$ and $p_{11} = 1$, while for an "OR" gate, we have $p_{00} = 0$ and $p_{11} = p_{10} = p_{01} = 1$. However, in our BN representation, these relationships are not pre-specified, as the dependency of system reliability to its components is unknown and needs to be evaluated by the data collected from the system and from other information sources such as expert opinions.

3.3.2 Bayesian Parameter Estimation in Bayesian Networks

In this section, we discuss Bayesian inference on model parameters in the context of a Bayesian network. Bayesian framework requires us to specify a joint distribution over the unknown parameters and the data instances. In this case, BN is parameterized by the marginal probabilities of components and conditional probabilities of the system given the states of the components. Suppose we want to estimate the parameters of the BN in Figure 5. Our network is parameterized by a parameter vector $\mathbf{p}$, where $\mathbf{p} = \{p_1, p_2, p_{11}, p_{10}, p_{01}, p_{00}\}$. Given the prior distribution of this parameter vector and the data collected from all nodes, Bayesian inference provides the posterior distribution of the parameter of interest and the posterior prediction of system or component reliability.

3.3.2.1 Complete System Log Data

Although it is uncommon in practice, we start our discussion with this naïve scenario – a complete history of the states of the system and its components are available. This is possible if this system and its components are continuously monitored by sensors and the log data from existing products can be obtained. In this example, each historical record is a tuple $C = \{C(i)\} = \{\langle C_0(i), C_1(i), C_2(i)\rangle\}$ for $i = 1, \ldots, N$ that describes a particular assignment (0 or 1) to nodes C_0, C_1 and C_2. The likelihood function is then given by

$$\begin{aligned}
L(C \mid \mathbf{p}) &= \prod_{i=1}^{N} P(C_0(i), C_1(i), C_2(i) \mid \mathbf{p}) \\
&= \prod_{i} P(C_1(i) \mid \mathbf{p}) P(C_2(i) \mid \mathbf{p}) P(C_0(i) \mid C_1(i), C_2(i), \mathbf{p}) \\
&= (\prod_{i} P(C_1(i) \mid \mathbf{p}))(\prod_{i} P(C_2(i) \mid \mathbf{p}))(\prod_{i} P(C_0(i) \mid C_1(i), C_2(i), \mathbf{p})) \quad (3.2)
\end{aligned}$$

According to the equation above, we have a separate factor for each node. These factors are called local likelihood functions and they depend on their corresponding node's conditional or marginal probability table parameters.

We can further decompose the conditional likelihood, $P(C_0(i) \mid C_1(i), C_2(i), \mathbf{p})$, as

$$\begin{aligned}
= &\prod_{C_1(i)=0, C_2(i)=0} P(C_0(i) \mid C_1(i), C_2(i), p_{00}) \prod_{C_1(i)=0, C_2(i)=1} P(C_0(i) \mid C_1(i), C_2(i), p_{01}) \\
&\prod_{C_1(i)=1, C_2(i)=0} P(C_0(i) \mid C_1(i), C_2(i), p_{10}) \prod_{C_1(i)=1, C_2(i)=1} P(C_0(i) \mid C_1(i), C_2(i), p_{11})
\end{aligned} \quad (3.3)$$

Assume that $M[C_0^x, C_1^y, C_2^z]$ represent the counts where $C_0(i) = x, C_1(i) = y$ and $C_2(i) = z$ ($x, y, z = 0$ or 1). Then, the terms in the right hand side of Eq. 3.3 can be reduced to be

$$\prod_{C_1(i)=1, C_2(i)=1} P(C_0(i) \mid C_1(i), C_2(i), p_{11}) = p_{11}^{M[C_0^1, C_1^1, C_2^1]} (1 - p_{11})^{M[C_0^0, C_1^1, C_2^1]} \quad (3.4)$$

As a result, the likelihood function of Eq. 3.2 becomes

$$\begin{aligned} L(C \mid \mathbf{p}) = & p_1^{M[C_1^1]}(1-p_1)^{M[C_1^0]} p_2^{M[C_2^1]}(1-p_2)^{M[C_2^0]} \\ & p_{00}^{M[C_0^1, C_1^0, C_2^0]}(1-p_{00})^{M[C_0^0, C_1^0, C_2^0]} p_{01}^{M[C_0^1, C_1^0, C_2^1]}(1-p_{01})^{M[C_0^0, C_1^0, C_2^1]} \\ & p_{10}^{M[C_0^1, C_1^1, C_2^0]}(1-p_{10})^{M[C_0^0, C_1^1, C_2^0]} p_{11}^{M[C_0^1, C_1^1, C_2^1]}(1-p_{11})^{M[C_0^0, C_1^1, C_2^1]} \end{aligned} \quad (3.5)$$

We can maximize the likelihood function above and get maximum likelihood function estimates for the parameters. However, even this simple formula could become troublesome in practice when there are many states for each component node. In such case, the number of combinations grows exponentially and the log file could be highly fragmented. There might be no observation for a particular combination. Therefore, it is better to combine the likelihood with expert opinions. This is equivalent to assigning a prior distribution to model parameter.

In this approach, we encode our prior knowledge about $\mathbf{p}$ with a probabilistic distribution. We now treat $\mathbf{p}$ as a random variable. According to the Bayes' formula, the posterior distribution over parameters given the observed data is

$$Pr(\mathbf{p} \mid C) = \frac{Pr(C \mid \mathbf{p})Pr(\mathbf{p})}{Pr(C)} \quad (3.6)$$

The term $Pr(\mathbf{p})$ is the prior distribution function of $\mathbf{p}$, $Pr(C \mid \mathbf{p})$ is the likelihood function, and $Pr(C)$ can be viewed as a normalizing constant.

Since all model parameters are probabilities, an appropriate prior is the beta distribution. A Beta distribution is specified by two hyperparameters – a and b, which are positive real numbers. The distribution is defined as follows:

$$\theta \sim Beta(a, b) \quad \text{with pdf of} \quad p(\theta) = \gamma \theta^{a-1}(1-\theta)^{b-1} \quad (3.7)$$

where γ is a normalizing constant, defined by

$$\gamma = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \quad (3.8)$$

where $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the Gamma function.

If a beta prior distribution, $Beta(a_i, b_i)$, is assumed for each p_i, it is easy to shown that the posterior distribution of $\mathbf{p}$ is given by

$$
\begin{aligned}
Pr(\mathbf{p} \mid C) \propto Pr(C \mid \mathbf{p}) Pr(\mathbf{p}) & \\
\propto\ & p_1^{M[C_1^1]}(1-p_1)^{M[C_1^0]} p_2^{M[C_2^1]}(1-p_2)^{M[C_2^0]} \\
& p_{00}^{M[C_0^1,C_1^0,C_2^0]}(1-p_{00})^{M[C_0^0,C_1^0,C_2^0]} p_{01}^{M[C_0^1,C_1^0,C_2^1]}(1-p_{01})^{M[C_0^0,C_1^0,C_2^1]} \\
& p_{10}^{M[C_0^1,C_1^1,C_2^0]}(1-p_{10})^{M[C_0^0,C_1^1,C_2^0]} p_{11}^{M[C_0^1,C_1^1,C_2^1]}(1-p_{11})^{M[C_0^0,C_1^1,C_2^1]} \\
& p_1^{a_1-1}(1-p_1)^{b_1-1} p_2^{a_2-1}(1-p_2)^{b_2-1} p_{00}^{a_{00}-1}(1-p_{00})^{b_{00}-1} \\
& p_{01}^{a_{01}-1}(1-p_{01})^{b_{01}-1} p_{10}^{a_{10}-1}(1-p_{10})^{b_{10}-1} p_{11}^{a_{11}-1}(1-p_{11})^{b_{11}-1} \\
=\ & p_1^{a_1+M[C_1^1]-1}(1-p_1)^{b_1+M[C_1^0]-1} p_2^{a_2+M[C_2^1]-1}(1-p_2)^{b_2+M[C_2^0]-1} \\
& p_{00}^{a_{00}+M[C_0^1,C_1^0,C_2^0]-1}(1-p_{00})^{b_{00}+M[C_0^0,C_1^0,C_2^0]-1} p_{01}^{a_{01}+M[C_0^1,C_1^0,C_2^1]-1} \\
& (1-p_{01})^{b_{01}+M[C_0^0,C_1^0,C_2^1]-1} p_{10}^{a_{10}+M[C_0^1,C_1^1,C_2^0]-1}(1-p_{10})^{b_{10}+M[C_0^0,C_1^1,C_2^0]-1} \\
& p_{11}^{a_{11}+M[C_0^1,C_1^1,C_2^1]-1}(1-p_{11})^{b_{11}+M[C_0^0,C_1^1,C_2^1]-1}
\end{aligned} \tag{3.9}
$$

Eq. 3.9 is comprised of beta distributions for all parameters in the BN. That is, their posterior distributions are analytically available as

$$
\begin{aligned}
p_i &\sim Beta(a_i + M[C_i^1], b_i + M[C_i^0]) \\
p_{ij} &\sim Beta(a_{ij} + M[C_0^1, C_1^i, C_2^j], b_{ij} + M[C_0^0, C_1^i, C_2^j])
\end{aligned}
$$

This result illustrates a conjugation property of the beta distribution when coupled with binomial likelihood (see Koller and Friedman (2009) for more details). Exact inference is possible in case of binomial likelihood (pass/fail tests) and fully observed variables. In literature, beta-binomial conjugation has been extensively used for reliability prediction (see Martz *et al.* (1988); Martz and Wailer (1990); Johnson *et al.* (2003); Wilson and Huzurbazar (2007)).

We can easily extend the discussion to multi-state models. In our BN model so far, we have assumed that all nodes have two distinct states: pass and fail. Now, consider the system and components having multiple states; for example, k states for the system, C_0, such as $\{0, \ldots, k-1\}$, where state $k-1$ represents state of failure, state 0 represents state of full functionality and the rest of the states between represent degraded states; l states for component C_1 and m states for the component C_2 (see Tables 5 and 6). The likelihood function of the system can be derived from multinomial distributions and the conjugate priors need to be specified by Dirichlet distributions.

Table 5. Marginal probability tables for the multi-state BN example

$C_1 = 0$	$C_1 = 1$	$\cdots$	$C_1 = l-1$
p_1^0	p_1^1	$\cdots$	p_1^{l-1}

$C_2 = 0$	$C_2 = 1$	$\cdots$	$C_2 = m-1$
p_2^0	p_2^1	$\cdots$	p_2^{m-1}

Table 6. Conditional probability table (CPT) for the multi-state BN example.

	$C_0 = 0$	$C_0 = 1$	$\cdots$	$C_0 = k-1$
$C_1 = 0, C_2 = 0$	p_{00}^0	p_{00}^1	$\cdots$	p_{00}^{k-1}
$C_1 = 0, C_2 = 1$	p_{01}^0	p_{01}^1	$\cdots$	p_{01}^{k-1}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$C_1 = i, C_2 = j$	p_{ij}^0	p_{ij}^1	$\cdots$	p_{ij}^{k-1}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$C_1 = l-1, C_2 = m-1$	$p_{l-1,m-1}^0$	$p_{l-1,m-1}^1$	$\cdots$	$p_{l-1,m-1}^{k-1}$

Note that $\sum_L p_1^L = 1$, $\sum_M p_2^M = 1$ and $\sum_K p_{ij}^K = 1$ for $\forall i, j$.

The likelihood function in this model has the same form as in Eq. 3.5. However, since there are multiple states, there will be a larger number of combinations of parent nodes. The likelihood function has the following multinomial form:

$$L(C \mid \mathbf{p}) = \prod_L (p_1^L)^{M[C_1^L]} \prod_M (p_2^M)^{M[C_2^M]} \prod_{\forall i,j} \left(\prod_K (p_{ij}^K)^{M[C_0^K, C_1^i, C_2^j]} \right) \tag{3.10}$$

In this case, an appropriate prior for the probabilities in the model is Dirichlet distribution, which is a generalization of beta distribution. A Dirichlet distribution is specified by a set of hyperparameters $\alpha_1, \ldots, \alpha_k$, so that

$$\theta \sim Dirichlet(\alpha_1, \ldots, \alpha_K) \quad \text{with the pdf of} \quad P(\theta) \propto \prod_k \theta_k^{\alpha_k - 1} \tag{3.11}$$

If we assume a Dirichlet prior distribution, $Dirichlet(\alpha_i^0, \ldots, \alpha_i^{k-1})$, for p_i, then the posterior distribution of $\mathbf{p}$ is given by

$$\begin{aligned}
P(\mathbf{p} \mid C) \propto & P(C \mid \mathbf{p})P(\mathbf{p}) \\
& \propto \prod_L (p_1^L)^{M[C_1^L]} \prod_M (p_2^M)^{M[C_2^M]} \prod_{\forall i,j} \left(\prod_K (p_{ij}^K)^{M[C_0^K, C_1^i, C_2^j]} \right) \\
& \prod_L (p_1^L)^{\alpha_1^L - 1} \prod_M (p_2^M)^{\alpha_2^M - 1} \prod_{\forall i,j} \left(\prod_K (p_{ij}^K)^{\alpha_{ij}^K - 1} \right) \\
& = \prod_L (p_1^L)^{\alpha_1^L + M[C_1^L] - 1} \prod_M (p_2^M)^{\alpha_2^M + M[C_2^M] - 1} \prod_{\forall i,j} \left(\prod_K (p_{ij}^K)^{\alpha_{ij}^K + M[C_0^K, C_1^i, C_2^j] - 1} \right)
\end{aligned} \tag{3.12}$$

Eq. 3.12 is comprised of posterior Dirichlet distributions for all parameters in the BN; that is,

$$\begin{aligned}
p_i &\sim Dirichlet(a_i^0 + M[C_i^0], \ldots, a_i^{k-1} + M[C_i^{k-1}]) \\
p_{ij} &\sim Dirichlet(a_{ij}^0 + M[C_0^0, C_1^i, C_2^j], \ldots, a_{ij}^{k-1} + M[C_0^{k-1}, C_1^i, C_2^j])
\end{aligned}$$

Bayesian conjugation is convenient for obtaining analytical results; however, in most scenarios, the prior distribution may not come from a conjugation family and the system is too complex to model with conjugate pairs. For such cases, computational Bayesian methods such as MCMC need to be employed.

It is also of research interest to examine the effect of prior distribution assumption on posterior estimation. In general, specifying a more informative prior reduces the variance of the posterior distribution, resulting in a more precise estimation. Therefore, eliciting prior distributions in Bayesian inference is rather important for representing prior knowledge more accurately and comprehensively. However, it is not usually a straightforward task to elicit prior distributions for the parameters of the model and special techniques must be used. One of the most commonly used techniques is expert elicitation, which converts an expert's opinions into a statistical expression of these opinions Garthwaite and O'Hagan (2000). Experts are asked to give their opinions about quantities for the distribution parameters such as the mean, mode and median values. As a result, we can obtain an appropriate prior for the parameters.

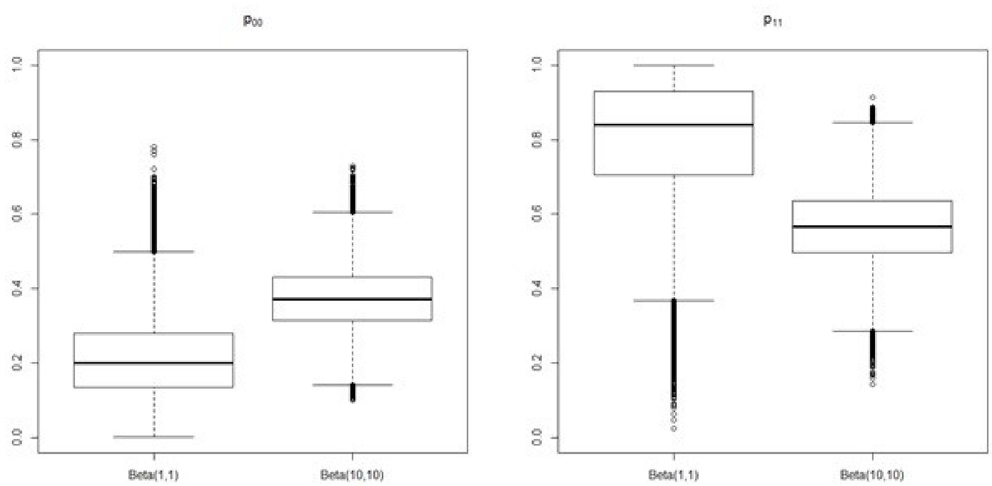

Figure 6. Box plots of conditional probabilities with different prior distributions

A sensitivity analysis has been carried out using the system in Figure 6 to show the effect of using a more informative prior. A dataset consisting of pass/fail data for all components was simulated and used as observations for calculating the likelihood. $Beta(1, 1)$ and $Beta(10, 10)$ were assigned as the priors for model parameters, separately. The box plots of the posterior samples of some model parameters are shown in Figure 6. According to the results, we get more precise results when $Beta(10, 10)$ is used as a prior. Therefore, we would like to emphasize that special cares to these prior distribution assignments are needed when Bayesian inference is in use.

3.3.2.2 Summarized System Failure Data

A system failure record is often maintained within an organization and it is the most common type of information that one can track for system failure diagnosis. In this case, once a system failure occurs, the components that are causing the failure are identified and this event is recorded. For example, given a checklist such as Table 7, one can see that a failure event occurred once due to C_1 and C_2, and once due to C_1 only, etc.

Table 7. System failure records.

System - C_0	C_1	C_2
Failure event 1	$\times$	$\times$
Failure event 2	$\times$	
$\cdots$	$\cdots$	$\cdots$

Notice that, with these records, we can directly estimate the probability of component failure given a system failure, i.e., $Pr(C_1, C_2 \mid C_0 = 1)$, but not the probability of system failure given the states of components. This is because, unlike the log data, Table 7 records only system failure events. The joint probability of component states and the marginal probability of system failure are required in order to obtain the conditional probability of system failure, because

$$Pr(C_0 \mid C_1, C_2) = \frac{Pr(C_1, C_2 \mid C_0)Pr(C_0)}{Pr(C_1, C_2)} \tag{3.13}$$

In Eq. 3.13, $Pr(C_0)$ represents the prior knowledge about system failure and $Pr(C_0 \mid C_1, C_2)$ represents the posterior failure distribution after observing the failure record data. Yontay *et al.* (2015) discussed a method for deriving the prior probability, $Pr(C_0)$. If each failure event is recorded with its time stamp, we can use failure times to estimate the failure rate of the system. Assuming the time to failure is exponentially distributed, after estimating the occurrence rate of the failure events, we can then convert failure times to event probabilities based on exponential distribution. That is,

$$Pr(C_0 = 1) = Pr(T < t) = 1 - e^{-\lambda_F t} \tag{3.14}$$

where λ_F is the occurrence rate of system failure event and t is the system lifetime. The next step is to calculate, $Pr(C_1, C_2 \mid C_0)$, which is the likelihood for each combination of component states, using Table 7.

As an example, consider the system in Figure 5. In this scenario, the system failure might be caused by C_1 or C_2, or C_1 and C_2 together, or the system might fail even when both of the components are functioning (by an unknown failure cause). Given the recorded failure times, we can obtain an initial estimate of the prior distribution for system failure, which is defined as $Beta(1.28, 1.30)$. The field observations of the system, which are summarized as the counts for each combination as shown in Table 8, can be modeled by a multinomial distribution.

We ran simulations in WinBUGS and obtained the results in Table 9. Since the system failure probability when at least one of the components is working is very small, we can conclude that the system behaves like a parallel system. However, since there exists an un-ignorable probability of system failure (its mean value is 0.0516 and 95% credible interval is $[0.01779, 0.1019]$) when both components are functional, it indicates some unknown factors that are influencing system reliability. As a result, we need to conduct further investigation of these unknown factors.

Table 8. Data from a system failure record.

Cause combinations	Counts
$C_1 = 0, C_2 = 0$	4
$C_1 = 0, C_2 = 1$	9
$C_1 = 1, C_2 = 0$	12
$C_1 = 1, C_2 = 1$	75

Table 9. Empirical mean, standard deviation, and quantiles for posterior failure probabilities.

	Mean	2.5%	25%	50%	75%	97.5%
$Pr(C_0 = 1 \mid C_1 = 0, C_2 = 0)$	0.0516	0.01779	0.03574	0.04863	0.06429	0.1019
$Pr(C_0 = 1 \mid C_1 = 0, C_2 = 1)$	0.1002	0.05038	0.07898	0.09769	0.1186	0.1646
$Pr(C_0 = 1 \mid C_1 = 1, C_2 = 0)$	0.1294	0.07218	0.1059	0.1271	0.1503	0.2001
$Pr(C_0 = 1 \mid C_1 = 1, C_2 = 1)$	0.7434	0.6542	0.715	0.7451	0.7735	0.8227

This approach can also been seen as an extension of the reliability parenting process presented in Sanchez and Pan (2011), in which the authors utilized the failure information of old-generation products stored in a failure database.

3.3.2.3 Incomplete Lower-Level Data

One big challenge in system reliability assessment is the lack of the complete lower-level data as presented in previous sections. A complex system does not necessarily have all components or subsystems being monitored at the same time. There can be a limited number of sensors deployed in the system to monitor the states of some components or subsystems, but not all of them. In addition, these sensor data are stored by sensor, not in the system format such as the row entries in Table 7.

Since system's functionality is conditional on the functionality of subsystems and components, collectively analyzing these data yields significant information about the reliability. However, data collected by multiple sensors in the same system at multiple system levels may contain duplicated system reliability information, thus they require different data analysis technique.

The basic problem for analyzing this type of data is that we cannot treat them as independent data although they come from individual sensors. The dependencies between the states of systems and components under monitoring must be taken into consideration in data analysis. Only a few previous studies have addressed this problem. Graves *et al.* (2008) proposed a method that incorporates overlapping data for traditional binary-state series/parallel systems. Their methodology relies on disjoint cut-set generation and considers each observation in isolation. Jackson (2011) extended this line of research by adding continuous failure time data. However, their methodology can only apply to the system failure that is represented by a fault tree. In addition, using their approach, generating all possible system failure cases was cumbersome. In this section, we consider the data scenario with simultaneous, multi-level sensor data from the same system and incorporate it into the BN model analysis. A Bayesian inference method is developed for dealing with simultaneous higher-level data and partial lower-level data.

Suppose that a system-level sensor monitors the system's health status. Some (not all) of its components/subsystems are also monitored by their own sensors. Each sensor will store the information such as how many failures occurred in a time interval (e.g., a day). These failures at different levels are correlated, as they come from the same system. For instance, considering a two-component series system, if the system is known to be functioning, this implies that both components must be functioning too. But, if both components are not monitored and the system is not functioning, it is impossible to know which component has failed or both of them failed. Only if we have one component monitored, the other component's state can be inferred by the observations at both system and component levels. In general, tracking and consolidating the states of monitored system and components can be done when a deterministic system reliability configuration is known. However, this process can be very tedious and varies according to system configuration. Using BN models, we are able to provide a generic algorithm of sensor data consolidation and code it into a computer program.

To develop the likelihood function of a BN model with simultaneous, multi-level sensor data, all possible instances of component and system states that imply the observed evidence by sensors need to be captured. To formulate the probability function for each of these combinations, we start by constructing state vector of all nodes in a Bayesian network. The state variable of the i^{th} node is denoted by x_i, (0 for functional and 1 for dysfunctional). The states of all nodes are given by the state vector, $\mathbf{x} = \{x_1, x_2, \ldots, x_n, x_0\}$, when the BN model has n component nodes and one system node (x_0).

Assume that all nodes are binary-state nodes, then there are $2^{(n+1)}$ possible combinations and hence $2^{(n+1)}$ possible state vectors. For example, for a 2-component system, there are $2^{(2+1)} = 8$ possible state vectors. They can be represented such that: $\mathbf{x_1} = \{0,0,0\}, \mathbf{x_2} = \{0,0,1\}, \mathbf{x_3} = \{0,1,0\}, \mathbf{x_4} = \{0,1,1\}, \mathbf{x_5} = \{1,0,0\}, \mathbf{x_6} = \{1,0,1\}, \mathbf{x_7} = \{1,1,0\}$, and $\mathbf{x_8} = \{1,1,1\}$. The probability of each state vector's occurrence is defined by the joint distribution of the BN (see Eq. 2.1). As an example, for the 2-component system in Figure 1, we can define the joint probability of each combination such as

$$
\begin{aligned}
Pr(\mathbf{x_1}) &= Pr(x_1 = 0)Pr(x_2 = 0)Pr(x_0 = 0 \mid x_1 = 0, x_2 = 0) \\
Pr(\mathbf{x_2}) &= Pr(x_1 = 0)Pr(x_2 = 0)Pr(x_0 = 1 \mid x_1 = 0, x_2 = 0) \\
Pr(\mathbf{x_3}) &= Pr(x_1 = 0)Pr(x_2 = 1)Pr(x_0 = 0 \mid x_1 = 0, x_2 = 1) \\
Pr(\mathbf{x_4}) &= Pr(x_1 = 0)Pr(x_2 = 1)Pr(x_0 = 1 \mid x_1 = 0, x_2 = 1) \\
Pr(\mathbf{x_5}) &= Pr(x_1 = 1)Pr(x_2 = 0)Pr(x_0 = 0 \mid x_1 = 1, x_2 = 0) \\
Pr(\mathbf{x_6}) &= Pr(x_1 = 1)Pr(x_2 = 0)Pr(x_0 = 1 \mid x_1 = 1, x_2 = 0) \\
Pr(\mathbf{x_7}) &= Pr(x_1 = 1)Pr(x_2 = 1)Pr(x_0 = 0 \mid x_1 = 1, x_2 = 1) \\
Pr(\mathbf{x_8}) &= Pr(x_1 = 1)Pr(x_2 = 1)Pr(x_0 = 1 \mid x_1 = 1, x_2 = 1) \qquad (3.15)
\end{aligned}
$$

with the constraint that $\sum_{i=1}^{8} Pr(\mathbf{x_i}) = 1$.

After formulating these state vector probabilities, we need to count how many times each state vector is observed in a specific evidence set. Thus, we represent the occurrence of each state vector by a count vector, $\mathbf{y} = \{y_1, y_2, \ldots, y_j, \ldots, y_{2^{n+1}}\}$, where y_j is the number of occurrences of the j^{th} state vector, $\mathbf{x_j}$.

Consider the 2-component system example in Figure 5. We need to keep track of the counts for each of the 8 state vectors. If we observe the state vector $\mathbf{x_8} = \{1, 1, 1\}$ 2 times in an evidence set, then $y_8 = 2$. If we also observe $\mathbf{x_4} = \{0, 1, 1\}$ once, then $y_4 = 1$. Combining them together, the count vector is given by $\mathbf{y} = \{0, 0, 0, 1, 0, 0, 0, 2\}$.

The likelihood function of specific evidence set is derived from a multinomial distribution. As a sensor signal only depends on the state of the node under its monitoring, each observation set from the system leads to exactly one state vector, then the count vector clearly follows a multinomial distribution with its parameters being the state vector probabilities defined in Eq. 3.15. That is, the random variables

y_j indicate the number of occurrence state vector $\mathbf{x_j}$ observed over N instances (total number of sensor signals).

Therefore, the likelihood function of one specific evidence set is given by

$$
\begin{aligned}
Pr(\mathbf{y} \mid \mathbf{p}) =& \frac{N!}{y_1!y_2!\dots y_{2^{n+1}}!}\,(Pr(\mathbf{x_1}))^{y_1}\,(Pr(\mathbf{x_2}))^{y_2}\dots(Pr(\mathbf{x_{2^{n+1}}}))^{y_{2^{n+1}}} \\
=& N!\prod_{j=1}^{2^{n+1}}\frac{\left(Pr(\mathbf{x_j})\right)^{y_j}}{y_j!} \\
=& N!\prod_{j=1}^{2^{n+1}}\{\frac{1}{y_j!}[\left(\prod_{i=1}^{n}(p_i)^{(x_i)_j}(1-p_i)^{[1-(x_i)_j]}\right) \\
& (p_{(x_1)_j\dots(x_n)_j})^{(x_0)_j}(1-p_{(x_1)_j\dots(x_n)_j})^{[1-(x_0)_j]}]\}
\end{aligned}
\tag{3.16}
$$

When there are only a partial set of components are monitored, it is important to realize that there could be more than one count vector that satisfy the evidence set from sensors. Thus, we need to keep track of the count vector for each possible scenario. Let the k^{th} possible count vector to be $\mathbf{y_k} = \{(y_1)_k, (y_2)_k, \dots, (y_l)_k, \dots, (y_{2^{n+1}})_k\}$, where $(y_j)_k$ is the number of occurrences of the j^{th} state vector, $\mathbf{x_j}$, in the k^{th} scenario that satisfies the given evidence. Then, the likelihood of observing the evidence, E, should be the sum of the probability of all possible count vectors that these evidences imply. That is,

$$
\begin{aligned}
Pr(E \mid \mathbf{p}) =& \sum_{\forall \mathbf{y_k}} Pr(\mathbf{y_k} \mid \mathbf{p}) \\
=& N!\sum_{\forall \mathbf{y_k}}[\prod_{j=1}^{2^{n+1}}\{\frac{1}{y_j!}[\left(\prod_{i=1}^{n}(p_i)^{(x_i)_j}(1-p_i)^{[1-(x_i)_j]}\right) \\
& (p_{(x_1)_j\dots(x_n)_j})^{(x_0)_j}(1-p_{(x_1)_j\dots(x_n)_j})^{[1-(x_0)_j]}]\}]
\end{aligned}
\tag{3.17}
$$

Therefore,

$$L(E \mid \mathbf{p}) \propto \sum_{\forall \mathbf{y_k}} [\prod_{j=1}^{2^{n+1}} \{\frac{1}{y_j!} [\left(\prod_{i=1}^{n} (p_i)^{(x_i)_j} (1-p_i)^{[1-(x_i)_j]} \right)$$

$$(p_{(x_1)_j \cdots (x_n)_j})^{(x_0)_j} (1-p_{(x_1)_j \cdots (x_n)_j})^{[1-(x_0)_j]}]\}] \tag{3.18}$$

To illustrate the computation, we use the BN model in Figure 5 as an example. In this 2-component system we assume there is one sensor placed on the component 1 node and another sensor on the system node (see Figure 7). Over the observation period, a series of 5 failure events were detected at the system level by sensor 1 and one failure event was detected at the component level by sensor 2. However, no direct information of component 2 is available, as it is not monitored by sensor.

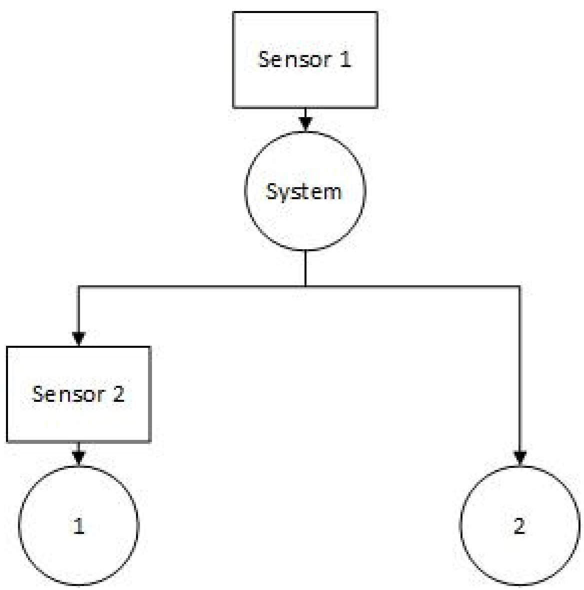

Figure 7. Basic two component BN system with sensors on the system and component 1.

Since there are two components (i.e. $n = 2$), the number of possible state vectors is $2^{n+1} = 2^3 = 8$. The state vectors are listed in Table 10, along with their probabilities.

Table 10. State Vectors of system in Figure 7.

State Vector # j	Node states			State vector $\mathbf{x_j}$	Probability $Pr(\mathbf{x_j} \mid \mathbf{p})$
	$(x_1)_j$	$(x_2)_j$	$(x_0)_j$		
1	0	0	0	$\{0,0,0\}$	$(1-p_1)(1-p_2)(1-p_{00})$
2	0	0	1	$\{0,0,1\}$	$(1-p_1)(1-p_2)p_{00}$
3	0	1	0	$\{0,1,0\}$	$(1-p_1)p_2(1-p_{01})$
4	0	1	1	$\{0,1,1\}$	$(1-p_1)p_2p_{01}$
5	1	0	0	$\{1,0,0\}$	$p_1(1-p_2)(1-p_{10})$
6	1	0	1	$\{1,0,1\}$	$p_1(1-p_2)p_{10}$
7	1	1	0	$\{1,1,0\}$	$p_1p_2(1-p_{11})$
8	1	1	1	$\{1,1,1\}$	$p_1p_2p_{11}$

The five observed system failure events are certainly related to the events at the component level. For each system event, it invokes one or more of the 8 possible state vectors. In this example, as we observe 5 failures at the system and 1 failure at component 1, the state vectors must be four $\{0, x_2, 1\}$ and one $\{1, x_2, 1\}$. As there are two possible states for the unobservable node x_2, the four events of $\{0, x_2, 1\}$ are distributed among two possible state vectors and there are 5 distinct arrangements. Similarly, there are 2 arrangements for the single event of $\{1, x_2, 1\}$. Thus, the total number of possible count vectors is $5 \times 2 = 10$.

For example, among the 5 observed system events, one possible scenario is that $\mathbf{x_2} = \{0,0,1\}$ occurred 4 times and $\mathbf{x_6} = \{1,0,1\}$ occurred 1 time. Correspondingly, we have $y_2 = 4$ and $y_6 = 1$, and other y_j's are zeros. This is the first row in Table 11. Another possible scenario is that $\mathbf{x_4} = \{0,1,1\}$ occurred 4 times and $\mathbf{x_8} = \{1,1,1\}$ occurred 1 time. This is the last row in Table 11. After enumerating all possible scenarios, their corresponding count vectors are listed in Table 11.

We used the likelihood function given by Eq. 3.18 along with uniform prior distributions of p to generate the posterior distributions of the parameters of BN model. MCMC was performed to draw samples from the unnormalized joint posterior distribution. We used the Bayesian software package, *WinBUGS*, to carry out the computation. One advantage of using *WinBUGS* software is that it can be also called from the statistical software *R* (R CORE TEAM *et al.*, 2012) through a package called *R2WinBUGS*, making it more convenient for the analysis of simulation results.

Table 11. Possible state vector combinations of system in Figure 7.

Count Vector # k	Count vector, $\mathbf{y_k}$ (no. of j^{th} state vectors)							
	$(y_1)_k$	$(y_2)_k$	$(y_3)_k$	$(y_4)_k$	$(y_5)_k$	$(y_6)_k$	$(y_7)_k$	$(y_8)_k$
1	0	4	0	0	0	1	0	0
2	0	4	0	0	0	0	0	1
3	0	3	0	1	0	1	0	0
4	0	3	0	1	0	0	0	1
5	0	2	0	2	0	1	0	0
6	0	2	0	2	0	0	0	1
7	0	1	0	3	0	1	0	0
8	0	1	0	3	0	0	0	1
9	0	0	0	4	0	1	0	0
10	0	0	0	4	0	0	0	1

The following results are based on discarding the first $20,000$ draws from the MCMC sampling chain and then keeping every other sample (to reduce the auto-correlation of drawn samples) until there were $100,000$ draws from the joint posterior distribution.

Based on the results in Table 12, we can conclude that the system behaves like a series system, because the system has high probability of failure when at least one of the components has failed. It is also found that there is a notable probability of system failure even when both components are functioning, so there might be some unknown factors that affect the working mechanism of the system. As a result, we are able to assess the dependencies between the system's health and the states of its components even when only a partial set of components are monitored.

Table 12. Empirical mean, standard deviation, and quantiles for parameters.

	Mean	**SD**	**2.5%**	**25%**	**50%**	**75%**	**97.5%**
p_1	0.148967691	0.062107504	0.02870975	0.1008	0.1535	0.2015	0.2452
p_2	0.169756133	0.062686748	0.02449975	0.1295	0.1849	0.2217	0.2475
p_{00}	0.183846816	0.055111599	0.0452995	0.1531	0.1987	0.2279	0.2481
p_{01}	0.799470654	0.138742057	0.524	0.6922	0.8218	0.9197	0.9926
p_{10}	0.772466599	0.142763382	0.5171	0.6534	0.7832	0.8973	0.9899
p_{11}	0.876228065	0.072134488	0.7562	0.8143	0.8767	0.9389	0.9938

3.4 Case Study

In this section, we demonstrate our methodology on a hypothetical mechatronic system: an active vehicle suspension (AVS), previously presented in Zhong *et al.* (2010). In the previous study, the system reliability configuration was deterministic, represented by a fault tree. In our study, we remodel one of its subsystems by a BN, assuming that this subsystem is redesigned and its reliability structure is more complex than the old generation. We start by introducing the AVS system.

The AVS system supports the vehicle body and reduces body vibration from the road surface. The system consists of tires, springs, dampers (shock absorbers) and linkages that connect a vehicle to its wheels and allows relative motion between the two. Suspension systems contribute to the vehicle's road handling and braking for good active safety, and keep vehicle occupants isolated from road noise and bumps. The suspension also protects the vehicle from damage and wear. Fully active suspension systems use electronic monitoring of vehicle conditions, in order to impact vehicle suspension and behavior in real time to directly control the motion of the car.

Figure 8 shows the fault tree of a simplified version of the system. The system has a parallel structure. The parallel system is composed of two subsystems: a passive subsystem and an actuator subsystem. The passive subsystem works in a series structure with the spring and damper (shock absorber) components, where the shock absorbers damp out the motions of a vehicle up and down on its springs. The actuator subsystem also works in a series structure with mechanical and electronic parts. Active suspensions use actuators to raise and lower the chassis independently at each wheel. The mechanical parts include components like pump, piston, and servovalve; whereas the electronic parts include power, sensors, and the controller. The suspension reacts to signals from the electronic controller (which means the suspension is externally controlled). Sensors continually monitor body movement and vehicle ride level, constantly supplying the computer with new data.

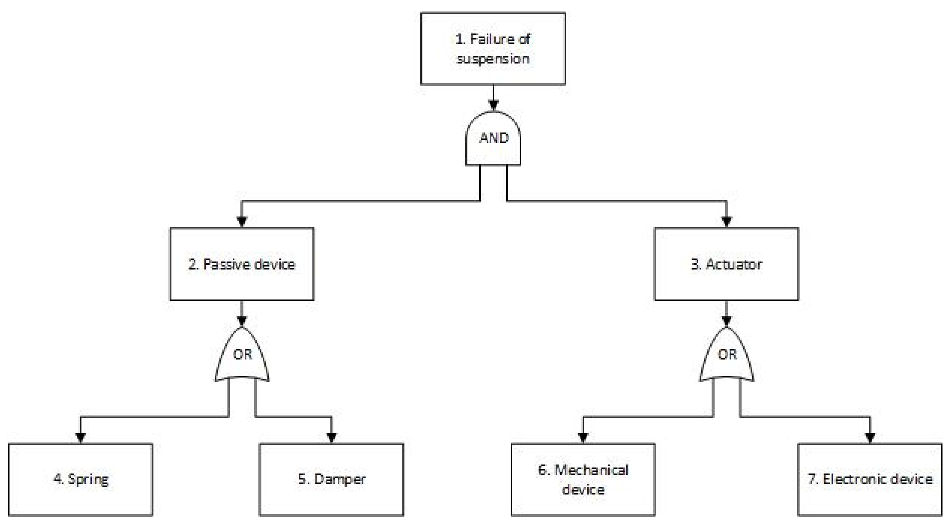

Figure 8. The fault tree of an active vehicle suspension.

Next, we model the AVS system as a Bayesian network (see Figure 9). Suppose that the parallel structure of the system reliability and the series structure of the actuator reliability are unchanged, but, due to a redesign, the reliability structure of the passive device reliability becomes uncertain. Therefore, we are interested in exploring the relationship between node X_2 and its parent nodes, X_4 and X_5, through conditional probabilities.

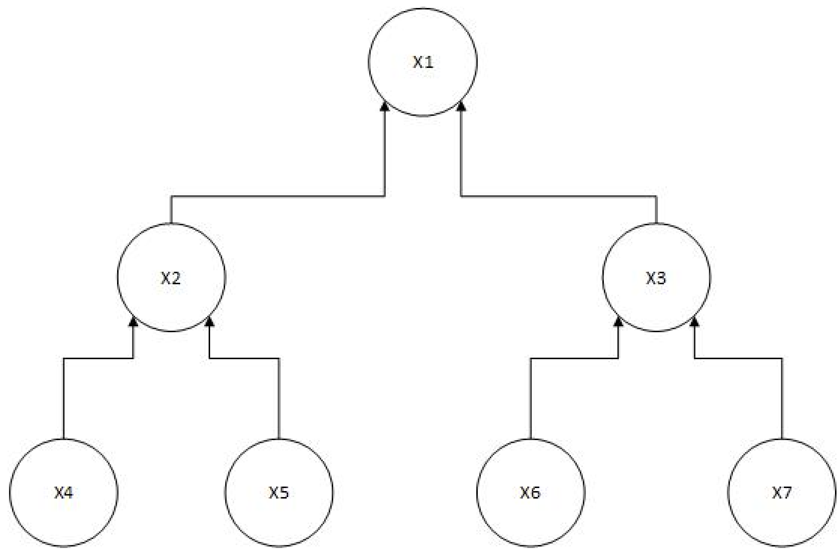

Figure 9. The corresponding BN model of the fault tree model in Figure 8.

In this scenario we continually monitor the system with sensors on nodes X_1, X_4 and X_7. We observe a series of 10 events where 10 failures were detected at the system level (by sensor 1), 2 failures were detected by sensor 2 and no failures were detected by sensor 3 (see Figure 10).

Since there are 7 components in the system, the number of possible state vectors would be $2^7 = 128$ if we did not observe any evidence. As some parts of system reliability structure are deterministic, we can eliminate a great amount of state vectors according to the evidence coming from the sensor.

The first step is to construct the state vectors as explained in Section 3.3.2.3. The states of X_6 and X_7 uniquely define the state of X_3, and the states of X_2 and X_3 uniquely define the state of X_1. Therefore, we only need to consider the stochastic nodes, X_2, X_4, X_5, X_6, X_7, in the model inference. Thus, we have $2^5 = 32$ state vectors. The joint probability is represented as

$$P(\mathbf{X}) = P(X_4)P(X_5)P(X_6)P(X_7)P(X_2 \mid X_4, X_5) \tag{3.19}$$

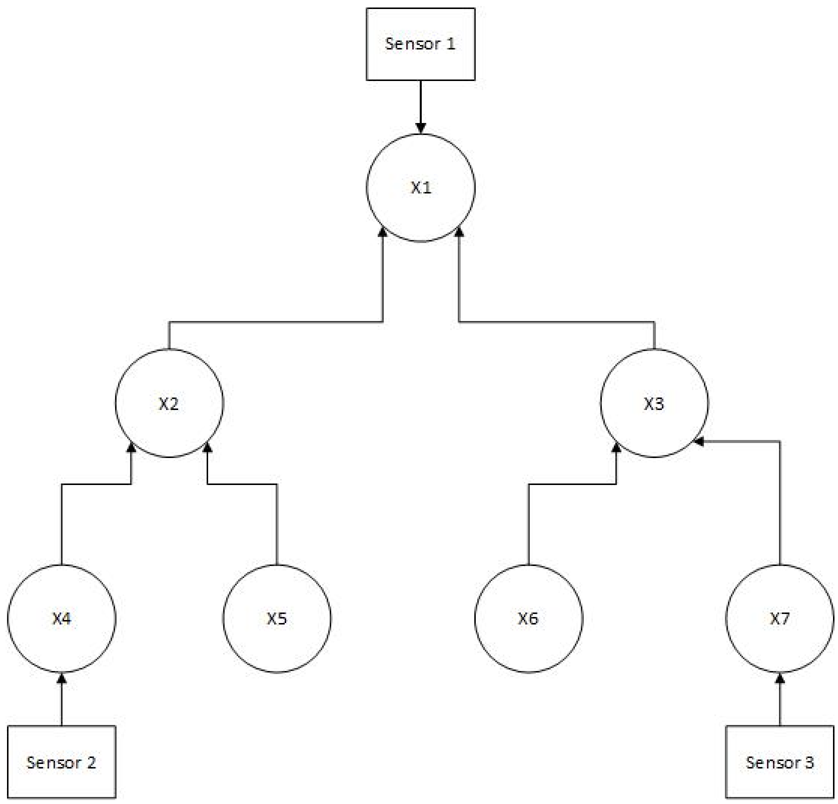

Figure 10. The AVS model with sensors.

Therefore, the parameters that we would like to estimate in this system are the failure probabilities $\mathbf{p} = \{p_4, p_5, p_6, p_7, p_{11}, p_{10}, p_{01}, p_{00}\}$ where $p_{ij} = P(X_2 = 1 \mid X_4 = i, X_5 = j)$. Note that $p_3 = (1 - (1 - p_6)(1 - p_7)$ (series system) and $p_1 = p_2 p_3$ (parallel system). The state vectors are listed in Table 13, along with their probabilities.

Table 13. State Vectors of system in Figure 10.

Vector	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Probability
1	0	0	0	0	0	0	0	$(1-p_4)(1-p_5)(1-p_6)(1-p_7)(1-p_{00})$
2	0	0	0	0	1	0	0	$(1-p_4)p_5(1-p_6)(1-p_7)(1-p_{01})$
3	0	0	0	1	0	0	0	$p_4(1-p_5)(1-p_6)(1-p_7)(1-p_{10})$
4	0	0	0	1	1	0	0	$p_4p_5(1-p_6)(1-p_7)(1-p_{11})$
5	0	0	1	0	0	0	1	$(1-p_4)(1-p_5)(1-p_6)p_7(1-p_{00})$
6	0	0	1	0	0	1	0	$(1-p_4)(1-p_5)p_6(1-p_7)(1-p_{00})$
7	0	0	1	0	0	1	1	$(1-p_4)(1-p_5)p_6p_7(1-p_{00})$
8	0	0	1	0	1	0	1	$(1-p_4)p_5(1-p_6)p_7(1-p_{01})$
9	0	0	1	0	1	1	0	$(1-p_4)p_5p_6(1-p_7)(1-p_{01})$
10	0	0	1	0	1	1	1	$(1-p_4)p_5p_6p_7(1-p_{01})$
11	0	0	1	1	0	0	1	$p_4(1-p_5)(1-p_6)p_7(1-p_{10})$
12	0	0	1	1	0	1	0	$p_4(1-p_5)p_6(1-p_7)(1-p_{10})$
13	0	0	1	1	0	1	1	$p_4(1-p_5)p_6p_7(1-p_{10})$
14	0	0	1	1	1	0	1	$p_4p_5(1-p_6)p_7(1-p_{11})$
15	0	0	1	1	1	1	0	$p_4p_5p_6(1-p_7)(1-p_{11})$
16	0	0	1	1	1	1	1	$p_4p_5p_6p_7(1-p_{11})$
17	0	1	0	0	0	0	0	$(1-p_4)(1-p_5)(1-p_6)(1-p_7)p_{00}$
18	0	1	0	0	1	0	0	$(1-p_4)p_5(1-p_6)(1-p_7)p_{00}$
19	0	1	0	1	0	0	0	$p_4(1-p_5)(1-p_6)(1-p_7)p_{10}$
20	0	1	0	1	1	0	0	$p_4p_5(1-p_6)(1-p_7)p_{11}$
21	1	1	1	0	0	0	1	$(1-p_4)(1-p_5)(1-p_6)p_7p_{00}$
22	1	1	1	0	0	1	0	$(1-p_4)(1-p_5)p_6(1-p_7)p_{00}$

Continued on next page

Table 13 – *Continued from previous page*

Vector	$\mathbf{x_1}$	$\mathbf{x_2}$	$\mathbf{x_3}$	$\mathbf{x_4}$	$\mathbf{x_5}$	$\mathbf{x_6}$	$\mathbf{x_7}$	Probability
23	1	1	1	0	0	1	1	$(1-p_4)(1-p_5)p_6p_7p_{00}$
24	1	1	1	0	1	0	1	$(1-p_4)p_5(1-p_6)p_7p_{01}$
25	1	1	1	0	1	1	0	$(1-p_4)p_5p_6(1-p_7)p_{01}$
26	1	1	1	0	1	1	1	$(1-p_4)p_5p_6p_7p_{01}$
27	1	1	1	1	0	0	1	$p_4(1-p_5)(1-p_6)p_7p_{10}$
28	1	1	1	1	0	1	0	$p_4(1-p_5)p_6(1-p_7)p_{10}$
29	1	1	1	1	0	1	1	$p_4(1-p_5)p_6p_7p_{10}$
30	1	1	1	1	1	0	1	$p_4p_5(1-p_6)p_7p_{11}$
31	1	1	1	1	1	1	0	$p_4p_5p_6(1-p_7)p_{11}$
32	1	1	1	1	1	1	1	$p_4p_5p_6p_7p_{11}$

For the 10 observed events (i.e. $N = 10$), there are many possible state vector combinations, as each event will invoke one of the 32 possible state vectors. We elicited these vectors by a *MATLAB* program (Hunt *et al.*, 2014) and counted the occurrence of each state vector for the given evidence. In this example, we obtained 27 possible counts vectors, $\mathbf{y_k}$, that imply the evidence; that is, there are 27 possible arrangements of state vectors that match the evidence coming from the sensors.

The likelihood function is constructed by summing up individual likelihoods defined by the multinomial distribution for each count vector, $\mathbf{y_k}$, as was formulated in Eq. 3.18. The posterior distributions for these failure probabilities are obtained by using uniform priors in *WinBUGS*.

The evidence set claims that, out of 10 system failures, sensor 2 only detected 2 failures, and sensor 3 did not detect any failures; therefore the probability of failure for node 4 and node 7 (p_4 and p_7) should be very small. This is confirmed by the MCMC output. Since we do not have any information about node 5, p_5 is around 0.5. The posterior failure probability of node 6 is very large because it is needed to compensate the low failure probability of node 7, for the series structure of their subsystem. More importantly, with the evidence set we are able to infer the reliability structure of the passive device subsystem (including nodes 2, 4 and 5). The conditional probabilities listed in Table 14 show that this subsystem has a high probability of failure when at least one of its components has failed. So, we can conclude that the reliability structure of the passive device subsystem is close to a series system.

Table 14. Empirical mean, standard deviation, and quantiles for parameters.

	Mean	**SD**	**2.5%**	**25%**	**50%**	**75%**	**97.5%**
p_4	0.184684	0.067958	0.05203	0.1333	0.1883	0.2409	0.2936
p_5	0.517456	0.076611	0.3156	0.4831	0.5398	0.5747	0.5978
p_6	0.917133	0.07627	0.7163	0.8821	0.9393	0.9745	0.9977
p_7	0.083205	0.07647	0.002211	0.0256	0.06082	0.118	0.2859
p_{00}	0.16214	0.06627	0.01738	0.1156	0.1773	0.2183	0.2471
p_{01}	0.885807	0.099819	0.6257	0.8343	0.914	0.9638	0.9968
p_{10}	0.774643	0.143001	0.517	0.656175	0.787	0.8993	0.9908
p_{11}	0.881724	0.072124	0.7572	0.8203	0.8849	0.9448	0.9945

The plots of prior and posterior distributions of these conditional probabilities are shown in Figure 11. From these plots we can see that, after combining evidence from sensors, the uniform prior evolves to a more narrowly distributed posterior. As a result, our method proves to be an effective way to assess dependencies in system reliability, even in the case of only a partial set of components being monitored.

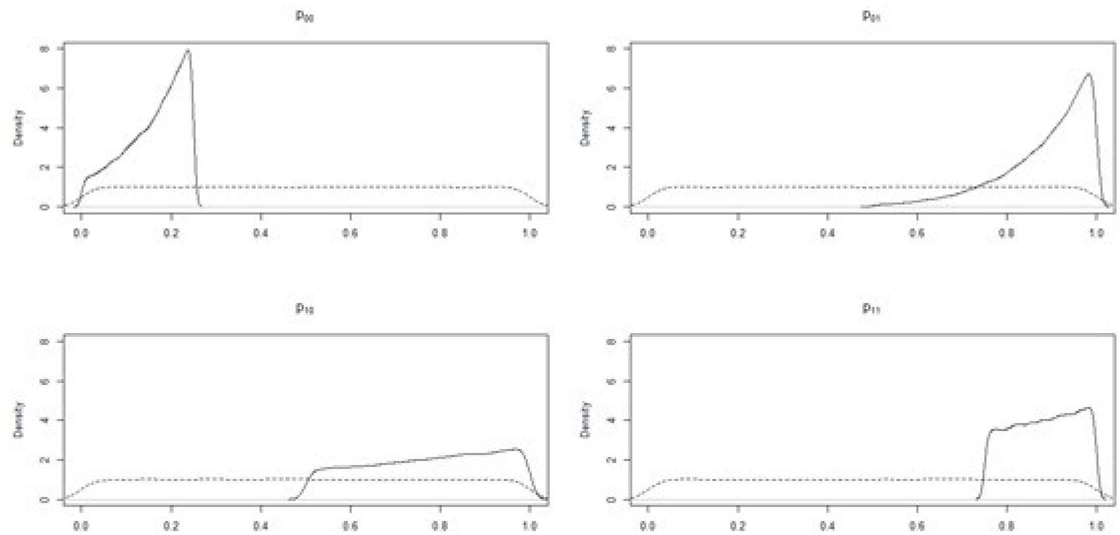

Figure 11. Prior (dashed lines) and posterior (solid lines) distributions of conditional probabilities.

3.4.1 Computational Complexity

As one can see from this case study, the computation complexity of our algorithm is not trivial. The evaluation of the likelihood function presented in this paper relies on identifying combinations of state vectors that are implied by the evidence. The speed of evaluation is largely dependent on the generation of possible state vectors for the system and then identifying all combinations of those state vectors. Once the state vector combinations are developed, the likelihood function can easily be calculated.

The generation of the set of combinations of state vectors is the most computationally intensive part for developing the likelihood function. We have developed an algorithm to rapidly identify these combinations. The first part of the algorithm, compiled in MATLAB, constructs all combinations of the count vectors for a given number of tests. The complexity of this part of the algorithm is $O(m^n)$, where m is the number of tests and n is the number of state vector combinations. Therefore, the number of count vectors increases exponentially with number of state vectors. As a result, the complexity of the algorithm is polynomial in the number of tests, but exponential in the number of state vector combinations. We admit that this might be problematic for very complex systems with hundreds of components.

We, however, suggest an alternative solution for combining state vectors. The matrix of count vectors is actually very sparse due to the fact that we do not observe all of the combinations. Therefore, it is not necessary to calculate all combinations. Note that the number of combinations of state vectors is related to number of ways distributing n identical objects among r groups and this can be done in $C(n + r - 1, r - 1)$ ways, where n is the number of counts of a specific vector combination observed and r is the number of possible combinations for unknown nodes. We have devised a formula that will rapidly give us the number of count vectors that satisfy the evidence and hence will provide rapid analysis of the likelihood function for subsequent Bayesian analysis.

For example, in the case study, according to the given evidence, we can specify what vectors are possible to be observed, so we do not need to combine all of the state vector combinations in our algorithm. Following the evidence, we infer that the state vectors must be eight $\{1, 1, 1, 0, x_5, 1, 0\}$'s ($n_1 = 8$) and two $\{1, 1, 1, 1, x_5, 1, 0\}$'s ($n_2 = 2$). As there are two possible states for the unobservable node x_5, $r_1 = r_2 = 2$. Therefore, total number of count vectors satisfying the evidence can be directly calculated as $C(9, 1) \times C(3, 1) = 27$. As a result, we can generate count vectors without going through all the possible combinations of state vectors.

Since probabilistic inference using BN is NP-hard (Cooper, 1990), we suggest designing efficient special-case algorithms, rather than using general probabilistic inference algorithms, for a specific problem. Stochastic simulation algorithms such as MCMC are very efficient, and they can be tuned to improve run times, especially in the incomplete data case.

3.5 Conclusion and Future Research

In this chapter we generalize the system reliability configuration of a complex system to a Bayesian network model. We are interested in exploring the relationship of system/subsystem reliability to its components. This research is particularly meaningful to a new system design where the system reliability configuration is uncertain. Using the Bayesian inference approach, we are able to combine information from multiple sources and multiple levels of the system to infer the conditional probabilities in BN.

Three data scenarios are discussed in this chapter. In a naïve scenario where the complete historical dataset of the states of the system and its components are available, we develop the conjugate Bayesian method for estimating the parameters in a binary state BN, and then extend it to a multi-state BN. When only failure records are available, we propose a method for quantifying the marginal distribution of system failure. Finally, we discuss the scenario of incomplete lower-level system information.

Data drawn simultaneously from the same system are fundamentally different from independent datasets. The dependencies between higher-level failure data and lower-level failure data are characterized by the conditional probabilities in a BN model. In the case of having incomplete lower-level data, the likelihood function of evidence becomes a summation of several likelihoods that correspond to all possible state vectors of the system. For such complicated function, it is impossible to find a closed form solution of posterior probability; therefore, we employed the computational Bayesian method, *MCMC*. The resulting method is successful at quantifying system reliability structure with incomplete data.

In this chapter, we studied simultaneous data analysis of binary-state systems. This research will be extended to Bayesian networks modeled by continuous life metric systems in the next chapter. Our proposed Bayesian network model can also be coupled with Hierarchical Bayesian (HB) inference to enable model parameter estimation without explicitly specifying its prior distribution. One concern is that, as the number of components and possible states increase, the exponentially increasing number of possible combinations of state vectors that comply with the observed evidence set will significantly worsen computational efficiency. In this research, we developed a *MATLAB* program to perform a rapid compilation of the set of combinations of state vectors to be used in the *MCMC* simulation in *WinBUGS*. However, a future research direction could be to develop more efficient algorithms that can handle multi-state systems and/or continuous state systems.

Furthermore, in the Bayesian inference of multi-level system, one may encounter the problem of the prior distribution of system reliability can be derived from two different channels. One is from the direct estimation on the system, such as expert opinions on the system reliability, and the other one is derived from component priors, because system reliability is a function of component reliability. Consequently, we need to combine the prior information from different channels. Guo (2011) used the Bayesian melding method originally proposed by Poole and Raftery (2000). In Chapter 4, we plan to incorporate Bayesian melding and other prior specification methods of system reliability into BN models.

Assessing the posterior distribution of conditional probabilities is critical to the understanding of both the functional and physical structure of a system. More research is needed on the techniques and tools for carrying out this activity. In our current study, we used *WinBUGS*, a tool for applying *MCMC* simulation in Bayesian inference. However, to reduce computational burden, other computational Bayesian methods should be investigated in future research.

Chapter 4

A BAYESIAN APPROACH TO SYSTEM RELIABILITY ASSESSMENT WITH INCOMPLETE HETEROGENEOUS DATA

4.1 Introduction

Estimating the reliability of complex systems has been a challenging problem as systems has grown more and more complex. Reliability engineers often have to deal with uncertain information in a complex environment, causing them to make decisions based on limited knowledge about the failure mechanisms of the system. Therefore, the statistical models used for representing complex systems should be mathematically robust, and at the same time easy to understand for reliability analysts. These models should be able to account for different sources of information, e.g., reliability tests, historical data, or expert judgments. These requirements have caused to a shift from traditional system reliability models, like fault trees and reliability block diagrams, to more flexible modeling frameworks, like Bayesian network (BN) (Wilson and Huzurbazar, 2007; Langseth and Portinale, 2007; Bobbio *et al.*, 2001).

In a complex system, even if many sensors have been deployed on various system levels for monitoring the health of the system and its components, it would be unrealistic to assume that the states of all components can be continuously observed. It is more common to have sensory data from some components/subsystems, but not all. In such case, a proper integration of multiple sources of information from different components or subsystems, as well as from expert opinions, for inferring the state of the system or some unobserved components becomes a crucial aspect for reliability assessment. This situation calls for a method to develop a reliability inference method that can combine simultaneous online information from various system levels for system and component reliability prediction.

BNs generalize fault trees by representing the relationship between components and subsystems by conditional probabilities instead of deterministic "AND" and "OR" gates, providing advantages when we are not sure about the reliability structure of a complex system, especially when there is uncertainty. In addition, BNs can accommodate different types of information, such as discrete, continuous or hybrid datasets. These uncertainties can be easily assessed using a BN, which would be rather difficult with conventional techniques, such as fault trees and reliability block diagrams since they are deterministic systems. However, parameter uncertainty of failure distributions of components has not received enough attention in the BN literature. This uncertainty of these parameters especially become apparent when there are some unobserved components in a complex system, and it is a challenging problem.

A great majority of this line of work considers the case of discrete Bayesian networks, i.e., networks that contain only discrete variables. Incorporation of discrete sources of data, such as pass/fail tests, from various levels in the system using Bayesian inference has been studied extensively (see Graves *et al.* (2007); Hamada *et al.* (2004); Johnson *et al.* (2003); Martz *et al.* (1988); Reese *et al.* (2005a). However, incorporating continuous data is more challenging, especially in the context of Bayesian inference, due to the integrals involved in calculations. Discrete networks are sometimes inadequate, since many important domains have continuous attributes as well as discrete ones. One can always discretize the continuous variables by partitioning their domain into some finite number of subsets, and transform the model to a discrete BN. However, this simple approach is often very problematic and might lead to poor performance. In our approach, we treat the continuous variables as continuous without trying to discretize them.

When a system is continuously monitored, the time at which the system or any of its components transitions from one state to another is a continuous random variable, thus the probability (or reliability) that they exist in a particular state is a function of time. Binary-state systems are those whose variables exist in either "failed" or "successful" states. The scope of this work is limited to binary-state systems. As failure times are observed, the likelihood function is a function of failure times. However, these observed failure data from different components and subsystems can be overlapping because in a coherent system they may represent the same event at different system levels. The probability of a component having failed at a given time is defined by the failure distribution parameters of that component.

This chapter presents a Bayesian network methodology for incorporating overlapping higher level data when making inferences about component reliability parameters associated with a time based reliability function. We develop a Bayesian model that accommodates lifetime information coming from some of the variables of a BN simultaneously. We show that our Bayesian network model can incorporate any parametric lifetime distribution for modeling the time-to-failure of the system components and can handle continuous variables without applying discretization.

An outline of this chapter is as follows. The next section presents a literature review of Bayesian network models with continuous and hybrid data structures. Our framework for system reliability and the inference method are discussed in Section 4.3. We start by discussing how to formulate the likelihood function with incomplete lifetime data, and then extend the case to the hybrid datasets where we also incorporate discrete pass/fail data into the likelihood formulation. We illustrate the proposed approach with an application to a missile guidance system in Section 4.4. Finally, we conclude the chapter with a discussion in Section 4.5.

4.2 Background

Bayesian networks have been used extensively in system reliability analysis due to their abilities in handling variables which are represented by a multivariate probability distribution (Bobbio *et al.*, 2001; Doguc and Ramirez-Marquez, 2009; Mahadevan *et al.*, 2001). Most of the research on Bayesian networks has focused on systems with discrete variables, or continuous variables with Gaussian distributions. Handling continuous variables have been a problematic issue for Bayesian networks due to the integrals involved in the likelihood calculations.

There are different types of Bayesian networks with respect to the type of their variables. These different types of BNs all require different analysis techniques because they all have different structures. In discrete state BNs, the state indicates whether the component works or fails, and it can be deducted from fault trees or reliability block diagrams. In this case, the variables of the BN are defined in discrete space and the BN is characterized by the conditional probability tables. Most of the research in the literature has focused on the discrete state BNs (Mahadevan *et al.*, 2001; Wilson and Huzurbazar, 2007). Exact inference in discrete state BNs is possible with some algorithms, such as variable elimination, belief propagation and junction trees (Heckerman, 1998; Koller and Friedman, 2009).

Continuous state BNs, on the other hand, assign a probabilistic distribution to the time-to-failure data of a component (Langseth and Portinale, 2007). The difference is that, in a continuous BN, the variables have a continuous state space. The state space represents the instant of time that the system component failed and covers the set of nonnegative real numbers. In literature, Hulting and Robinson (1994) extended the Martz *et al.* (1988) and Martz and Wailer (1990) methods to lifetime data. Like the binomial data method, Hulting and Robinson (1994) employed approximations in building up from component-reliability assessment to a system-reliability assessment. Boudali and Dugan (2005) presented a non-parametric discrete-time time-to-failure model, and Boudali and Dugan (2006) modeled a continuous-time time-to-failure in close-form without considering model uncertainty. Their continuous BN framework was able to capture the system components' behaviors and interactions, proposing a temporal Bayesian network reliability modeling and analysis method. However, it is still a challenging task to model the time-to-failure distribution because of the complexity of modeling a probability density in continuous space. Johnson *et al.* (2003) modeled the distribution parameters of time-to-failure as a continuous unknown variable, such as the scale and the shape of a 2-parameter Weibull density. This facilitates passing information through the network and the reliability analysis at system level based on the characteristics of the lifetime distributions of components. However, the integral in continuous state space makes the calculations intractable when the systems grow more complex.

Some researchers have proposed non-parametric methods for continuous Bayesian networks. Zhong *et al.* (2010) formulated the problem of system reliability assessment as a BN considering the parameter uncertainty. They modeled the time-to-failure of the system/components by the parametric distributions whose parameters are considered as random variables in the BN. For reasoning in a continuous BN, their method provided an alternative solution to the other methods, such as mixture of truncated exponentials, dynamic discretization and Markov chain Monte Carlo (MCMC). Warr and Collins (2014) also presented a hierarchical nonparametric framework, using Dirichlet processes, in which time-to-event distributions may be estimated from sample data or derived based on physical failure mechanisms. Their goal was to develop reliability estimates for complex systems, including estimates of uncertainty, using component, subsystem, and system data, and all available data types, which may include subjective data such as expert opinion as well as data collected from various formal tests.

Finally, hybrid-state BNs contain mixtures of discrete and continuous variables. Continuous and hybrid state BNs show similar characteristics and hybrid state BNs are also imposed to the same difficulties when it comes to computing posterior distributions. Previous research has suggested some discretization methods to perform the inference in the continuous/hybrid Bayesian networks (Langseth *et al.*, 2009; Neil *et al.*, 2007, 2008). Neil *et al.* (2008) have modeled time-to-failure distributions by continuous random variables as well as by discrete random variables. Marquez *et al.* (2010) showed how BN algorithms can be used to model time to failure distributions and performed reliability analysis of complex systems. Their hybrid BN approach extended fault trees by defining the time-to-failure of the fault tree constructs as deterministic functions of the corresponding input components' time-to-failure. Their approach incorporated an approximate inference algorithm for hybrid BNs, based on a process of dynamic discretization of the domain of all continuous variables in the BN. Iamsumang *et al.* (2015) also presented a hybrid BN-based methodology for component degradation modeling and efficient algorithm development with an application to online health monitoring of complex systems. They introduced a hybrid dynamic Bayesian network with component-based structure to represent complex engineering systems with underlying physics of failure by modeling an empirical degradation model with continuous variables. However, all these methods have not paid enough attention to the stochastic nature of parametric time-to-failure models in system reliability. They also have not considered simultaneous and incomplete data in their analyses.

Unfortunately, exact inference in continuous BNs with general distributions does not exist, especially when the data is incomplete, although the inference for the case where the distributions are Gaussians can be found in literature. Moral *et al.* (2001) described a theory for exact inference where distributions are specified as a mixture of truncated exponentials. However, at this point, in order to get the closed-form solution, one needs to go through multiple integrations. This process is very time consuming. Moreover, a closed-form solution can only be obtained if the integral is analytically solvable. An approximate solution has to be performed in the case where the closed-form solution cannot be explicitly derived.

All the aforementioned studies for continuous/hybrid BNs had good contributions. However, in highly complex systems, algorithms require large amount of computational time for inference in a continuous/hybrid BN. The computation time grows exponentially with each additional layer of network and becomes infeasible with a large number of nodes. As a result, for continuous BNs and hybrid BNs containing both discrete and continuous variables with non-Gaussian distributions, exact inference becomes computationally intractable (Boyen and Koller, 1998).

Markov chains have also been used for modeling continuous Bayesian networks in the literature (Boudali and Dugan, 2006). However, they present some limitations. Specifying a Markov chain for a large system becomes a cumbersome and tedious task. Markov chain modeling is limited to Markov processes, which generally requires all failure times to be exponentially distributed. Markov chains are also faced with the state space explosion problem; in fact, the number of states grows exponentially with the size of the system. Consequently, the number of differential equations to be solved grows exponentially with the size of the system. The state space explosion is one of the main limitations in using Markov chains for modeling large systems.

As a result, due to the limitations of Markov chains, discretization of the variables and assumption of Gaussian distributions, Bayesian researchers have focused on developing more efficient methods for incorporating continuous variables in their frameworks. For example, Wilson *et al.* (2006) showed how to combine reliability data that change over time, with an example that had binomial data at the system and one component, lifetime data at a second component, and degradation data at a third component. However, this paper did not demonstrate how to incorporate lifetime data at the system level. Guo (2011) proposed a model that considers lifetime data at every component. Their solution was to simply re-express system and subsystem lifetime distributions in terms of component lifetime distributions using deterministic relations derived from the system structure. However, their model was based on reliability block diagrams and they used independent and complete data in their analysis. Reese *et al.* (2011) presented a Bayesian model for assessing the reliability of multicomponent systems. In their model, lifetime data collected at the component, subsystem, or system level were integrated with prior information at any level. However, they also assumed that the test data are completely observed and independent from each other.

In literature, there have been very few studies developed for overlapping data at various levels of a system. Jackson (2011) developed an overlapping data likelihood function to incorporate inherent dependencies between the datasets and generate the correct inference within Bayes' theorem for systems. Their overlapping data Bayesian method incorporates all information and evidence that can possibly be generated or observed by complex time based systems represented by a fault tree. In this research, we focus on the problem of inference of the reliability model parameters in a BN in system reliability context using simultaneous and incomplete hybrid data.

4.3 Methodology

In this section Bayesian inference methods are discussed for simultaneously estimating parameters of lifetime distributions in a Bayesian network when lifetime and pass/fail data are collected from different levels of the system. We develop the likelihood function for simultaneous continuous and discrete data in a time based system represented by a Bayesian network structure, thus presenting a generalization to the standard series and parallel systems. We assume the structure of the network is given and the conditional probabilities of the network are known.

4.3.1 Bayesian Networks in Reliability Assessment

In recent years, Bayesian networks (BNs) have been increasingly used in a wide range of applications including computer science, bioinformatics, data fusion, decision support systems and others. A Bayesian network is a directed acyclic graph (DAG) that represents a joint probability distribution among a set of variables, where the nodes denote random variables and the arcs between these variables denote the conditional dependencies (represented by conditional probability distributions) among variables (Koller and Friedman, 2009).

A Bayesian network model allows for efficient calculation of belief revisions, i.e. calculating the quantitative belief changes of variables when new evidence is observed. This is very useful for diagnostic and prediction purposes in decision support environments, such as reliability assessment domains. An example of a BN can be seen in Figure 12.

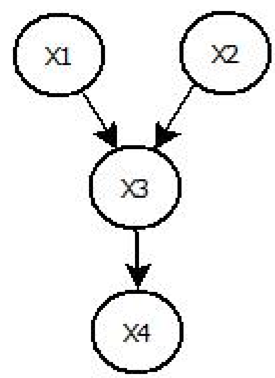

Figure 12. A sample Bayesian network

In a BN, the nodes without any arrows directed into them are called root nodes (also called parent nodes) and they are described according to their marginal probability distributions (nodes X_1 and X_2 in Figure 12). The nodes that have arrows directed into them are called child nodes. Each child has a conditional probability distribution associated with it, given the values of parent nodes.

Nodes in a Bayesian network are said to be uncorrelated if there is no arc between these nodes. This situation is called conditional independence. The conditional independence structure reduces significantly the complexity of inference and allow to decompose the underlying joint probability distribution as a product of local conditional probability distributions (CPDs) associated to each node and its respective parents (Spiegelhalter and Lauritzen, 1990). If the variables are discrete, the CPDs can be represented by conditional probability tables (CPTs), which list the probability that the child node takes on each of its different values for each combination of values of its parents.

Let G be the BN in Figure 12 with nodes X_1, X_2, X_3, X_4. The joint distribution of G over the variables can be written as

$$P(X_1, X_2, X_3, X_4) = \prod_{i=1}^{4} P(X_i \mid pa(X_i)) = P(X_1)P(X_2)P(X_3 \mid X_1, X_2)P(X_4 \mid X_3) \tag{4.1}$$

where the multiplication is replaced by an integral in case of continuous variables.

The simplest of Bayesian networks are binary-state BNs, where components are either in the "functional" or "failed" states. "Multi-state" BNs involve components that can be classified by order of severity in various degraded states ranging from "functional" to "failed". However, restricting our attention to models containing only discrete variables seems very unsatisfactory in the domain of reliability analysis. Bayesian networks based on continuous data are those whose failure probability is a function of a time variable, which is our main focus in this chapter.

4.3.2 Integrating Incomplete Lifetime Data Using Bayesian Inference

Learning Bayesian networks from data has drawn lots of attention by researchers in order to be able to apply BNs to real-world applications. Learning from complete data has been studied extensively in the last decade. If the data is complete, or fully observed, so that each of the network variables is observed, learning BN parameters is not difficult, however, in many applications, data can be incomplete for various reasons. In the complete data case, we can use binomial likelihood and beta priors and obtain a closed form solution for the distribution of the parameters (Martz *et al.*, 1988; Johnson *et al.*, 2003; Hamada *et al.*, 2004). If the data is complete, the learning problem reduces to a set of local learning problems, one for each variable (Koller and Friedman, 2009). However, in reality one frequently has to deal with incomplete data. The problem gets more complicated for the incomplete data case, where the variables are partially observed, so that, in each instance, some variables (known in advance) are not observed in the Bayesian network. Learning Bayesian networks from incomplete data is a very difficult problem. The occurrence of missing values leads to analytical intractability and high computational complexity compared to the complete data scenario. The existing methods either use inference algorithms to get the expected values of statistics or delete the missing values. Approaches like the expectation-maximization might get stuck at local optima (Lauritzen, 1995). In our methodology, we prefer to use Bayesian inference for parameter learning, as it is a powerful tool when used with probabilistic graphical models such Bayesian networks. A commonly adopted technique for applying Bayesian inference is Markov Chain Monte Carlo (MCMC) methods, for their efficiency in sampling from the joint probability distribution of the model (Gelman *et al.*, 2014). To implement MCMC, we use WinBUGS, a statistical software for Bayesian inference (Spiegelhalter *et al.*, 2003). WinBUGS is a general purpose modeling language, which takes as its input a BN model and returns samples that can be used for estimating the posterior probability distributions of the model parameters. In this paper, we develop a method that is statistically valid, and correctly reflects the increased uncertainty due to missing data. We also demonstrate that the MCMC method can learn Bayesian networks from incomplete data efficiently.

In order to make inferences about the failure parameters of the components, we need to update the prior beliefs as in Bayes' theorem such that

$$\pi_1(\boldsymbol{\theta} \mid E) = \frac{L(E \mid \boldsymbol{\theta})\pi_0(\boldsymbol{\theta})}{\int_{\forall \boldsymbol{\theta}} L(E \mid \boldsymbol{\theta}')\pi_0(\boldsymbol{\theta}')d\boldsymbol{\theta}'} \tag{4.2}$$

where $L(E \mid \boldsymbol{\theta})$ is the likelihood of observing evidence set E for given parameter set $\boldsymbol{\theta}$. The challenge here is how to formulate the likelihood function due to the fact that the collected data at different system levels are overlapping, causing dependency among them. Therefore, the likelihood function cannot be a multiplication of separate likelihoods coming from different nodes. A special consideration is needed for the formulation of the total likelihood function.

In our system representation, we adopt the following convention. Given a component, we represent the state of the component with either one of the two states: 1 for failed component, and 0 for working component. The quantification of the Bayesian network requires the assignment of a probability value to each node. Since the computation is performed according to a given time t, the failure probabilities of the components at time t should be provided. For discrete systems, the failure probability is a parameter of Binomial distribution. For systems and components dealing with continuous data, on the other hand, the probability of being in a "failed" or "working" state is a function of time. For example, the probability of a component being in the "failed" state is $Pr(C = 1 = failed) = Pr(C = 1, t) = F_C(t)$ and the probability of the component being in the working state is $Pr(C = 0 = working) = Pr(C = 0, t) = 1 - F_C(t) = R_C(t)$, where $F_C(t)$ and $R_C(t)$ are the cumulative distribution and reliability functions of that component, respectively.

We next define our system reliability representation as follows. Our system is represented as a multi-level directed acyclic graph (DAG) which contains a hierarchical structure. The nodes of a multi-level DAG can be partitioned into levels $L_1, \ldots, L_m$, such that there is no edge within a level and all the edges are between nodes in level

L_i and the nodes in the adjacent levels L_{i-1} and L_{i+1} (see Figure 13 for an illustration). We call the nodes in the adjacent lower level L_{i+1} as the direct subordinates of the nodes in level L_i.

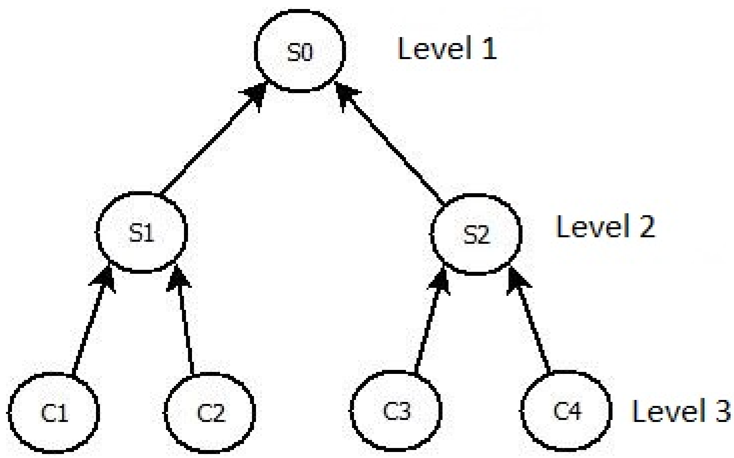

Figure 13. BN representation of a hierarchical system.

As we develop the model, we will use the following notation. The components, subsystems, and system in the BN are referred to as nodes, such as C_i and S_j in Figure 13. The components are denoted by C_i, and subsystems and the system are denoted by S_j. The direct subordinates of S_j are the nodes in the next lower level, which constitute node S_j. The set of direct subordinates of S_j is denoted as A_j. In Figure 13, for instance, system S_0 has a direct subordinates set $A_0 = (S_1, S_2)$. The evidence set E contains the simultaneous lifetime information collected at several nodes in the BN.

In our Bayesian framework, we represent data and their information by likelihood contributions. We would like to assess system reliability as a function of time; therefore, we need to formulate the probability of failure as a function of time. The probability of a component having failed at a given time t is equal to the cumulative distribution function (CDF), $F(t)$. The CDF is defined by a set of parameters, which for the i^{th} component is represented as θ_i. The set of component parameters for the system is:

$$\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_n\}$$

Our primary goal is to assess these failure parameters in order to monitor reliability of the system and its components through time. The probability of the i^{th} component having failed at a given time t is defined by the set of reliability parameters of that component. Thus, the failure probability of component C_i at time t can be calculated by

$$p_{C_i}(t \mid \theta_{C_i}) = F_{C_i}(t \mid \theta_{C_i}) \tag{4.3}$$

For simplicity, we will denote the component i failure probability as $p_{C_i}(t)$. The failure probability is a function of time, as opposed to the discrete case, where we model discrete data as multinomial likelihood. In order to formulate the likelihood function, we also need to model the lifetime distribution of each component, C_i, which we will denote $f_{C_i}(t \mid \theta_{C_i})$, where $f(\cdot)$ is the probability density function and represents the probability that the i^{th} component will fail at time t. However, for the system node and subsystem nodes (S_j), we need to infer the probability density function using the relationship of components to the system/subsystem represented by the Bayesian network, which will not result in a standard distribution. Therefore, the probability density function of the system is calculated by taking the derivative of the distribution function, such that

$$f_{S_j}(t \mid \theta_{S_j}) = \frac{d}{dt}F_{S_j}(t \mid \theta_{S_j}) = \frac{d}{dt}(1 - R_{S_j}(t \mid \theta_{S_j})) = -\frac{d}{dt}R_{S_j}(t \mid \theta_{S_j}) \tag{4.4}$$

where the reliability function, $R_{S_j}(t \mid \theta_{S_j})$, is calculated by using the relationship between components and subsystem using the BN conditional dependence structure:

$$R_{S_j}(t \mid \theta_{S_j}) = 1 - F_{S_j}(t \mid \theta_{S_j}) = \psi_{S_j}(R_{S_k}, R_{C_k} : \forall S_k \in A_j, \forall C_k \in A_j) \tag{4.5}$$

where $\psi_{S_j}(\cdot)$ is the structure function of node S_j defined by the BN structure (conditional probabilities), which describes the reliability relationship between the node and its direct subordinates. S_k, and C_k are, respectively, the subordinate subsystems, and components of node S_j within the direct subordinates set A_j. Differently from fault trees and reliability block diagrams, the dependence relations among variables in a BN are not restricted to be deterministic. This corresponds to being able to model uncertainty in the interaction between components, by suitably specifying conditional probabilities, instead of using deterministic "AND" and "OR" gates. Probabilistic gates may reflect an imperfect knowledge of the system behavior, also helping us quantify the dependency structure among the components (see Bobbio *et al.* (2001)). As an example, for a 2-component system, the reliability of the system can be formulated such that

$$\begin{aligned}R_{sys}(t) =&\psi_{sys}(R_1(t), R_2(t))\\ =&p_{00}R_1(t)R_2(t) + p_{01}R_1(t)(1 - R_2(t)) + p_{10}(1 - R_1(t))R_2(t)\\ &+ p_{11}(1 - R_1(t))(1 - R_2(t))\end{aligned}$$

where the conditional failure probabilities are defined as $p_{ij} = Pr(C_{sys} = 1 \mid C_1 = i, C_2 = j)$ $(i, j = 0$ or $1)$. In our work, we make the assumption that these conditional probabilities are known in advance.

As a result, our BN framework is a generalization of the series and parallel system structures, where the conditional probabilities are either 0 or 1. Let us for instance consider the problem of calculating the reliability of a parallel system of three components. The components have life-lengths T_1, T_2 and T_3 respectively, and the system's life-length is thus given as $T_{sys} = max(T_1, T_2, T_3)$. However, if the system is connected in series, then $T_{sys} = min(T_1, T_2, T_3)$. Since Bayesian network is a stochastic system, we do not have a deterministic relationship for the lifetime of the system.

The system lifetime largely depends on the conditional probabilities between the components and the system.

We would like to emphasize the importance of using simultaneous data in our framework due to the dependencies inherent in a Bayesian network structure. If we would like to learn about these dependencies, we need to avoid using independent data, which will make it impossible to quantify the relationships between components of a complex system. Therefore, in this work, we only use simultaneous data, which means all observations come from the same system such that they are dependent to each other. In case of simultaneous and incomplete data, the likelihood of observed data is not a simple multiplication of likelihoods of the nodes anymore, so we cannot apply Eq. 4.2 easily. We, therefore, develop a method to formulate the likelihood function for the BN system by using conditional independencies implied by the network structure. In the next section, we explain the concept of d-separation and how we use this concept in the formulation of the likelihood function.

4.3.3 D-Separation in Bayesian Networks

Probabilistic graphical models such as Bayesian networks are efficient in portraying conditional independencies and causal relations, and the criterion called d-separation can be used to read them off the graph (Pearl, 2014). Since every d-separation in the graph implies conditional independence in the distribution, using this criterion in our framework proves very efficient in our overlapping data methodology.

To better understand the nature of overlapping data sets, the concept of "d-separation in Bayesian networks" is introduced. The main idea is that each observed variable constitutes a subset of variables that is d-separated from the rest of the variables in the Bayesian network. In effect, d-separation helps us generate separate likelihood functions given each evidence, and then we can generate an overall likelihood function. Thus, our aim in this section is to understand when we can guarantee that an independence holds in a distribution associated with a BN structure.

Definition 4.1 *D-separation (see Pearl (2014); Koller and Friedman (2009)) A path p is said to be d-separated by a set of nodes Z if and only if:*

1. *p contains a chain $i \rightarrow m \rightarrow j$ or a fork $i \leftarrow m \rightarrow j$ such that the middle node m is in Z, or*
2. *p contains an inverted fork $i \rightarrow m \leftarrow j$ such that the middle node m is not in Z and such that no descendant of m is in Z.*

As a result, when influence can flow from a node to another node thorough **Z**, we say that the trail between those two nodes is active. Due to the hierarchical structure in the reliability representation of our BN system, we only have two types of trails: chain trail (also called causal trail) ($i \rightarrow m \rightarrow j$) and inverted fork trail (also called common effect trail) ($i \rightarrow m \leftarrow j$). Looking back at Figure 13, we can see the causal trails: $C_1 \rightarrow S_1 \rightarrow S_0$, $C_2 \rightarrow S_1 \rightarrow S_0$, $C_3 \rightarrow S_2 \rightarrow S_0$ and $C_4 \rightarrow S_2 \rightarrow S_0$. The common effect trails are: $C_1 \rightarrow S_1 \leftarrow C_2$, $C_3 \rightarrow S_2 \leftarrow C_4$ and $S_1 \rightarrow S_0 \leftarrow S_2$.

We will use the concepts of "active trails" and "d-separation" to formulate conditionally independent likelihood functions. A causal trail ($i \rightarrow m \rightarrow j$) is active if and only if m is not observed. This means that every time we observe a variable, it will block the path of influence between the upstream and downstream nodes. A common effect trail ($i \rightarrow m \leftarrow j$) is actived if m is observed. This structure is also called a v-structure (Koller and Friedman, 2009). As a result, every observed variable activates the v-structure (which consists of the components attached to that variable) and cuts off the path of influence from other variables, thus creating a region of influence conditionally independent of the rest of the network given the observed variable. For example, if we observe variable S_1 in Figure 14, then

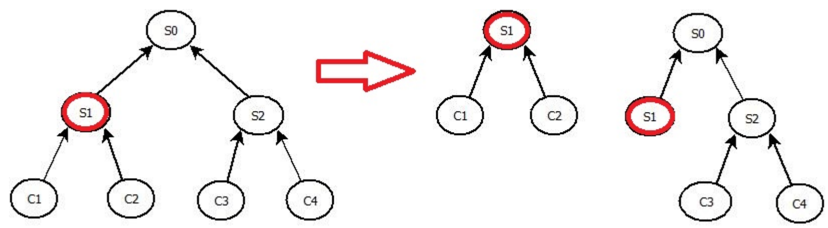

Figure 14. Two conditionally independent sub-systems given S_1.

As seen in the figure above, evidence in variable S_1 breaks the BN into two conditionally independent subsystems. Nodes C_1 and C_2 are d-separated from the rest of the network given evidence about S_1; however, they are not d-separated from each other, so they belong to the same sub-system. In our research, it is useful to view probabilistic influence as a flow in the graph. One node can influence another if there is any trail along which influence can flow. As a result, d-separation provides us with a notion of separation between nodes in a directed graph (hence the term d-separation, for directed separation). As a result, the set of independencies derived from d-separation is a complete characterization of the independence properties that are implied by the network structure.

4.3.4 Formulation of the Likelihood Function for Incomplete Lifetime Data

The likelihood function plays a central role in Bayesian learning. Our approach addresses how to parametrically model the multilevel system structure to preserve the probabilistic constructs defined by the BN, and to coherently combine the simultaneous data sets through the derivation of their joint likelihood function. A descriptive flowchart of the proposed Bayesian approach is given in Fig. 15.

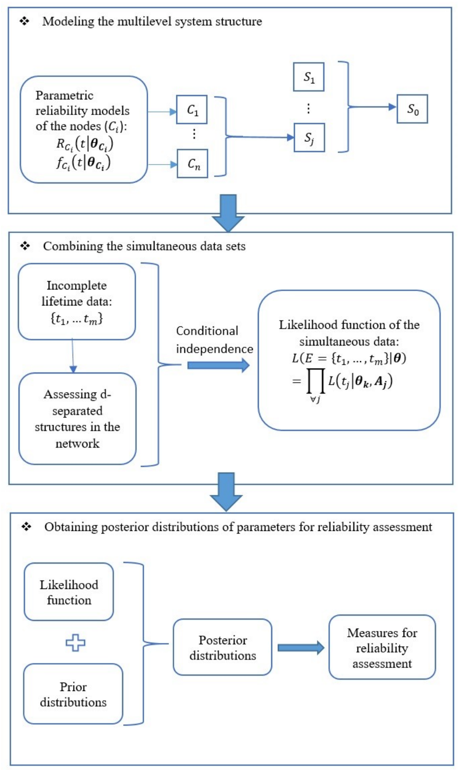

Figure 15. Descriptive flowchart of the proposed approach.

The first framework is a substitution strategy for modeling the multilevel system structure. It is carried out by re-expressing the reliability function and distribution function of high level node in terms of the corresponding functions of its direct subordinates, which are contained in set A_j as explained in Section 4.2 (see Eq. 4.5). The structure function ψ_{S_j} derived from the BN is used to construct the inherent functional relationship. The second framework is a combining strategy for integrating the overlapping data sets. It is implemented by formulating the likelihood function based on d-seperation. These likelihood contributions are developed according to the evidence and parametric models of the nodes. The third framework is a Bayesian inference strategy for information integration. The Bayesian model is constructed by deriving the posterior distribution of model parameters using the joint likelihood function and specified prior distributions. After the joint posterior distribution of model parameters are obtained, some reliability measures such as the failure rate and predicted reliability are generated by averaging over the posterior distribution of related model parameters.

In our model framework, lifetime data collected at individual component and lifetime data collected at the system/subsystem level are incorporated. The data collected at the higher level provide both direct information both about the system (or subsystem) at which it was collected, and also partial information about the components that comprise the system (or subsystem). As depicted in Figure 15, the multilevel system structure is modeled based on parametric models of components C_i, $i = 1, \ldots, n$. As explained in Section 4.2, we use structure functions (ψ_j) as a substitution strategy for modeling the high level nodes (S_j). This substitution is implemented by formulating the reliability function of S_j with reliability functions of its direct subordinates, that is, the nodes in A_j. The PDF, $f_{C_i}(t \mid \theta_{C_i})$, and reliability function, $R_{C_i}(t \mid \theta_{C_i})$, of the higher level node S_j are expressed as shown in Eq. 4.4 and 4.5, respectively.

Since calculating the PDF of a higher level node requires derivation, one needs a standard and efficient way to calculate this derivation. Note that the reliability function $R_{C_i}(t \mid \theta_{C_i})$ is only a function of nodes in the next lower level, which are composed of at most three elements: the observed component/subsystem, the unobserved subsystem and the unobserved component in the immediate subordinate set A_j. Next, the PDF, $f_{C_i}(t \mid \theta_{C_i})$, is calculated by taking the negative derivative of the reliability function as in Eq. 4.4. For simplicity, we can use the chain rule of calculus, by splitting the derivation in 3 different parts and Eq. 4.4 becomes

$$\begin{aligned} f_{S_j}(t \mid \boldsymbol{\theta}_{\mathbf{S_j}}) &= -\frac{d}{dt} R_{S_j}(t \mid \boldsymbol{\theta}_{\mathbf{S_j}}) \\ &= -\sum_{\substack{\forall O_i \in E \\ \forall O_i \in A_j}} \frac{\partial R_{S_j}(t)}{\partial R_{O_i}(t)} \times \frac{dR_{O_i}(t)}{dt} - \sum_{\substack{\forall S_k \in Aj \\ \forall S_k \notin E}} \frac{\partial R_{S_j}(t)}{\partial R_{S_k}(t)} \times \frac{dR_{S_k}(t)}{dt} \\ &\quad - \sum_{\substack{\forall C_l \in Aj \\ \forall C_l \notin E}} \frac{\partial R_{S_j}(t)}{\partial R_{C_l}(t)} \times \frac{dR_{C_l}(t)}{dt} \end{aligned} \tag{4.6}$$

where subscript O_i belongs to the observed variables ($\forall O_i \in E$), S_k belongs to unobserved subsystems in the direct subordinates set ($\forall S_k \in Aj, \forall S_k \notin E$), and C_k belongs to unobserved components in the direct subordinates set ($\forall C_l \in Aj, \forall C_l \notin E$).

When evidence data is introduced, the d-separated portions of the Bayesian network structure are assessed as previously discussed in Section 4.3. To capture the temporal dependencies found in the Bayesian network model, we will use a different form of a special function called the unit step function (also called the Heaviside unit-step function). In literature, unit-step and impulse functions have been used to represent evidence (see Boudali and Dugan (2006); Jackson (2011). When the evidence is observed, it changes the form of the distribution function of the corresponding observed variable, because we know that the failure time is equal to that instant. Since we are working with reliability functions instead of cumulative distribution functions in this work, we take a different approach and represent the reliability function using the unit-step function. CDF of the time to failure of that component becomes the unit step function, such that $H(t) = 1$ when $t \geq 0$ and $H(t) = 0$ otherwise. Since $R(t) = 1 - F(t)$, we can formulate the unit-step function as the reliability function when there is evidence such that $H(t_F - t) = 1$ when $t \leq t_F$ and $H(t_F - t) = 0$ otherwise, where t_F is the time failure was observed in the evidence set. As a result, the reliability function will make a jump from 1 to 0 at the specific time the component failed, and we can use this unit-step function directly in the likelihood, modifying Eq. 4.6 to

$$f_{S_j}(t \mid \boldsymbol{\theta}_{\mathbf{S_j}}) = -\sum_{\substack{\forall O_i \in E \\ \forall O_i \in A_j}} \frac{\partial R_{S_j}(t)}{\partial R_{O_i}(t)} \times \frac{dH(t_{O_i} - t)}{dt} - \sum_{\substack{\forall S_k \in Aj \\ \forall S_k \notin E}} \frac{\partial R_{S_j}(t)}{\partial R_{S_k}(t)} \times \frac{dR_{S_k}(t)}{dt}$$
$$- \sum_{\substack{\forall C_l \in Aj \\ \forall C_l \notin E}} \frac{\partial R_{S_j}(t)}{\partial R_{C_l}(t)} \times \frac{dR_{C_l}(t)}{dt} \tag{4.7}$$

This modification introduces steps into the reliability function of components based on the evidence observed. As Eq. 4.7 is the PDF of time to failure of an observed node, it is equivalent to the likelihood function for each observation given observations and reliability parameters from the next lower level nodes, that is, the set of nodes in A_j, and each observed variable constitutes a subset of variables that is d-separated from the rest of the Bayesian network, therefore

$$\begin{aligned}
L_{S_j}(t_j \mid \boldsymbol{\theta}_{\mathbf{j}}, \mathbf{A_j}) =& f_{S_j}(t \mid \boldsymbol{\theta}_{\mathbf{S_j}}) = - \left.\frac{dR_{S_j}(t)}{dt}\right|_{t=t_j} \\
=& -\sum_{\substack{\forall O_i \in E \\ \forall O_i \in A_j}} \frac{\partial R_{S_j}(t)}{\partial R_{O_i}(t)} \times \left.\frac{dH(t_{O_i} - t)}{dt}\right|_{t=t_j} - \sum_{\substack{\forall S_k \in Aj \\ \forall S_k \notin E}} \frac{\partial R_{S_j}(t)}{\partial R_{S_k}(t)} \times \left.\frac{dR_{S_k}(t)}{dt}\right|_{t=t_j} \\
& - \sum_{\substack{\forall C_l \in Aj \\ \forall C_l \notin E}} \frac{\partial R_{S_j}(t)}{\partial R_{C_l}(t)} \times \left.\frac{dR_{C_l}(t)}{dt}\right|_{t=t_j} \\
=& -\sum_{\substack{\forall O_i \in E \\ \forall O_i \in A_j}} \frac{\partial R_{S_j}(t)}{\partial R_{O_i}(t)} \times \left.\frac{dH(t_{O_i} - t)}{dt}\right|_{t=t_j} - \sum_{\substack{\forall S_k \in Aj \\ \forall S_k \notin E}} \frac{\partial R_{S_j}(t)}{\partial R_{S_k}(t)} \times f_{S_k}(t_j \mid \boldsymbol{\theta}_{\mathbf{S_k}}) \\
& - \sum_{\substack{\forall C_l \in Aj \\ \forall C_l \notin E}} \frac{\partial R_{S_j}(t)}{\partial R_{C_l}(t)} \times f_{C_l}(t_j \mid \boldsymbol{\theta}_{\mathbf{C_l}})
\end{aligned} \tag{4.8}$$

where $f_{S_k}(t_j \mid \boldsymbol{\theta}_{\mathbf{S_k}})$ is calculated as in Eq. 4.4, causing a recursive structure. Note that we have as many conditionally independent likelihoods as the number of observed variables. After identifying the d-separated network with respect to the evidence, we need to start from the last level of the network (the component level), and go to the next upper level, this enabling us to use the recursive structure. If the observed variable is a component, we can use the pdf of the component directly, instead of Eq. 4.8.

Suppose m nodes are observed in a Bayesian network. The likelihood of observing the failure times of the observed components, $E = \{t_1, \ldots, t_m\}$ given the set of parameters that define the failure distributions of the components $\theta = \{\theta_1, \ldots, \theta_n\}$, is defined below. The likelihood functions derived from the data set of each observed component can be multiplied as they have been isolated into conditionally independent sets of likelihoods.

$$L(E \mid \boldsymbol{\theta}) = L(\{t_1, \ldots, t_m\} \mid \{\theta_1, \ldots, \theta_n\}) = \prod_{k=1}^{m} L_k(t_k \mid \boldsymbol{\theta_k}, \mathbf{A_k}) \tag{4.9}$$

According to Bayesian inference, given the prior distribution of model parameters, the posterior distributions can be obtained by

$$p(\boldsymbol{\theta} \mid E) \propto L(E \mid \boldsymbol{\theta}) \times p(\boldsymbol{\theta}) \tag{4.10}$$

where $p(\boldsymbol{\theta})$ is the joint prior distribution for system model parameters and $p(\boldsymbol{\theta} \mid E)$ is the joint posterior distribution of model parameters $\boldsymbol{\theta}$. The analysis of reliability assessment and prediction will be based on this posterior distribution. After obtaining the posterior distributions, we can make some assessments on the system reliability. For this task, Peng *et al.* (2013) suggested calculating some measures, such as the failure rate of the system and reliability as a function of mission time. Therefore, based on the system reliability function $R_{S_0}(t)$ and the joint posterior distribution of the parameters θ, the failure rate of the system at time t can be obtained from

$$\lambda_{S_0}(t \mid E) = \int_{\Theta} \frac{f_{S_0}(t \mid \boldsymbol{\theta})}{R_{S_0}(t \mid \boldsymbol{\theta})} p(\boldsymbol{\theta} \mid E) d\boldsymbol{\theta} \tag{4.11}$$

where E denotes the available simultaneous data. $f_{S_0}(t \mid \boldsymbol{\theta})$,$R_{S_0}(t \mid \boldsymbol{\theta})$ and $\lambda_{S_0}(t \mid E)$ are separately the PDF, reliability function, and failure rate of the system.

Given that the system has survived up to the present time t_p, the probability that the system will survive another interval of mission time Δt can be calculated by

$$R(t_p + \Delta t \mid t_p, E) = \int_\Theta \frac{R(\Delta t + t_p \mid \boldsymbol{\theta})}{R(t_p \mid \boldsymbol{\theta})} p(\boldsymbol{\theta} \mid E) d\boldsymbol{\theta} \tag{4.12}$$

Similar to the joint posterior distribution of parameters, Eq. 4.11 and 4.12 cannot be specified analytically. The MCMC is used to collect samples from these distributions. By substituting the generated posterior samples into the corresponding PDF and reliability functions above, samples for these reliability measures are obtained. Summary statistics can be easily obtained based on these random samples. For instance, the integrations above are approximated by the mean of relative samples. Moreover, the variances and confidence intervals for these measures can be obtained within this Bayesian framework as well.

When a new system is running, it is necessary to predict system reliability at future time points. Such predictions are usually adopted to set strategies for system operation and warranty. Therefore, the reliability as a function of mission time is obtained from

$$R(t \mid E) = \int_\Theta R(t \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid E) d\boldsymbol{\theta} \tag{4.13}$$

Similar to Eq. 4.11 and 4.12, Eq. 4.13 above have no analytical forms. The calculations are based on the posterior samples of model parameters using simulation based integration.

4.3.5 Integrating Incomplete Hybrid Data Structures by Bayesian Inference

Restricting our attention to models containing only discrete or only continuous variables might be very unrealistic in real applications. Therefore, we need to also consider Bayesian inference with overlapping hybrid data structures.

The proposed framework with hybrid data structure can be summarized as follows: Our BN model combines high-level system functionality data with low-level component failure time data. System node indicates whether the system is working as intended. We observe a system's functionality and there are sensors on some (not all) of the components of the system. When we observe if a system is working or failed, we analyze the components. The sensors on the components record the failure times of these components. As a result, we have discrete data from the system, and continuous life time data from the components.

System reliability problems typically have two types of information, component tests and system tests. However, in the literature, these component tests and system tests are modeled separately because they are independent tests. In this research, we seek a model which provides flexibility for incorporating both types of information coming from the same test, making the data simultaneous. As stated previously, dealing with simultaneous data is not a trivial task, and integrating data and prior information at different levels within a BN has often proven problematic from both the perspectives of computational tractability and model consistency.

We can consider the pass/fail data coming from the system as censored observations. When we observe a system has failed, we do not know the exact failure time, but we know that the system failed within that time period. So we can consider it as a left-censored observation. On the other hand, when we observe that the system is still functioning, we know that it has not failed until that time, so we can consider it as a right-censored observation. The contribution of a right-censored observation to the likelihood function is the reliability function, $(1-F(t))$, evaluated at the censored value at the appropriate level in the Bayesian network; whereas the contribution of a left-censored value observation is $F(t)$, the cumulative distribution function. Incorporating censored data into our model framework is thus straightforward and can be accomplished by simply substituting the appropriate expression for the censored observation for the system in Eq.4.9.

As a result, if we observe that the system has failed, then the likelihood of the corresponding observation is

$$L_{S_0}(t_{S_0} \mid \boldsymbol{\theta}, \mathbf{A_0}) = p_{S_0}|_{t=t_{S_0}}$$

where p_{S_0} is a function of conditional probabilities and failure distributions of the components in the next lower level (A_0), just as explained in Section 4.3.4. If we observe that the system is still functioning at a specific mission time, then the likelihood of the corresponding observation is

$$L_{S_0}(t_{S_0} \mid \boldsymbol{\theta}, \mathbf{A_0}) = 1 - p_{S_0}|_{t=t_{S_0}}$$

In the continuous data case, we use the chain rule of calculus, by splitting the derivation in 3 different parts, which makes it easier to incorporate evidence from the lower level. However, in this case, since we only need the failure function, not the probability density function, we do not need to calculate derivatives as in Eq. 19. Therefore, we do not need to take the derivative of the unit step function, $H(t_{O_i}-t)$, which represents the jump in reliability function of an observed variable. The key point here is that we only need to substitute 1 for $H(t_{O_i}-t)$ when $t \leq t_{O_i}$ and 0 otherwise for observed variables while calculating p_{S_0} from the Bayesian network structure.

4.4 Illustrative Example

Reliability assessment and prediction for missiles in a guidance system is carried out in this section to demonstrate the proposed Bayesian network approach, which was also studied by Jackson (2011) as a fault tree. A missile has a guidance system to allow it to steer and change course towards its intended target, and also a propulsion system that self-drives it. The missile's flight path can be guided by use of guidance information transmitted from the control point via. As a result, guidance systems improve the performance of the missile, which is the missile accuracy. Over the years, more and more sophisticated systems have been developed to implement guidance control rules. Accordingly, operation and management of a guided missile system requires precise assessment and prediction of the system reliability using available data and information.

4.4.1 The Guided Missile System Structure

Every missile guidance system consists of an attitude control system and a flight path control system. The attitude control system functions to maintain the missile in the desired attitude on the ordered flight path by controlling the missile in pitch. The attitude control system operates as an auto-pilot, damping out fluctuations that tend to deflect the missile from its ordered flight path. The function of the flight path control system is to determine the flight path necessary for target interception and to generate the orders to the attitude control system to maintain that path. The reliability block diagram (RBD) of a simplified system structure is depicted in Figure 16, where subsystem S_1 represents flight-path control, subsystem S_2 represents attitude control and component C_6 represents the power supply. The system consists of two subsystems and a component: with S_1 and S_2 being parallel structures, and C_6 being a component connected in series to subsystems S_1 and S_2. Note that S_3 is a series structure and is one of the parallel components of subsystem S_1.

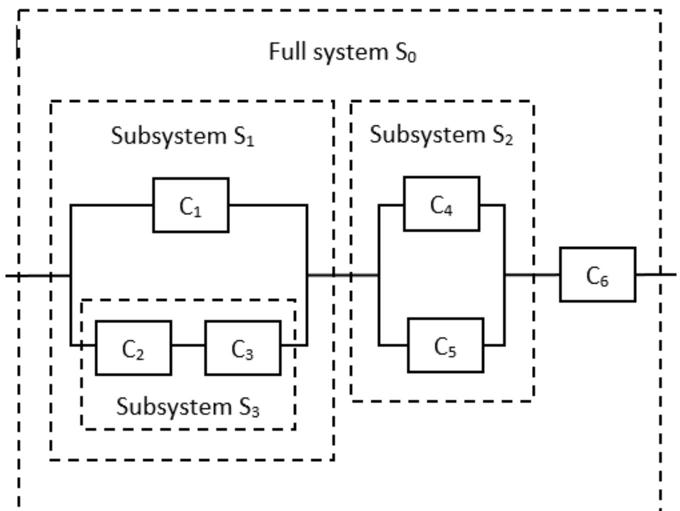

Figure 16. Reliability block diagram of a simplified missile guidance system.

4.4.2 Bayesian Network Model for the Guided Missile with Incomplete Data

In our case study, our focus is on a new guided missile system being tested and it has sensors embedded that relay information back to a ground station. We model the system as a Bayesian network (see Figure 17). Note that, we need to add nodes for the subsystems and system while constructing a BN, even though they are not actual components (Bobbio *et al.*, 2001). In this system, we can only monitor 3 nodes:

system node (S_0), subsystem S_1 and component C_5. We get discrete data from S_0 and continuous data from S_1 and C_5 such that: We observe the state of the system at a specific time (functional or failed), and then we analyze the components with sensors, which provide lifetime data from the components.

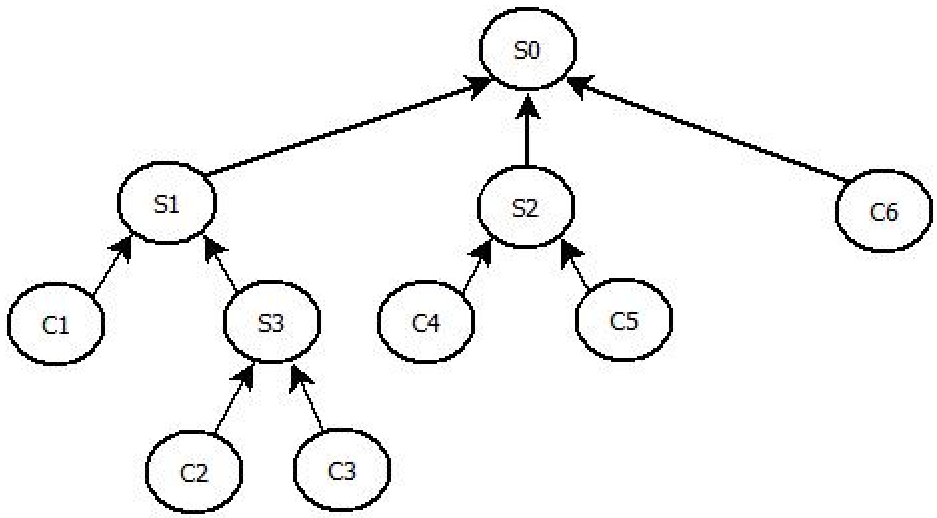

Figure 17. BN representation of the missile guidance system.

As described in Fig. 15, the first step is to define parametric models for the components of the guided missile system. Prior information exists for the reliability parameters from previous testing regimes and expert solicitation. The exponential distribution is adapted to model the reliability of components C_1, C_4, C_5 and C_6 as $T_i \sim Exponential(\lambda_i)$,$i = 1, 4, 5, 6$. The 2-parameter Weibull distribution is employed to model the lifetime of the component C_2 as $T_2 \sim Weibull(\beta_2, \eta_2)$. Its CDF is given as $F_{C_2}(t \mid \beta_2, \eta_2) = 1 - e^{-\left(\frac{t}{\eta_2}\right)^{\beta_2}}$. The Lognormal distribution is used to model the reliability of component C_3 as $T_3 \sim Lognormal(\mu_3, \sigma_3)$ with $F_{C_3}(t \mid \mu_3, \sigma_3) = \frac{1}{2}\left[1 + erf\left(\frac{lnt-\mu_3}{\sigma_3 sqrt(2)}\right)\right]$. The selection of these reliability models for the components is based on their respective goodness-of-fit test of these models, and the testimony of experts. Therefore, the parameter vector is defined as $\theta = \{\lambda_1, \beta_2, \eta_2, \mu_3, \sigma_3, \lambda_4, \lambda_5, \lambda_6\}$.

Meanwhile, the prior information is quantify into prior distributions for the model parameters given above. The prior is based on the testimony of experts and information from previous guided missiles. The priors used in this example are depicted in Table 15.

Table 15. Missile guidance system's basic component reliability characteristics.

Components	Parameters	Priors
1	λ_1	$\pi_0(\lambda_1) = Gamma(0.5, 0.15)$
2	β_2	$\pi_0(\beta_2) = Uniform(2, 3)$
	η_2	$\pi_0(\eta_2) = Uniform(100, 150)$
3	μ_3	$\pi_0(\mu_3) = Uniform(10, 250)$
	σ_3	$\pi_0(\sigma_3) = Gamma(1, 0.007)$
4	λ_4	$\pi_0(\lambda_4) = Uniform(0.01, 0.04)$
5	λ_5	$\pi_0(\lambda_5) = Uniform(0, 0.02)$
6	λ_6	$\pi_0(\lambda_6) = Uniform(0, 0.1)$

Suppose we know that components C_2 and C_3 are connected to in series (forming subsystem S_3), S_3 and component C_1 are connected to in parallel (forming subsystem S_1), and components C_4 and C_5 are connected in parallel (forming subsystem S_2). However, the system is connected to subsystems S_1, S_2 and component C_6 by a probabilistic gate. We are interested in exploring how this structure affects the working mechanism between the system and its components and we would like to make inferences about reliability parameters of all the components with data coming from a limited number of nodes (only 3 nodes: S_0, S_1 and C_5).

We start by determining the d-separated structures in the Bayesian network as explained in Section 4.3 (see Figure 18) with respect to the evidence. In Figure 18, $L_1(t \mid \boldsymbol{\theta})$, $L_2(t \mid \boldsymbol{\theta})$ and $L_3(t \mid \boldsymbol{\theta})$ are the likelihoods of the evidence for each node.

The multilevel system structure of the guided missile is modeled following the substitution strategy depicted in Fig. 15 and Eq. 4.7. The reliability function of the system depends on the reliability of components through the probabilistic gate of the BN model.

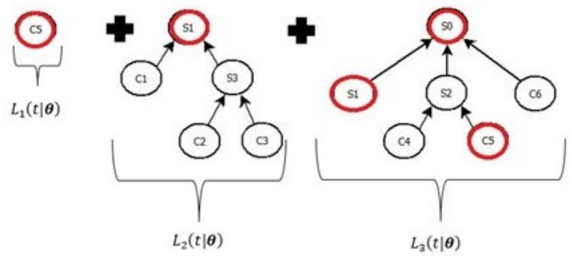

Figure 18. Three conditionally independent subsystems given S_0, S_1 and C_5.

Using $C_i = 0(1)$ to denote that component i is working (not working), the relationships given in Eq. 4.14 describing the dependence among the components are used to fully specify the Bayesian network.

$$
\begin{aligned}
&Pr(S_0 = 1 \mid S_1 = 1, S_2 = 1, C_6 = 1) = p_{111} = 0.9 \\
&Pr(S_0 = 1 \mid S_1 = 0, S_2 = 1, C_6 = 1) = p_{011} = 0.4 \\
&Pr(S_0 = 1 \mid S_1 = 1, S_2 = 0, C_6 = 1) = p_{101} = 0.3 \\
&Pr(S_0 = 1 \mid S_1 = 1, S_2 = 1, C_6 = 0) = p_{110} = 0.5 \\
&Pr(S_0 = 1 \mid S_1 = 0, S_2 = 0, C_6 = 1) = p_{001} = 0.1 \\
&Pr(S_0 = 1 \mid S_1 = 1, S_2 = 0, C_6 = 0) = p_{100} = 0.05 \\
&Pr(S_0 = 1 \mid S_1 = 0, S_2 = 1, C_6 = 0) = p_{010} = 0.25 \\
&Pr(S_0 = 1 \mid S_1 = 0, S_2 = 0, C_6 = 0) = p_{000} = 0
\end{aligned} \tag{4.14}
$$

The reliability functions of the system S_0 and subsystems S_1, S_2 and S_3 are obtained as follows (Note that $p_i(t)$ is the failure probability of basic components such that $i = 1, \ldots, 6$).

$$
\begin{aligned}
R_{S_0}(t) &= 1 - p_{S_0}(t) \\
R_{S_1}(t) &= 1 - p_{S_1}(t) \\
R_{S_2}(t) &= 1 - p_{S_2}(t) \\
R_{S_3}(t) &= 1 - p_{S_3}(t)
\end{aligned} \tag{4.15}
$$

where

$$
\begin{aligned}
p_{S_0}(t) =& p_{111} p_{S_1}(t) p_{S_2}(t) p_6(t) + p_{011}(1 - p_{S_1}(t)) p_{S_2}(t) p_6(t) + p_{101} p_{S_1}(t)(1 - p_{S_2}(t)) p_6(t) \\
&+ p_{110} p_{S_1}(t) p_{S_2}(t)(1 - p_6(t)) + p_{001}(1 - p_{S_1}(t))(1 - p_{S_2}(t)) p_6(t) \\
&+ p_{100} p_{S_1}(t)(1 - p_{S_2}(t))(1 - p_6(t)) + p_{010}(1 - p_{S_1}(t)) p_{S_2}(t)(1 - p_6(t)) \\
&+ p_{000}(1 - p_{S_1}(t))(1 - p_{S_2}(t))(1 - p_6(t)) \\
p_{S_1}(t) =& p_1(t) p_{S_3}(t) = p_1(t)[p_2(t) + p_3(t) - p_2(t) p_3(t)] \\
p_{S_2}(t) =& p_4(t) p_5(t) \\
p_{S_3}(t) =& p_2(t) + p_3(t) - p_2(t) p_3(t)
\end{aligned}
$$

and failure functions of the basic components are formulated by their CDFs. In our proposed method, we formulate the likelihood functions for each observed node starting with the lowest level (component level) of the Bayesian network, making it easier to use the functional relationships since the reliability of the nodes depend on the next lower level.

Next, we collect the evidence. The evidence is simulated with respect to the reliability distributions of the components, as represented in Table 16.

For calculating the likelihood function, $L(E \mid \boldsymbol{\theta})$, we need to formulate the reliability functions of the nodes with evidence data. As mentioned in Section 4.3.4, we use a specific form of step function for representing evidence. When the evidence is observed, it changes the form of the distribution function and therefore reliability function of the corresponding observed variable. When we are calculating the pdf of an upper level node with observation in the lower level nodes, we need to substitute $H(t_{O_i} - t)$ with the reliability function of the observed node. For example, for test #4, we use $H(51 - t)$ for node S_1 and $H(16 - t)$ for node C_5. Note that we do not need to use the step function for node S_0, as it is the last level of the Bayesian network and the causal flow stops at the system level.

Table 16. Simulated evidence data. (Data with superscript (p for pass) are right-censored observations where the unit worked at the specific time. Data with superscript (f for fail) are left-censored observations where the unit has failed.)

Test #	$\{S_0, S_1, C_5\}$
1	$\{20^{(p)}, 20^{(p)}, 17\}$
2	$\{40^{(p)}, 40^{(p)}, 28\}$
3	$\{60^{(p)}, 60^{(p)}, 1\}$
4	$\{80^{(f)}, 51, 16\}$
5	$\{100^{(f)}, 100^{(p)}, 41\}$
6	$\{120^{(f)}, 120^{(f)}, 27\}$
7	$\{140^{(f)}, 19, 63\}$
8	$\{160^{(f)}, 28, 82\}$
9	$\{180^{(f)}, 180^{(p)}, 180^{(p)}\}$
10	$\{200^{(f)}, 175, 34\}$

As a result, the reliability function will make a jump from 1 to 0 at the specific time the component failed, and we can use this unit-step function directly in the likelihood calculations. We calculate the likelihoods starting with the component level and then going upwards towards the system level.

$$
\begin{aligned}
L_1(t \mid \boldsymbol{\theta}, \mathbf{A_j}) =& f_{C_5}(t=16) = \lambda_5 e^{-16\lambda_5} \\
L_2(t \mid \boldsymbol{\theta}, \mathbf{A_j}) =& f_{S_1}(t=51) = - \left.\frac{dR_{S_1}(t)}{dt}\right|_{t=51} \\
=& - \frac{\partial R_{S_1}(t)}{\partial R_{S_3}(t)} \times \left.\frac{dR_{S_3}(t)}{dt}\right|_{t=51} - \frac{\partial R_{S_1}(t)}{\partial R_{C_1}(t)} \times \left.\frac{dR_{C_1}(t)}{dt}\right|_{t=51} \\
=& [p_2(51) + p_3(51) - p_2(51)p_3(51)] \times f_{C_1}(51) \\
& + [p_1(51)(1-p_3(51))(f_2(51))] \times f_{C_2}(51) \\
& + [p_1(51)(1-p_3(51))(f_2(51))] \times f_{C_3}(51)
\end{aligned}
$$

$$
\begin{aligned}
& L_3(t \mid \boldsymbol{\theta}, \mathbf{A_j}) = p_{S_0}(t=80) \\
& = p_{111} p_{S_1}(80) p_{S_2}(80) p_6(80) + p_{011}(1-p_{S_1}(80)) p_{S_2}(80) p_6(80) \\
& + p_{101} p_{S_1}(80)(1-p_{S_2}(80)) p_6(80) + p_{110} p_{S_1}(80) p_{S_2}(80)(1-p_6(80)) \\
& + p_{001}(1-p_{S_1}(80))(1-p_{S_2}(80)) p_6(80) + p_{100} p_{S_1}(80)(1-p_{S_2}(80))(1-p_6(80)) \\
& + p_{010}(1-p_{S_1}(80)) p_{S_2}(80)(1-p_6(80)) + p_{000}(1-p_{S_1}(80))(1-p_{S_2}(80))(1-p_6(80))
\end{aligned} \tag{4.16}
$$

where $p_{S_1}(80) = 1$ and $p_5(80) = 1$ are substituted as evidence from lower levels.

Then the joint likelihood function of the system is obtained. With the prior distributions given in Table 15, the joint posterior distribution for model parameters of the guided missile is given as

$$
p(\boldsymbol{\theta} \mid E) \propto L(E \mid \boldsymbol{\theta}) \times p(\boldsymbol{\theta}) \tag{4.17}
$$

where $L(E \mid \boldsymbol{\theta})$ is calculated from the multiplication of likelihoods of the 10 test datasets, each of which is calculated by Eq. 4.16.

The next step is sampling from the posterior distribution of the model. As described in Section 4.4, the assessment and prediction of the system reliability are carried out by generating samples from the joint posterior distribution in Eq. 4.17. The

WinBUGS software is used to implement the sampling procedure. $220,000$ samples were generated from this joint posterior distribution with $20,000$ samples for burn-in and then every other sample was kept (to reduce the auto-correlation of drawn samples) until there were $100,000$ draws from the joint posterior distribution. The posterior sample statistics of the model parameters are summarized in Table 17.

Table 17. Summary statistics of the posterior samples for the parameters.

	Mean	**SD**	**2.5%**	**25%**	**50%**	**75%**	**97.5%**
λ_1	3.3476	4.6711	0.009227	0.3653	1.561	4.435	16.68
β_2	2.5177	0.2874	2.029	2.273	2.525	2.768	2.977
η_2	127.8334	13.8123	102.1	116.7	128.9	139.7	149
μ_3	132.3514	69.1485	16.19	72.9875	133.5	192.3	244.3
σ_3	122.8305	128.1004	3.231	34.45	82.35	166.9	473.9024
λ_4	0.01947	0.007456	0.0103	0.01332	0.01762	0.02419	0.03695
λ_5	0.01143	0.005074	0.002014	0.007409	0.01162	0.01572	0.01956
λ_6	0.01167	0.01089	0.000326	0.003664	0.008554	0.01642	0.04043

4.4.3 Reliability Assessment and Prediction

Let us assume that we are analyzing a new system. Given the system has survived up to the present time ($t_p = 10$), our primary interest is on the reliability of the guided missile at this point in time. According to Eqs. 4.11, 4.12 and 4.13 derived in Section 4.3.4, the reliability and the failure rate of the system at the present time, and the probability that the system will survive another mission time $\Delta t = 5$ are obtained and presented in Table 18. The results are obtained based on $100,000$ posterior samples. The simulation based integration method is implemented.

Suppose a new system is going to be launched, and we are interested in the reliability of this new system. The mean value for the predicted reliability distribution of the new guided missile is obtained and presented in Fig. 19. It is generated based on the 100,000 posterior samples using simulation based integration.

Table 18. Summary statistics for reliability assessment of the system.

	Mean	**SD**	**2.5%**	**25%**	**50%**	**75%**	**97.5%**
$\lambda_{S_0}(t_p)$	0.02787	0.01304	0.008779	0.01811	0.02582	0.03557	0.05798
$R_{S_0}(t_p + 5 \mid t_p)$	0.9051	0.0422	0.8011	0.8835	0.9135	0.9355	0.9622
$R_{S_0}(t_p)$	0.7693	0.09899	0.551	0.7061	0.7801	0.8434	0.928

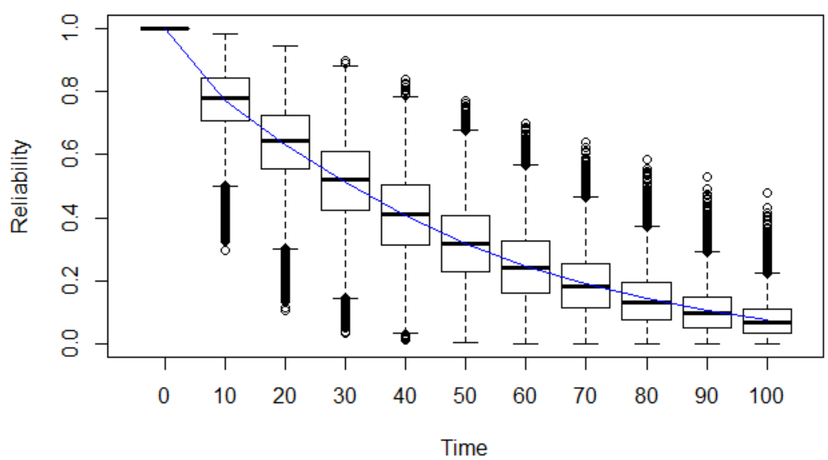

Figure 19. The reliability distribution of the new system with respect to mission time.

4.5 Conclusion and Future Research

In this chapter, a Bayesian network approach for integrating multilevel heterogeneous data sets for reliability assessment is developed. Our objective is to assess failure distribution parameters of the components and make inferences and predictions about system reliability. We start by developing the likelihood function for overlapping continuous datasets coming from some of the nodes (not all) in the network. Next, we extend this case by adding pass/fail data and provide a coherent framework for integrating multilevel heterogeneous data sets. We calculate some reliability measures like predicted reliability and failure rate of the system using an integration by simulation based method on the proposed Bayesian network framework. These measures could be use during decision making for system operation and management.

A key aspect of our method is the ability to incorporate heterogeneous overlapping data. Non-overlapping data ignores the dependencies between the datasets and removes useful information; and therefore using overlapping data is crucial in a Bayesian network framework. An overlapping data likelihood function was developed to incorporate these inherent dependencies through the use of Bayesian inference. A case study was demonstrated to highlight the effect of overlapping data and how it can be used to correctly improve our knowledge about the failure distribution parameters of the system.

The basis of our methodology is specifying the conditional independencies imposed by the Bayesian network using d-separation of the nodes. We use d-separation to formulate the conditionally independent likelihoods coming from overlapping data. The hierarchical system representation provides a good system structure so that we can separate the paths of influence easily through d-separation. For future work, we plan to work on more complex system structures. In our current framework, the components only belong to a certain subsystem. However, in reality, subsystems might share some components. The d-separation structure will change with a more complex system. Therefore, it will be more challenging to formulate the likelihood function.

Another area of future research could be to estimate the distributions of conditional probabilities. In this work, we assumed a given system structure and conditional probabilities. We would like to analyze situations where the conditional probabilities are unknown and their distributions need to be estimated from the likelihood data.

Chapter 5

A BAYESIAN FRAMEWORK FOR INCORPORATING DIFFERENT SOURCES OF PRIOR KNOWLEDGE IN RELIABILITY ASSESSMENT

5.1 Introduction and Background

Bayesian methods grow more and more complex as the systems get larger, causing an increase in the complexity of the computational methods used. Using conjugate priors somewhat overcomes this complexity problem and provides us with exact form solutions. However, when the data come from different sources and in different structures, it becomes impossible to use conjugate priors. Therefore, Bayesian researchers are showing more interest in working with non-conjugate priors. As a result, it becomes imperative that elicitation of prior distributions from different resources be done effectively. Eliciting prior distributions is rather important for representing prior knowledge more accurately and comprehensively. Thus, there is a need to develop a methodology to elicit complex, non-standard distributions coming from different sources. Although there is a broad literature in elicitation techniques, there is still a lot of aspects to consider for further research.

Bayesian statistical methods are based on the personal (or subjective) interpretation of probability. Bayesian prior and posterior distributions describe the uncertainties in the unknown parameters of the statistical model. Point estimates of parameters do not capture the uncertainty in the assessment of parameters. Therefore, Bayesian models are used to represent and quantify uncertainties and dependencies of the parameters of a complex model.

However, reliability data is not usually available for new systems or systems with modifications, so the use of expert judgment is unavoidable. In the simplest case there is only one expert. In order to include as much information as possible in the model, analysts often try to combine the distributions of several experts.

The aim of our research in this paper is to obtain as much from data (from components and system) as we can, and to elicit expert opinion accurately and combine these different streams of data to derive prior distributions for the parameters of a Bayesian model. An advantage of using Bayesian models in this context is that we can incorporate "non-data information" (also called pseudo-data) into the model. The pseudo-data can take the form of elicited data from the experts.

There are two big challenges to the problem of combining prior information. The first challenge is that specifying prior distributions for systems comprising of many components requires special thought. In the system reliability context, the reliability and lifetime of systems are functions of the parameters of the components. Therefore, the prior distributions specified on the parameters of components induce prior distributions on the reliability and lifetime of systems. We might also have direct prior information on the system parameters. Consequently, if we also have prior information about the reliability or lifetime of systems, we need a way to combine the information. There might be even cases when these two streams may have conflicts, so we also need to reconcile any difference between them. Guo (2011) used the Bayesian melding method for this problem, which was originally proposed by (Poole and Raftery, 2000). The second challenge comes from handling the pseudo-data. Quantifying non-data information is not always straightforward especially when it comes from expert opinions and it must be handled with care. Therefore, there is a need for a solid method to convert expert opinions to equivalent pseudo data for quantifying and combining prior opinions.

Our motivation for this paper is the lack of a solid unified approach for quantifying expert opinions and combining these with data coming from other sources to obtain a prior distribution for the system being studied. We propose a Bayesian methodology that incorporates different sources of prior information and reconciles these different sources, such as expert opinions and component information in order to form a prior distribution for the system. The next section presents some background information about obtaining prior distributions from the literature.

5.1.1 Elicitation Techniques

Elicitation of prior distributions is a key task for the Bayesian methodology. It is the process of formulating beliefs about uncertain quantities into a probability distribution for those quantities. That is, it converts an expert's opinions into a statistical expression of these opinions. In the context of Bayesian models, elicitation mostly arises as a method for specifying the prior distribution for the unknown parameters of the model. In the literature, the first methods involved choosing hyperparameters using conjugate prior families. With the advance in Bayesian computational methods, such as Markov chain Monte Carlo (MCMC), researchers are now able to obtain posterior distributions in the case of non-conjugate priors. However, different techniques may produce different distributions because the method of elicitation may have some effect on the way the expert states his opinions (Smith and Winkler, 1967). Bayesian modeling with informative priors based on expert opinion can provide very useful for reliability analysts (Garthwaite and O'Hagan, 2000).

In Bayesian statistical modeling, expert elicitation refers to the process of obtaining expert opinion, together with uncertainty, which is then carefully formulated into informative prior distributions (O'Hagan *et al.*, 2006). The main steps involved in elicitation as experienced by the expert are well documented (see Garthwaite and O'Hagan (2000); Clemen and Reilly (2013); Renooij (2001); Walls and Quigley (2001); Jenkinson (2005)). Direct approaches ask experts directly about parameters in the model, so experts not only require adequate statistical understanding of the role of parameters in the underlying model, but their knowledge should also be easily communicated in this way. That is why sometimes a facilitator (also called analyst or decision maker) is appointed to handle the conversion of the expert opinion to statistical form. In contrast, indirect approaches ask experts only about what they have observed. This typically involves asking experts to predict the response given particular scenarios, such as in a regression model for known covariate values.

Common approaches elicit quantiles at fixed probabilities or alternately elicit probabilities of fixed quantiles (O'Hagan, 1998). Other summary statistics may be elicited, such as moments and the mode or changes to estimates in light of hypothetical new information. Once the summary statistics about the unknown quantity has been quantified using expert knowledge, then it is necessary to estimate the prior distribution of that quantity. In most cases additional information about expert uncertainty is required, such as the equivalent sample size of their knowledge, in order to estimate the variance of prior distributions.

There has been considerable debate about using subjective opinion to construct priors (Cox, 2000; O'Hagan *et al.*, 2006). However, representation of probabilities and uncertainty under Bayesian inference contains a subjective element (Lindley, 2000; Dawid *et al.*, 2004), and other choices such as model and data are similarly subjective (Pearce *et al.*, 2001; Ferrier *et al.*, 2002). An advantage of the Bayesian inference is that it requires subjective information in the form of priors to be stated explicitly and precisely before modeling (Wintle *et al.*, 2003).

Despite abundant research on elicitation techniques, research into methods for quantifying expert opinion has never kept pace with the growing importance of Bayesian methods and we aim to reduce this deficit. As more and more Bayesian belief networks are being developed for complex real-life problem domains, it is becoming increasingly apparent that the construction of the qualitative part with the help of domain experts is feasible; the elicitation of the large number of probabilities required, however, is a far harder task. In fact, the elicitation of probabilities is often referred to as a major obstacle in building complex Bayesian models. Most methods tend to be time-consuming that it is infeasible to apply them when hundreds of probabilities are to be assessed, especially for very complex models. Faster elicitation methods are available, but are prone to even more biased answers. Renooij (2001) presented an overview of some of the issues to consider when relying on expert judgments and described the methods that are available for expert elicitation, along with their benefits and drawbacks. They discussed various issues that are to be taken into consideration when faced with the task of probability elicitation.

Garthwaite and O'Hagan (2000) proposed modeling approaches to use the elicited assessments to form subjective probability distributions. They performed statistical analysis to evaluate the objective accuracy of elicited distributions. According to their study, eliciting quantiles is the most common approach to estimating the spread of an expert's subjective distribution.

O'Hagan and Oakley (2004) outlined a Bayesian technique that allows the imprecision in elicitation to be formulated explicitly. They assumed the expert's true probability distribution is unknown to the analyst and represented the uncertainty about the expert's distribution as being the analyst's uncertainty. Oakley and O'Hagan (2007) also presented a non-parametric Bayesian analysis from this perspective. In their study, the analyst's prior beliefs about the expert's probability density function were represented by a prior distribution. These beliefs were then updated by Bayes' theorem, treating the expert's elicited summaries as data. Then the expert's probability density function can be estimated by the analyst's posterior mean.

O'Hagan *et al.* (2006) addressed applied approaches to extract information and distributional forms for use in modeling and prediction. They emphasized using distributional summaries such as probabilities, quantiles, intervals, location measures, scale and dispersion measures and measures of shape, all of which can be used as frameworks for developing survey questions in an elicitation process. They analyzed the problem of extracting critical information from experts, which will then be combined with observed data to build statistical models which can be used for prediction and inference.

In their paper Choy *et al.* (2009) outlined a framework for statistical design of expert elicitation processes for quantifying expert knowledge, in a form suitable for input as prior information into Bayesian models for ecological applications. They demonstrated the steps that need to be taken in the elicitation process, providing a useful overall description of elicitation design.

O'Hagan (2012) provided an overview and an outline of the process of eliciting knowledge from experts in probabilistic form. They explored approaches to probabilistic uncertainty specification including direct elicitation and Bayesian analysis.

Another major problem in prior knowledge elicitation is that, most of the reliability models are not able to account for prior expert opinion and data when such information is simultaneously obtained at several levels within a system. In many applications, expert opinion plays an important role in assessing system reliability, especially in large complex systems for which data collected on components may be sparse. However, Bayesian researchers overlooked the problem of incorporating pseudo-data information coming from expert opinions. Furthermore, expert opinion may be available from several experts, and the quality of information obtained from each expert may vary. Johnson *et al.* (2003) assumed that the prior density obtained from an expert concerning a specific probability takes the form of a beta density, and obtained point estimates for the probability value from each expert. They assigned an expert precision parameter for each expert and assumed that each expert precision parameter was drawn from a gamma density with known parameters. For example, if the posterior mean for the distribution of precision parameter of an expert is 12.2, this suggests that the expert's opinion is worth approximately 12 full system tests. As a result, their method simply treated expert opinion as "imprecisely-observed" data.

Another method for integrating pseudo-data into the assessment of prior distributions in literature is the "equivalent prior sample (EPS) method" (Garthwaite *et al.*, 2005). In the EPS method, an expert expresses his or her knowledge as an equivalent prior sample. However, Garthwaite *et al.* (2005) also stated that this method might tend to produce prior distributions that are unrealistically tight. Experts might equate their knowledge to too large a sample size because they might not realize the value of sample information. As a result, specification of a "prior sample" whose information content would approximately equate to an expert's knowledge is not a straightforward task, and there is also need for an objective method for relating an expert's opinion to an equivalent prior sample size. As a result, in our work, we would like to also explore the pseudo data and pseudo sample size method.

5.1.2 Verification and Validation of Experts

In this research, we define "expert" as someone that has special knowledge about the subject that we are interested in eliciting opinion about. For the sake of a more formal definition, Czembor *et al.* (2011) defined an expert as someone with:

- A minimum of 5 years of education, research experience or technical training in the specific application.
- High levels of theoretical and/or practical experience working in the specific application.
- Published research on the topic in peer-reviewed journals or reports.
- Peer nomination of being an expert.

The process of expert elicitation is basically about extracting beliefs from someone with knowledge and experience. A Bayesian model might be dominated by expert opinions, especially in case of scarce data; therefore, proper verification and validation of the experts should be be conducted. There are various techniques for evaluating the experts in the literature. According to Kadane and Wolfson (1998), reliability, coherence and calibration components can be used to validate an expert. The expert's assessments should be coherent and valid such that his assessments should follow the same pattern for the same variable. The reliability of an expert depends on the performance of the expert and it can be measured. Finally, calibration deals with the bias component in the expert's assessments and the biases can be evaluated by setting some scoring rules (Refer to Morgan *et al.* (1992) for more details about scoring rules and measuring calibration.). Cooke (1991) defined scoring as a numerical evaluation of probability assessments on the basis of observations. Scoring is of great importance for evaluating expert opinions. The expert is scored on the basis of his assessment and the observed value of that quantity. Cooke (1991) discussed two basic properties for scoring: entropy and calibration.

Entropy is defined as a good measure of degree to which the density function is spread out. Let $H(P)$ be the entropy associated with a probability density function and $P(x)$ be the probability that the elicited parameter is x. When $P(x) = 1$, $H(P) = 0$; hence an expert whose probability function has low entropy is desired. The entropy function is represented as

$$H(P) = -\int P(x)ln(P(x))dx$$

In order to define the calibration, Cooke (1991) presented a statistical hypothesis: $C(P)$:= the uncertain quantities are independent and identically distributed with the probability density function (P) provided by the expert. Let S be a sample distribution generated by observing the true values for all parameters. Then, the discrepancy between S and P is given in the following equation.

$$I(S, P) = \int S(x)ln\frac{S(x)}{P(x)}dx$$

As a result, calibration and entropy can be used to analyze expert probability assessments. Usage of these techniques can open help the experts to get adjusted to the process and give better assessments. It can be concluded that good experts should have good entropy scores and good calibration scores. However, calibrating the bias might be very tricky and it should not be skipped during the validation process. Experts are not usually accustomed to quantifying their beliefs, and there might be a number of psychological issues that make the task difficult (Denham *et al.*, 2007). Wolfson (1995) discussed some of the key psychological issues and biases that commonly occur in the elicitation process.

The biases usually represent misperceptions of probabilities. There might also be domain biases connected with experts' preferences relating to their specific fields. Identification of the bias errors generally require knowledge of the experts involved the elicitation, and require substantial amount of data. Gavasakar (1988) introduced a hierarchical model component to model elicitation errors. They tested the elicitation methods by assuming that the prior distribution had a certain form, and then adding random errors to what the answers should have been, given the specified prior. The results from the elicitation were used to compare the estimated hyperparameters with the true hyperparameters.

Overconfidence might be another cause of bias and might be the result of poor calibration. As a result, calibration provides a form of control on experts and their subjective probability assessments. There is always room for improvement for the elicitation process and training in "elicitation of subjective probabilities" can be worthwhile. Therefore, using suitable measures for calibration is a very important step in the process of expert verification and validation.

5.1.3 Combination of Several Prior Distributions

In many applied problems, the construction of informative priors using expert opinions is a delicate problem, because it might be difficult to quantify qualitative knowledge for people (O'Hagan *et al.*, 2006). With more than one expert, we may elicit from each expert a different prior and in many situations it is desirable to combine these different priors into a single "consensus" prior for the parameter θ. The more information you have, the better the results will be; therefore, it can be preferable to elicit the opinions of several experts. However, what is often needed is not a collection of different distributions but one distribution that represents the combined opinion of the experts, the result of their combined expertise, that can be used as a prior distribution in a Bayesian analysis. A good review of the issues surrounding the combination of probability distributions is given by Clemen and Winkler (1999).

There are many possible ways of combining probability distributions, which can be classified in 2 major approaches: mathematical and behavioral approaches. Our scope is only on mathematical approaches in this research. Mathematical approaches are also divided into two different approaches: axiomatic approaches (opinion pools) and Bayesian approaches. The two main axiomatic approaches are the linear opinion pool and the logarithmic opinion pool. There is a substantial literature on opinion pooling. For a detailed review of this literature, refer to Genest and Zidek (1986); Givens and Roback (1999); Jacobs (1995); O'Hagan *et al.* (2006). Let $p_i(\theta)$ represent the i^{th} expert's probability density function and w_i be the weight for the i^{th} expert's opinion. Then, the linear opinion pool is given by

$$p(\theta) = \sum_{i=1}^{n} w_i p_i(\theta) \tag{5.1}$$

with non-negative weights w_i such that $\sum_{i=1}^{n} w_i = 1$. This combination method satisfies the "marginalization property", that is, for a multivariate θ the marginal probability from the combined density for any of the variables in θ is the same as what is obtained when the elicited marginal distributions for that variable are combined. Linear pooling is the only combination method that satisfies the marginalization property.

The logarithmic opinion pool, on the other hand, is a weighted geometric mean of the densities such that

$$p(\theta) = k \prod_{i=1}^{n} p_i(\theta)^{w_i} \tag{5.2}$$

where k is the normalizing constant. The logarithmic method does not satisfy the marginalization property; however, it does satisfy the ”external Bayesian” principle. The external Bayesian principle is satisfied if the result of updating the individual expert distributions and then combining the updated distributions provides the same posterior distribution as updating the combined distribution (Poole and Raftery, 2000). Unlike the linear opinion pool, it is typically uni-modal and less dispersed. Thus, it is more likely to indicate consensual values, making it a preferable option when experts' elicited distributions are similar. Except in trivial cases, the linear opinion pool fails to have this property, while the logarithmic pool does have it, when the weights sum to one.

Despite its advantages, the logarithmic opinion pool suffers from the same problem as the linear opinion pool in that it lacks a standard method for choosing the pooling weights. It also suffers from the fact that a single expert's opinion that a probability being zero implies that the pool must also assign zero probability to that event.

Cooke (1991) described a method of choosing weights based on the experts' performance in assessing distributions for seed variables, which are quantities whose true value is known to the facilitator but not to the experts. Weights are based on p-values for evaluating how well expert assessments on seed variables align with empirical results. This method produces better elicitation than equal weighting of the experts (Cooke and Goossens, 2000). Cooke (1991) also generalized the pooling methods by raising the individual densities to the r^{th} power, taking a weighted average, raising it to the $1/r^{th}$ power and then multiplying by a constant to ensure that the combined density integrates to one.

In conclusion, the linear and logarithmic opinion pools have both their advantages and disadvantages and it is not possible to find an opinion pooling method that satisfies all good qualities like the externally Bayesian and the marginalization criteria, without making any assumptions.

A quite different approach to combining multiple experts' opinions together is the Bayesian approach, which involves experts giving information about certain events or quantities to a decision maker (DM - sometimes called a supra-Bayesian) who then updates a prior distribution using Bayes' Theorem. There are difficulties with obtaining the likelihood function required by the Bayesian methods (Clemen and Winkler, 1990). From the viewpoint of the DM, the opinions expressed by the experts are "data". The DM combines the probability distributions provided by the experts with his own prior distribution using Bayes' rule. Therefore, in the supra-Bayesian method, the pooling operator becomes the Bayes' rule and the DM's posterior distribution is the combined distribution. However, selecting the DM's prior might be problematic. Moreover, defining an appropriate likelihood function for the experts' opinions can be tedious and computationally expensive. However, due to the advancements in Markov chain Monte Carlo (MCMC) methods, we can nowadays evaluate complex posterior distributions. For example, Gelfand *et al.* (1995) modeled the likelihood function for the experts' opinions as a finite mixture of Beta distributions, and used Gibbs sampling to evaluate the DM's posterior distribution.

O'Hagan and Oakley (2004) and Oakley and O'Hagan (2007) both outlined a supra-Bayesian technique and assumed the experts' true probability distribution is unknown to the DM and represented the uncertainty about the experts' distribution as being the DM's uncertainty. In both their studies, the DM's prior beliefs about the expert's probability density function were updated by Bayes' theorem, treating the expert's elicited summaries as data. Then they estimated the expert's probability density function by the DM's posterior mean.

There are several different approaches to the problem of combining prior distributions in the literature. Savchuk and Martz (1994) developed Bayes estimators for the true binomial survival probability p when there exist multiple sources of prior information. For each source of prior information, incomplete (partial) prior information is assumed to exist in the form of either a prior mean of p or a prior credibility interval on p. Both maximum entropy and maximum posterior risk criteria are used to determine a beta prior for each source. A mixture of these beta priors is then taken as the combined prior, after which Bayes theorem is used to obtain the final mixed beta posterior distribution. Pulkkinen (1993) also discussed the problem of combining expert probability distributions. Their approach was based on the use of information theory. They derived combination procedures based on minimization of the sums of the Kullback-information between the expert distributions and the aggregated distribution. Pulkkinen and Holmberg (1997) described a method for using expert judgments, in which the combination of experts judgments is based on a Bayesian framework utilizing hierarchic models. The posterior distributions were determined by applying MCMC methods. Lipscomb *et al.* (1998) adopted a hierarchical approach that reflects a different statistical perspective on how to conceptualize and model the expert judgment synthesis problem within the supra-Bayesian framework. They presented a general approach to opinion pooling based on hierarchical modeling. Rosqvist (2000) used a Bayesian aggregation approach for experts' judgments on the failure intensity function of repairable systems. Their Bayesian statistical approach yielded posterior distributions of the parameters of the Power Law and the Log-Linear intensity functions using MCMC methods.

5.2 Methodology

In our approach, we deviate from the traditional approaches of averaging and pooling, by treating the elicited information as data and converting these pseudo data to equivalent samples of observations. Our prior combination model is based on a Bayesian approach. In this section, we will incorporate different experts with different confidence levels (that is, different pseudo sample sizes), we also combine the pseudo data with actual data coming from the components of the system, which will also induce a prior on the system parameters.

5.2.1 Incorporating Priors From Experts

In Bayesian probability theory, if the posterior distributions $p(\theta \mid x)$ are in the same family as the prior probability distribution $p(\theta)$, the prior and posterior are then called conjugate distributions, and the prior is called a conjugate prior for the likelihood function. For example, the Gaussian family is conjugate to itself with respect to a Gaussian likelihood function: if the likelihood function is Gaussian, choosing a Gaussian prior will ensure that the posterior distribution is also Gaussian. This means that the Gaussian distribution is a conjugate prior for the likelihood that is also Gaussian.

A conjugate prior gives a closed-form expression for the posterior in Bayesian analysis; otherwise a difficult numerical integration may be necessary. Conjugate priors also show how a likelihood function updates a prior distribution. All members of the exponential family have conjugate priors.

It is often useful to think of the hyperparameters of a prior distribution as corresponding to having observed a certain number of pseudo-observations with properties specified by the parameters. This is the main logic we will follow in this work. For example, the hyperparameters α and β of a beta distribution can be thought of as corresponding to $\alpha - 1$ successes and $\beta - 1$. In general, for nearly all conjugate prior distributions, the hyperparameters can be interpreted in terms of pseudo-observations. This can help to choose reasonable hyperparameters for a prior in a Bayesian framework. In a reliability based point of view, the failure probabilities can well be represented by Beta distributions, so beta-binomial models are used extensively to model pass/fail data with a probability of failure modeled by a Beta distribution. The Beta hyperparameters are often called pseudo-counts and therefore we can think of these hyperparameters as the number of times we have seen the different outcomes (pass or fail) in our prior experience before seeing actual data (Koller and Friedman, 2009). The total $n = \alpha + \beta$ of the pseudo-counts reflects how confident we are in our prior, and is often called the equivalent sample size. The relative magnitude of $\alpha + \beta$ therefore represents total weight of the pseudo-counts.

Christensen *et al.* (2011) calls the priors that allow the hyperparameters to be represented as pseudo-samples as "data augmentation priors" (DAPs). For example, the beta-binomial model is represented by the posterior distribution such that

$$p \mid x \sim Beta(x + a, n - x + b) \tag{5.3}$$

where p is the failure probability, x is the observed number of failures in a test, n is the total number of observations, and a and b are the hyperparameters of the prior beta distribution for p. In the posterior $p \mid x \sim Beta(x + a, n - x + b)$, the number of "failures" x and the hyperparameter from the prior a play similar roles. Also, the number of "successes" $n - x$ and b play similar roles. Therefore, we can think of the prior as augmenting the data with a failures and b successes out of $a + b$ trials. In DAPs, the prior density $\pi(p)$ has the same functional form as the sampling density $f(x \mid p)$ when viewed as a function of p.

However, assigning hyperparameters of a prior distribution might not reflect the actual uncertainty of experts. An expert is not usually a statistician in reality, and in most situations in reality, he is not. They might not understand the value of "sample size". It is usually an analyst who elicits experts' estimates and converts them to distributions. Therefore, after obtaining a prior distribution from the experts, we need to calibrate the experts' hyperparameters so that their uncertainty is represented in the prior distribution as accurately as possible.

In many industrial applications, expert opinion plays an important role in assessing system reliability, particularly in large complex systems because data collected on specific components and the system might be sparse. Furthermore, expert opinion may be available from several experts, and the quality of information obtained from each expert may be different due to the difference in their expertise and confidence. Incorporating expert knowledge into estimates of system reliability can therefore be a complicated task. Our solution to this problem is to elicit information from experts in the form of pseudo-observations. We analyze the continuous data case, by analyzing a gamma-exponential model and demonstrate how to incorporate several expert opinions in our Bayesian framework.

In our framework, we assume that lifetime data and prior expert opinion are available at different levels of the system, and that our primary goal is to evaluate the system reliability function, $R_0(t \mid \theta)$, defined as the probability that the system will function beyond time t, given the value of a parameter vector θ. $R_i(t \mid \theta_i)$ denotes the reliability of the component i. We are also interested in assessing the posterior distribution of the parameters, which are the failure distribution parameters of the components.

Several sources of information relevant to estimating system reliability are incorporated into our model framework. The first is lifetime data collected at individual components. The second is lifetime data collected at the system level. A third source of information is expert opinion regarding the failure rate of particular components and the system. That is, we ask each expert to provide a value for the failure rate for each component. We could ask a question such as "How often do you think this component would fail?" We then formulate a prior distribution representing the expert opinions, also including a "weight" parameter for each expert in the prior. This "weight" parameter adjusts the precision of the information solicited from each expert. We can elicit other quantities from the experts such as the failure probability, or average lifetime for a component, and formulate our priors based on these quantities. In this work, we choose to elicit the failure rate because it is directly related the failure time distributions in our system.

We assume that the prior information obtained from expert e concerning the lifetime distribution of component C_i can be formulated by a Gamma distribution such that

$$Gamma(\lambda_i \mid N_e + 1, \frac{N_e}{\mu_{i,e}}) \equiv \frac{(N_e/\mu_{i,e})^{N_e+1}}{\Gamma(N_e+1)} \lambda_i^{N_e} e^{-\frac{N_e}{\mu_{i,e}}\lambda_i} \tag{5.4}$$

In Eq. 5.4, $\mu_{i,e}$ represents the failure rate estimate that we get from expert e for λ_i, and N_e represents the weight assigned to information collected from expert e, representing the number of observations assigned to the expert e's assessment; that is, the number of the pseudo-counts. The reason we derive the expert distribution as in Eq. 5.4 is due to the interpretation of hyperparameters of Gamma distribution. Consider a gamma-exponential model such that

$$\lambda \mid t_i \sim Gamma(\alpha + n, \beta + \sum_{i=1}^{n} t_i)$$

where λ is the failure rate, t_i is the lifetime likelihood data, n is the number of lifetime observations, and α and β are the hyperparameters of the gamma prior. α and n have the same interpretation, and β and $\sum_{i=1}^{n} t_i$ have the same interpretation. Therefore, the hyperparameters are interpreted as "α observations that sum to β". As a result, we calibrate the expert parameters such that it will correspond to the interpretation of gamma priors.

We model N_e as a random parameter, by assigning a prior distribution to it. N_e also represents the consistency of the expert's assessment with observed data. We assume that each expert weight parameter N_e is drawn from a gamma density with parameters α_e and β_e, such that

$$Gamma(N_e \mid \alpha_e, \beta_e) \equiv \frac{\beta_e^{\alpha_e}}{\Gamma(\alpha_e)} N_e^{\alpha_e - 1} e^{-\beta_e N_e} \tag{5.5}$$

Let $E = \{t_i\}$ denote the test data available for constructing the likelihood function and $E_e = \{\mu_{i,e}\}$ denote the set containing expert e's elicited opinion on component i. Then the posterior distribution on model parameters is proportional to

$$\begin{aligned} p(\theta, \eta \mid \mathbf{E}) \propto \prod_{\forall i} \prod_{t \in E_i} [f_i(t_i \mid \theta_i)] \times \prod_{\forall e} \left[N_e^{\alpha_e - 1} e^{-\beta_e N_e} \prod_{i \in E_e} \lambda_i^{N_e} e^{-\frac{N_e}{\mu_{i,e}} \lambda_i} \right] \\ \times \pi(\theta \mid \eta) \times \pi(\eta) \end{aligned} \tag{5.6}$$

where $\pi(\theta \mid \eta)$ is the hierarchical priors of the parameters coming from the components and $\pi(\eta)$ is the hyper prior distribution on the η. In 5.6, we represent the system failure time distribution as a function of component life time distributions with respect to the reliability structure posed by the system reliability block diagram.

5.3 An Application to an Anti-Aircraft Missile System

As a simple demonstration of the proposed methodology, consider a weapon system (previously studied by Guo (2011) and Reese *et al.* (2005b)). The system (C_0) works if all of the components (C_1, C_2, C_3) work. The reliability block diagram for this system is depicted in Figure 20, which shows that this system consists of three components connected in series.

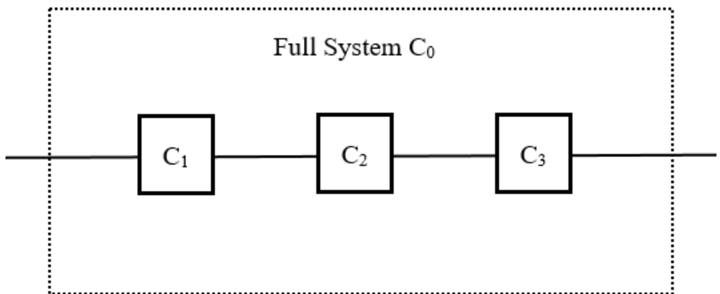

Figure 20. Reliability block diagram for a weapon system.

Test data available for estimating the reliability functions for this system are provided in Table 19. Twenty tests were conducted for each component, and ten system tests were performed. Failure times for each test are depicted in the table.

Table 19. Test data.

Component	Data (hours)
System (C_0)	23.9, 18, 53.1, 27.6, 53.7, 34.5, 47.2, 25.7, 20.8, 7.1
C_1	5.3, 65.9, 15.5, 39.4, 47.2, 28.2, 91.7, 33.6, 13.4, 13.9
	117.7, 29.3, 35.5, 4.4, 150.4, 15.7, 47, 5.1, 23.5, 25.1
C_2	65.5, 51.9, 120.2, 32, 51.5, 70.5, 37.7, 9.7, 78, 24.9
	47.7, 46.6, 105.8, 70.5, 39.9, 29.8, 48.3, 25.4, 17.7, 27.6
C_3	28.8, 51.3, 41.2, 59.2, 19.9, 57.5, 64.4, 15.7, 75, 35.2
	57.5, 49.2, 18.2, 48.8, 57.5, 35.7, 29.4, 14.6, 46.2, 9

Table 20. Expert opinions for the weapon system.

Component	Expert	Failure rate
C_0	$e1$	0.03
C_0	$e2$	0.02
C_1	$e1$	0.01
C_1	$e2$	0.01
C_3	$e2$	0.01

Two experts provided prior assessments for the system or component failure rates (see Table 20). Expert 1 provided information about the system and component 1. Expert 2 provided information about the system, and components 1 and 3. No expert opinion is available for component 2. For example, expert 1 claims that the failure rate of the system is 0.03 per hour. This means that expert 1 thinks that the system will function for about 33 hours on average.

In this application, we use an Exponential distribution to model the component failure times. The Exponential density for failure times for component C_i, $i = 1, 2, 3$, is represented by

$$f_i(t \mid \lambda_i) = \lambda_i e^{-\lambda_i t} \tag{5.7}$$

so that $\theta_i = \{\lambda_i\}$. All values of λ_i are drawn mutually independently from gamma distributions; that is

$$\pi(\lambda_i \mid \delta_\lambda, \zeta_\lambda) \propto \lambda_i^{\delta-1} e^{-\delta_\lambda \lambda_i} \tag{5.8}$$

We assume that $\delta_\lambda, \zeta_\lambda$ have independent exponential distributions with mean 1. We assigned a $Gamma(5, 1)$ prior density to the expert weight parameters N_1 and N_2, which means that each expert's assessment is considered to be worth approximately 5 observations before observing the data.

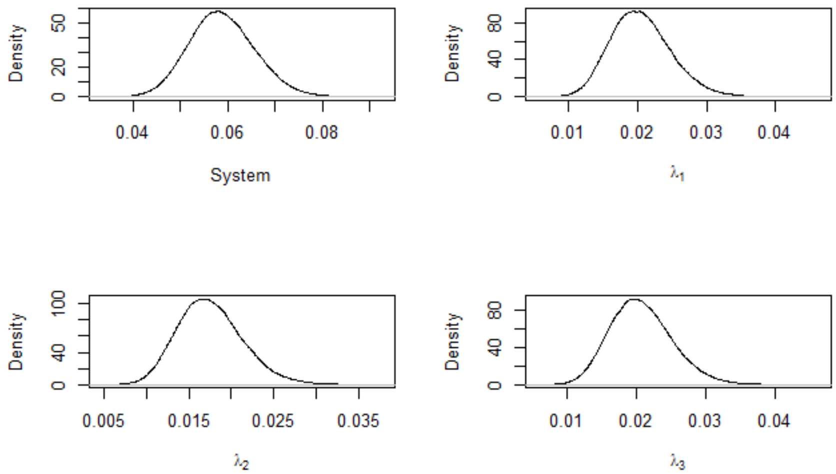

Figure 21. Posterior distributions of the reliability parameters.

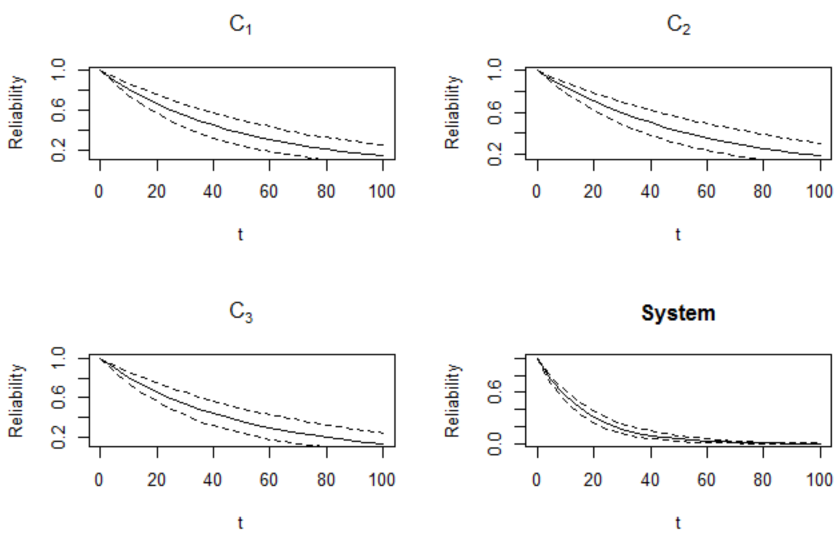

Figure 22. Posterior reliability distributions of the system and components. The solid line is the posterior mean and the dashed lines are the 90% credible interval.

To sample from the posterior distribution on model parameters and reliabilities, we ran MCMC simulations through the Bayesian software package, WinBUGS. The posterior distributions were based on $100,000$ draws from the joint posterior distri-

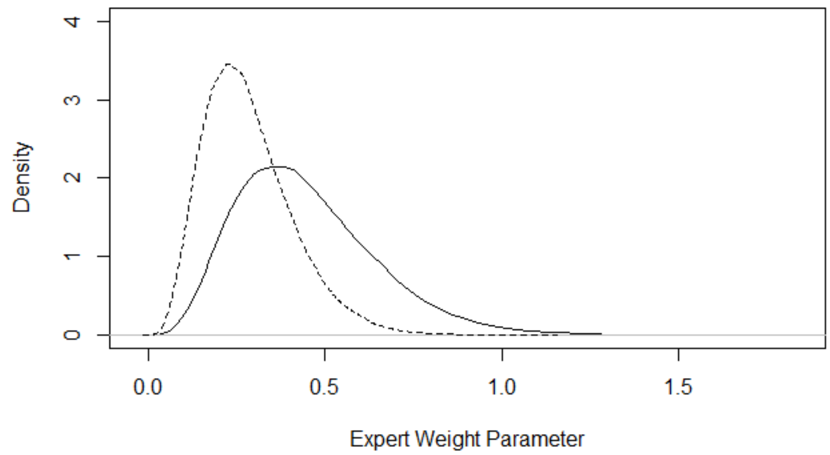

Figure 23. Posterior distributions of the experts' weight parameters. The solid line is the posterior distribution for the first expert and the dashed line is the posterior distribution for the second expert.

bution with a $20,000$ burn-in period. The posterior distribution for each parameter is plotted in Figure 21. The reliability functions of the system and components are plotted in Figure 22.

The posterior distributions for the expert precision parameters are depicted in Figure 23. These plots suggest that assessments from expert 1 were more consistent with observed data than were those from expert 2, due to the fact that the distribution obtained from expert 1 is closer to 1. Parameters for both expert 1 and expert 2 turn out to be less than 1, because the sample size of the data is much greater than the number of expert assessments, thus dominating the likelihood. We can say that the 2 experts are worth around 1 system test.

In order to analyze the effect of priors on the posterior distribution, the simulation was run with different prior distributions (see Figure 24).

According to Figure 24, the posterior distribution for the system failure rate was analyzed. In this comparison, four different models are compared: model with prior specifications from components (without expert data), pseudo data method (with expert data), logarithmic and linear opinion pools (using equal weights). As can be seen from the figure, adding the pseudo data into the model clearly improves the posterior. Linear opinion pool performs poorly, in terms of the variance and precision. We can therefore conclude that prior distributions do have an effect on the posterior distribution, and special care must be taken when combining priors in a Bayesian model.

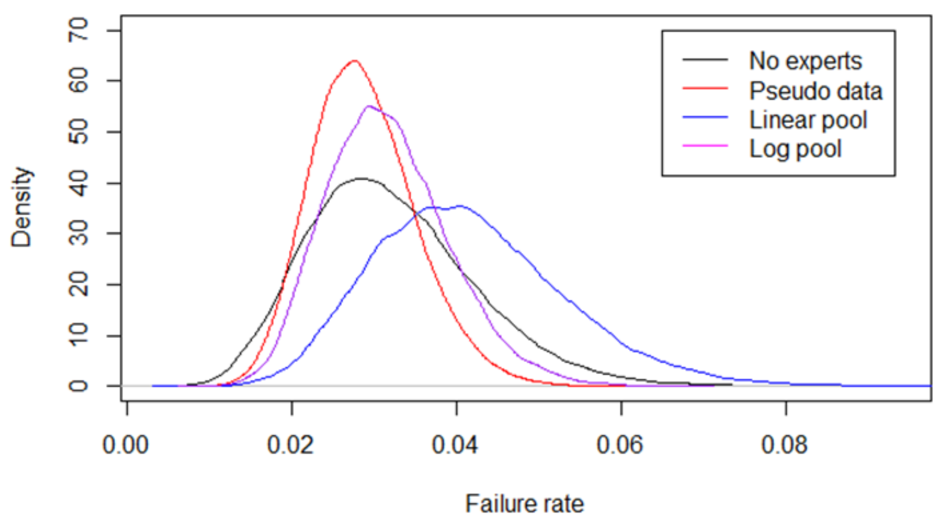

Figure 24. Posterior distributions of the failure rate given different priors.

5.4 Conclusion

In this chapter, we present a Bayesian framework for incorporating multiple sources of prior information through the treatment of expert opinion as imprecisely-observed data (pseudo-data). Our proposed hierarchical model for system reliability offers several advantages over other existing models for system reliability. Firstly, incorporating expert opinion in the form of pseudo-observations substantially simplifies statistical modeling. We can use the hierarchical priors directly in our Bayesian model, without having to use a mathematical aggregation method to combine different priors. The linear and logarithmic pooling techniques and supra Bayesian methods used for combining prior distributions in the literature require complex calculations and might be tedious to work with. Therefore, converting the experts' distributions to pseudo data proves as an effective method in a Bayesian framework.

Another advantage of our methodology is that experts are assigned a "weight" parameter representing their pseudo sample size, thus calibrating the experts' beliefs with respect to their accuracy. We formulate this weight parameter as a random variable with gamma distribution, and our hierarchical Bayesian model updates this parameter with the likelihood data. This method is especially useful when we do not have enough likelihood data, because it increases our observed sample size.

An example from the literature, a weapon system, is used as a case study in this work. We present a gamma-exponential model, modeling the lifetime data with Exponential distribution and parameters with Gamma priors. We elicit estimates about the failure rate parameter of several components from each expert and derive a Gamma distribution by calibrating the hyperparameters of the Gamma prior. As a result, we were able to obtain posterior densities for both the failure parameters and expert weight parameters.

In future work we plan to extend this framework to include more complex distributions. We would like to analyze the case of non-conjugate priors. In this case, assigning a prior distribution to the expert becomes more challenging, as it gets harder to evaluate the hyperparameters. Determining the pseudo sample size therefore might require more complicated calculations. We also would like to analyze the situations in which we elicit different quantities than the failure rate from the experts, such as the failure probability during mission time. There might be situations in which, our pseudo samples are discrete and our likelihood data are continuous, thus causing a mixture likelihood. As a result, there are many scenarios to extend this study to, creating many future research areas.

Chapter 6

CONCLUSIONS AND FUTURE WORK

The fundamental problem that this dissertation addresses is the reliability analysis of complex engineering systems through the use of Bayesian networks coupled with Bayesian inference. In the preceding chapters, we present Bayesian methods for assessing system reliability (Chapter 2, Chapter 3 and Chapter 4) and for combining prior distributions coming from different resources (Chapter 5). In this final chapter, we summarize the main contributions and discuss promising directions for further research.

6.1 Summary of Methods and Contributions

Chapter 2 proposes a Bayesian network model for assessing the system reliability at the system's early design stage. Information from parent products that was stored as a function failure record are used for inference. In our framework, failure modes and failure causes represent the nodes of the Bayesian network, whereas the conditional probabilities represent the dependencies between these causes and modes. The objective is to quantify the relationships and dependencies between failure modes and failure causes using historical records from parent products. A Bayesian network methodology is provided for early reliability prediction problem by integrating both objective and subjective reliability information. After analyzing the functional dependencies in the system, these dependencies are established in a Bayesian network model. Then, belief propagation is used to update the current knowledge about the system. Using our method, we can identify functions with high failure risk and offer suggestions for improvement.

Chapter 3 also presents a Bayesian network methodology with a deeper analysis of a complex system. In this chapter, the relationship of system/subsystem reliability to its components are examined using simultaneous pass/fail data. Information from multiple sources and multiple levels of the system to infer the conditional probabilities in a BN is combined. Firstly, a naïve scenario is presented where the complete historical dataset of the states of the system and its components are available. Then, this case is extended to a multi-state Bayesian network. Finally, the scenario of incomplete lower-level system information is discussed. Since Bayesian networks represent dependencies between the system and its components, overlapping data instead of independent data should be used in the analysis. Therefore, in this research, only data drawn simultaneously from the same system are used for inference. The dependencies between higher-level failure data and lower-level failure data are characterized by the conditional probabilities in a BN model; therefore, the objective of Chapter 3 is to infer the parameters of a Bayesian network given overlapping pass/fail data. In the independent data case, the likelihood is a multiplication of individual likelihood data coming from each component. However, in the incomplete simultaneous data case the likelihood function of evidence becomes a summation of several likelihoods that correspond to all possible state vectors of the system. For such complicated function, it is impossible to find a closed form solution of posterior probability; therefore, the computational Bayesian method, MCMC is employed. The resulting method is successful at quantifying system reliability structure with incomplete data. A MATLAB program is developed to perform compilation of the set of combinations of state vectors to be used in the MCMC simulation in WinBUGS.

Chapter 4 extends the work in Chapter 3 to systems with continuous likelihood data. A Bayesian network model has been developed for overlapping lifetime data at various levels within a complex system. A key aspect of this methodology is its ability to incorporate overlapping data. An overlapping data likelihood function is developed using d-separation in the Bayesian network model. The model developed highlights the effect of the information overlapping data contains and how it can be used to correctly improve our state of knowledge (which is the set of component reliability characteristics parameters). The resulting method completely incorporates all information taking into account the dependencies imposed by the system structure.

Chapter 5 proposes a fully Bayesian model for incorporating expert opinions with different precision and offers several advantages over other existing models. Among these are an efficient Bayesian framework for incorporating multiple sources of prior information through the treatment of expert opinion as imprecisely-observed data (also called pseudo data), and evaluating the experts' precision with a weight parameter assigned as a random variable in the model. Proposed method provides efficiency in calculations, avoiding the computational complexity posed by the pooling methods proposed in the literature.

6.2 Suggestions for Future Research

The discussions at the end of Chapter 2, Chapter 3, Chapter 4 and Chapter 5 have addressed some future research directions. We organize those that are promising and suggest other possibilities.

In system reliability, the first goal is to address more complex and general systems. As discussed previously, we can extend the proposed methods to more complex systems by modeling the relationship between different levels. If we model a very complex system using a large BN, we would have too many parameters as the complexity of systems increases because there would be too many nodes and parameters. As a result, it would be interesting to address assessing system reliabilities for a very complex Bayesian network and develop more efficient algorithms for inference. Developing more efficient simulation techniques for the proposed models is therefore very crucial. In this direction, further work could be done to propose better MCMC algorithms, especially for overlapping data.

One of the promising areas for BN related applications is safety assessment of software based systems. Software reliability is very challenging to compute, since many of the aspects of the software are not directly measurable. Therefore, BNs could be used to model software based systems to constitute a systematic way to combine quantitative reliability data with qualitative data and show the link between these components. The BN methodology can provide a useful and practical framework that supports decision-making in software engineering because of the ease of representation of causal relationships among variables (Fenton *et al.*, 2008; Fenton and Neil, 2012). Lewis (1999) discussed some of the issues surrounding Bayesian network software process modeling and outlined directions for future research. Dahll (2000) discussed how to combine disparate sources of information in the safety assessment of software-based systems using Bayesian networks. Bibi and Stamelos (2004) suggested the use of Bayesian networks for representing software process models. Misirli and Bener (2014) investigated the applications of Bayesian networks in software engineering in terms of techniques used to learn causal relationships among variables and techniques used to infer the parameters. They proposed a hybrid BN to improve evidence-based decision-making in software engineering, showing that hybrid BNs are powerful frameworks that combine expert knowledge with quantitative data.

Bayesian networks can provide a network of software work flows and their interdependencies. They are highly visual tools that can indicate which work flows affect others. They enable evolution of the process as they can be used for sensitivity analysis in order to explore the impact of some changes in software process before actually implementing them. To satisfy this objective, the software process needs to be analyzed and carefully modeled in order to encourage it's understanding, assessment and improvement. Therefore, it would be a good research direction to develop generic Bayesian network process models for software based systems.

Another future research area is assessing prior distributions from experts and combining these distributions. In real life, we might get very complex distributions, so it would be an interesting research direction and more work could be done on how to assess the hyperparameters of different prior distributions.

MCMC simulation techniques were used all throughout this dissertation. Further work could be done on inference for Bayesian networks using other approximation techniques.

REFERENCES

Augustine, M., O. P. Yadav, R. Jain and A. Rathore, “Cognitive map-based system modeling for identifying interaction failure modes”, Research in Engineering Design **23**, 2, 105–124 (2012).

Bibi, S. and I. Stamelos, “Software process modeling with bayesian belief networks”, in “Proceedings of 10th International Software Metrics Symposium (Metrics 2004)”, vol. 14, p. 16 (2004).

Bobbio, A., L. Portinale, M. Minichino and E. Ciancamerla, “Improving the analysis of dependable systems by mapping fault trees into bayesian networks”, Reliability Engineering & System Safety **71**, 3, 249–260 (2001).

Boudali, H. and J. B. Dugan, “A discrete-time bayesian network reliability modeling and analysis framework”, Reliability Engineering & System Safety **87**, 3, 337–349 (2005).

Boudali, H. and J. B. Dugan, “A continuous-time bayesian network reliability modeling, and analysis framework”, Reliability, IEEE Transactions on **55**, 1, 86–97 (2006).

Boyen, X. and D. Koller, “Tractable inference for complex stochastic processes”, in “Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence”, pp. 33–42 (Morgan Kaufmann Publishers Inc., 1998).

Chandrasekaran, B., A. K. Goel and Y. Iwasaki, “Functional representation as design rationale”, Computer **26**, 1, 48–56 (1993).

Choy, S. L., R. O’Leary and K. Mengersen, “Elicitation by design in ecology: using expert opinion to inform priors for bayesian statistical models”, Ecology **90**, 1, 265–277 (2009).

Christensen, R., W. Johnson, A. Branscum and T. E. Hanson, *Bayesian ideas and data analysis: an introduction for scientists and statisticians* (CRC Press, 2011).

Clark, G. and R. Paasch, “Diagnostic modeling and diagnosability evaluation of mechanical systems”, Journal of Mechanical Design **118**, 3, 425–431 (1996).

Clemen, R. and T. Reilly, *Making hard decisions with DecisionTools* (Cengage Learning, 2013).

Clemen, R. T. and R. L. Winkler, “Unanimity and compromise among probability forecasters”, Management Science **36**, 7, 767–779 (1990).

Clemen, R. T. and R. L. Winkler, "Combining probability distributions from experts in risk analysis", Risk analysis **19**, 2, 187–203 (1999).

Cooke, R. and L. Goossens, "Procedures guide for structured expert judgment", EUR(Luxembourg) (2000).

Cooke, R. M., *Experts in uncertainty: opinion and subjective probability in science* (New York, NY (United States); Oxford University Press, 1991).

Cooper, G. F., "The computational complexity of probabilistic inference using bayesian belief networks", Artificial intelligence **42**, 2, 393–405 (1990).

Cox, D., "Comments on paper by lindley", Statistician **49**, 3, 321–324 (2000).

Czembor, C. A., W. K. Morris, B. A. Wintle and P. A. Vesk, "Quantifying variance components in ecological models based on expert opinion", Journal of Applied Ecology **48**, 3, 736–745 (2011).

Dahll, G., "Combining disparate sources of information in the safety assessment of software-based systems", Nuclear Engineering and Design **195**, 3, 307–319 (2000).

Dawid, A. P. *et al.*, "Probability, causality and the empirical world: a bayes–de finetti–popper–borel synthesis", Statistical Science **19**, 1, 44–57 (2004).

Denham, R., K. Mengersen *et al.*, "Geographically assisted elicitation of expert opinion for regression models", Bayesian Analysis **2**, 1, 99–135 (2007).

Derelöv, M., "Qualitative modelling of potential failures: on evaluation of conceptual design", Journal of engineering design **19**, 3, 201–225 (2008).

Doguc, O. and J. E. Ramirez-Marquez, "A generic method for estimating system reliability using bayesian networks", Reliability Engineering & System Safety **94**, 2, 542–550 (2009).

Eubanks, C. F., S. Kmenta and K. Ishii, "Advanced failure modes and effects analysis using behavior modeling", in "ASME Design Engineering Technical Conferences", pp. 14–17 (1997).

Fenton, N. and M. Neil, *Risk assessment and decision analysis with Bayesian networks* (CRC Press, 2012).

Fenton, N., M. Neil and D. Marquez, "Using bayesian networks to predict software defects and reliability", Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability **222**, 4, 701–712 (2008).

Ferrier, S., G. Watson, J. Pearce and M. Drielsma, "Extended statistical approaches to modelling spatial pattern in biodiversity in northeast new south wales. i. species-level modelling", Biodiversity & Conservation **11**, 12, 2275–2307 (2002).

Garthwaite, P. H., J. B. Kadane and A. O'Hagan, "Statistical methods for eliciting probability distributions", Journal of the American Statistical Association **100**, 470, 680–701 (2005).

Garthwaite, P. H. and A. O'Hagan, "Quantifying expert opinion in the uk water industry: an experimental study", Journal of the Royal Statistical Society: Series D (The Statistician) **49**, 4, 455–477 (2000).

Gavasakar, U., "A comparison of two elicitation methods for a prior distribution for a binomial parameter", Management Science **34**, 6, 784–790 (1988).

Gelfand, A. E., B. K. Mallick and D. K. Dey, "Modeling expert opinion arising as a partial probabilistic specification", Journal of the American Statistical Association **90**, 430, 598–604 (1995).

Gelman, A., J. B. Carlin, H. S. Stern and D. B. Rubin, *Bayesian data analysis*, vol. 2 (Taylor & Francis, 2014).

Genest, C. and J. V. Zidek, "Combining probability distributions: A critique and an annotated bibliography", Statistical Science pp. 114–135 (1986).

Givens, G. H. and P. J. Roback, "Logarithmic pooling of priors linked by a deterministic simulation model", Journal of Computational and Graphical Statistics **8**, 3, 452–478 (1999).

Graves, T. L., M. S. Hamada, R. Klamann, A. Koehler and H. F. Martz, "A fully bayesian approach for combining multi-level information in multi-state fault tree quantification", Reliability Engineering & System Safety **92**, 10, 1476–1483 (2007).

Graves, T. L., M. S. Hamada, R. Klamann, A. Koehler and H. F. Martz, "Using simultaneous higher-level and partial lower-level data in reliability assessments", Reliability Engineering & System Safety **93**, 8, 1273–1279 (2008).

Guo, J., *Bayesian methods for system reliability and community detection*, Ph.D. thesis, Iowa State University (2011).

Hamada, M., H. F. Martz, C. S. Reese, T. Graves, V. Johnson and A. G. Wilson, "A fully bayesian approach for combining multilevel failure information in fault tree quantification and optimal follow-on resource allocation", Reliability Engineering & System Safety **86**, 3, 297–305 (2004).

Heckerman, D., *A tutorial on learning with Bayesian networks* (Springer, 1998).

Hirtz, J., R. B. Stone, D. A. McAdams, S. Szykman and K. L. Wood, "A functional basis for engineering design: reconciling and evolving previous efforts", Research in engineering Design **13**, 2, 65–82 (2002).

Huang, Z. and Y. Jin, "Conceptual stress and conceptual strength for functional design-for-reliability", in "ASME 2008 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference", pp. 437–447 (American Society of Mechanical Engineers, 2008).

Hulting, F. L. and J. A. Robinson, "The reliability of a series system of repairable subsystems: a bayesian approach", Naval Research Logistics (NRL) **41**, 4, 483–506 (1994).

Hunt, B. R., R. L. Lipsman and J. M. Rosenberg, *A guide to MATLAB: for beginners and experienced users* (Cambridge University Press, 2014).

Iamsumang, C., A. Mosleh and M. Modarres, "Hybrid dbn monitoring and anomaly detection algorithms for on-line shm", in "Reliability and Maintainability Symposium (RAMS), 2015 Annual", pp. 1–7 (IEEE, 2015).

Jackson, C. S., *Bayesian Inference with Overlapping Data: Methodology and Application to System Reliability Estimation and Sensor Placement Optimization*, Ph.D. thesis, University of Maryland (2011).

Jacobs, R. A., "Methods for combining experts' probability assessments", Neural computation **7**, 5, 867–888 (1995).

Jenkinson, D., "The elicitation of probabilities: A review of the statistical literature", Tech. rep., Citeseer (2005).

Johnson, V. E., T. L. Graves, M. S. Hamada and C. S. Reese, "A hierarchical model for estimating the reliability of complex systems", in "Bayesian Statistics 7: Proceedings of the Seventh Valencia International Meeting", pp. 199–214 (Oxford University Press, USA, 2003).

Kadane, J. and L. J. Wolfson, "Experiences in elicitation", Journal of the Royal Statistical Society: Series D (The Statistician) **47**, 1, 3–19 (1998).

Koller, D. and N. Friedman, *Probabilistic graphical models: principles and techniques* (MIT press, 2009).

Kurtoglu, T. and I. Y. Tumer, "A graph-based fault identification and propagation framework for functional design of complex systems", Journal of Mechanical Design **130**, 5, 051401 (2008).

Langseth, H., T. D. Nielsen, R. Rumí and A. Salmerón, "Inference in hybrid bayesian networks", Reliability Engineering & System Safety **94**, 10, 1499–1509 (2009).

Langseth, H. and L. Portinale, "Bayesian networks in reliability", Reliability Engineering & System Safety **92**, 1, 92–108 (2007).

Lauritzen, S. L., "The em algorithm for graphical association models with missing data", Computational Statistics and Data Analysis **19**, 2, 191–201 (1995).

Lewis, N. D., "Continuous process improvement using bayesian belief networks", Computers & industrial engineering **37**, 1, 449–452 (1999).

Li, M., J. Liu, J. Li and B. U. Kim, "Bayesian modeling of multi-state hierarchical systems with multi-level information aggregation", Reliability Engineering & System Safety **124**, 158–164 (2014).

Lindley, D. V., "The philosophy of statistics", Journal of the Royal Statistical Society: Series D (The Statistician) **49**, 3, 293–337 (2000).

Lipscomb, J., G. Parmigiani and V. Hasselblad, "Combining expert judgment by hierarchical modeling: An application to physician staffing", Management Science **44**, 2, 149–161 (1998).

Mahadevan, S., R. Zhang and N. Smith, "Bayesian networks for system reliability reassessment", Structural Safety **23**, 3, 231–251 (2001).

Marquez, D., M. Neil and N. Fenton, "Improved reliability modeling using bayesian networks and dynamic discretization", Reliability Engineering & System Safety **95**, 4, 412–425 (2010).

Martz, H. and R. Wailer, "Bayesian reliability analysis of complex series/parallel systems of binomial subsystems and components", Technometrics **32**, 4, 407–416 (1990).

Martz, H., R. Wailer and E. Fickas, "Bayesian reliability analysis of series systems of binomial subsystems and components", Technometrics **30**, 2, 143–154 (1988).

Misirli, A. T. and A. B. Bener, "Bayesian networks for evidence-based decision-making in software engineering", Software Engineering, IEEE Transactions on **40**, 6, 533–554 (2014).

Moral, S., R. Rumí and A. Salmerón, "Mixtures of truncated exponentials in hybrid bayesian networks", in "Symbolic and Quantitative Approaches to Reasoning with Uncertainty", pp. 156–167 (Springer, 2001).

Morgan, M. G., M. Henrion and M. Small, *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis* (Cambridge university press, 1992).

Neil, M., M. Tailor and D. Marquez, "Inference in hybrid bayesian networks using dynamic discretization", Statistics and Computing **17**, 3, 219–233 (2007).

Neil, M., M. Tailor, D. Marquez, N. Fenton and P. Hearty, "Modelling dependable systems using hybrid bayesian networks", Reliability Engineering & System Safety **93**, 7, 933–939 (2008).

Oakley, J. E. and A. O'Hagan, "Uncertainty in prior elicitations: a nonparametric approach", Biometrika **94**, 2, 427–441 (2007).

O'Hagan, A., "Eliciting expert beliefs in substantial practical applications", Journal of the Royal Statistical Society: Series D (The Statistician) **47**, 1, 21–35 (1998).

O'Hagan, A., "Probabilistic uncertainty specification: Overview, elaboration techniques and their application to a mechanistic model of carbon flux", Environmental Modelling & Software **36**, 35–48 (2012).

O'Hagan, A., C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley and T. Rakow, *Uncertain judgements: eliciting experts' probabilities* (John Wiley & Sons, 2006).

O'Hagan, A. and J. E. Oakley, "Probability is perfect, but we can't elicit it perfectly", Reliability Engineering & System Safety **85**, 1, 239–248 (2004).

Otto, K. N. and K. L. Wood, "Product evolution: a reverse engineering and redesign methodology", Research in Engineering Design **10**, 4, 226–243 (1998).

Pahl, G. and W. Beitz, *Engineering design: a systematic approach* (Springer Science & Business Media, 2013).

Pan, R. and S. E. Rigdon, "Bayes inference for general repairable systems", Journal of Quality Technology **41**, 1 (2009).

Pearce, J., K. Cherry, G. Whish *et al.*, "Incorporating expert opinion and fine-scale vegetation mapping into statistical models of faunal distribution", Journal of Applied Ecology **38**, 2, 412–424 (2001).

Pearl, J., *Probabilistic reasoning in intelligent systems: networks of plausible inference* (Morgan Kaufmann, 2014).

Peng, W., H.-Z. Huang, M. Xie, Y. Yang and Y. Liu, "A bayesian approach for system reliability analysis with multilevel pass-fail, lifetime and degradation data sets", Reliability, IEEE Transactions on **62**, 3, 689–699 (2013).

Poole, D. and A. E. Raftery, "Inference for deterministic simulation models: the bayesian melding approach", Journal of the American Statistical Association **95**, 452, 1244–1255 (2000).

Pulkkinen, U., "Methods for combination of expert judgements", Reliability Engineering & System Safety **40**, 2, 111–118 (1993).

Pulkkinen, U. and J. Holmberg, *A method for using expert judgement in PSA* (Finnish Centre for Radiation and Nuclear Safety Helsinki, 1997).

Qian, L. and J. S. Gero, "Function–behavior–structure paths and their role in analogy-based design", Artificial Intelligence for Engineering, Design, Analysis and Manufacturing **10**, 04, 289–312 (1996).

R CORE TEAM, A. *et al.*, "R: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria. 2012", Tech. rep., ISBN 3-900051-07-0, URL http://www. R-project. org (2012).

Reese, C. S., M. Hamada and D. Robinson, "Assessing system reliability by combining multilevel data from different test modalities", Qual. Technol. Quant. Manag **2**, 177–188 (2005a).

Reese, C. S., V. E. Johnson, M. Hamada and A. Wilson, "A hierarchical model for the reliability of an anti-aircraft missile system", UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series, Houston, TX (2005b).

Reese, C. S., A. G. Wilson, J. Guo, M. S. Hamada and V. E. Johnson, "A bayesian model for integrating multiple sources of lifetime information in system-reliability assessments", Journal of quality technology **43**, 2, 127 (2011).

Renooij, S., "Probability elicitation for belief networks: issues to consider", The Knowledge Engineering Review **16**, 03, 255–269 (2001).

Rosqvist, T., "Bayesian aggregation of experts' judgements on failure intensity", Reliability Engineering & System Safety **70**, 3, 283–289 (2000).

Sanchez, L. M., *Reliability Information and Testing Integration for New Product Design*, Ph.D. thesis, Arizona State University (2014).

Sanchez, L. M. and R. Pan, "An enhanced parenting process: Predicting reliability in product's design phase", Quality Engineering **23**, 4, 378–387 (2011).

Savchuk, V. P. and H. F. Martz, "Bayes reliability estimation using multiple sources of prior information: binomial sampling", Reliability, IEEE Transactions on **43**, 1, 138–144 (1994).

Smith, A. and H. Winkler, "A simple method for the isolation of adrenal chromaffin granules on a large scale", Biochemical Journal **103**, 2, 480 (1967).

Spiegelhalter, D., A. Thomas, N. Best and D. Lunn, "Winbugs user manual", (2003).

Spiegelhalter, D. J. and S. L. Lauritzen, "Sequential updating of conditional probabilities on directed graphical structures", Networks **20**, 5, 579–605 (1990).

Sridharan, P. and M. I. Campbell, "A study on the grammatical construction of function structures", AIE EDAM **19**, 03, 139–160 (2005).

Stone, R. B. and K. L. Wood, "Development of a functional basis for design", Journal of Mechanical design **122**, 4, 359–370 (2000).

Tumer, I. Y. and R. B. Stone, "Analytical method for mapping function to failure during high-risk component development", in "Proceedings of the Design Engineering Technical Conferences", (2001).

Walls, L. and J. Quigley, "Building prior distributions to support bayesian reliability growth modelling using expert judgement", Reliability Engineering & System Safety **74**, 2, 117–128 (2001).

Wang, K.-L. and Y. Jin, "An analytical approach to functional design", in "ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference", pp. 449–459 (American Society of Mechanical Engineers, 2002).

Warr, R. L. and D. H. Collins, "Bayesian nonparametric models for combining heterogeneous reliability data", Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability **228**, 2, 166–175 (2014).

Wilson, A. G., C. M. Anderson-Cook and A. V. Huzurbazar, "A case study for quantifying system reliability and uncertainty", Reliability Engineering & System Safety **96**, 9, 1076–1084 (2011).

Wilson, A. G., T. L. Graves, M. S. Hamada and C. S. Reese, "Advances in data combination, analysis and collection for system reliability assessment", Statistical Science pp. 514–531 (2006).

Wilson, A. G. and A. V. Huzurbazar, "Bayesian networks for multilevel system reliability", Reliability Engineering & System Safety **92**, 10, 1413–1420 (2007).

Wintle, B. A., M. A. McCarthy, C. T. Volinsky and R. P. Kavanagh, "The use of bayesian model averaging to better represent uncertainty in ecological models", Conservation Biology **17**, 6, 1579–1590 (2003).

Wolfson, L. J., "Elicitation of priors and utilities for bayesian analysis", Unpublished Ph. D. Thesis, Department of Statistics, Carnegie Mellon University **5**, 6, 7 (1995).

Yontay, P., L. M. Sanchez and R. Pan, "Bayesian network for reliability prediction in functional design stage", in "Reliability and Maintainability Symposium (RAMS), 2015 Annual", pp. 1–6 (IEEE, 2015).

Zhong, X., M. Ichchou and A. Saidi, "Reliability assessment of complex mechatronic systems using a modified nonparametric belief propagation algorithm", Reliability Engineering & System Safety **95**, 11, 1174–1185 (2010).