Extending Cyber-Physical Systems to Support Stakeholder Decisions Under Resource and User Constraints: Applications to Intelligent Infrastructure and Social Urban Systems

by

Katherine A. Flanigan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Civil Engineering)
in the University of Michigan
2020

Doctoral Committee:

Professor Jerome P. Lynch, Chair
Associate Professor Branko Kerkez
Professor Elizabeth B. Moje
Associate Professor Jeffrey T. Scruggs

Katherine A. Flanigan

kaflanig@umich.edu

ORCID iD: 0000-0002-2454-5713

ACKNOWLEDGEMENTS

I am indebted to my advisor, Professor Jerry Lynch, who has played an integral role in my personal and professional development. Through his mentorship, I have learned to measure my success not just by the quality and quantity of my research, but also by the impact of my research, teaching, and commitment to the academic community. As a role model, Professor Lynch has displayed great personal and professional integrity, which are qualities that I will continue to value and strive for as I enter my new academic position. I look forward to continuing to engage with Professor Lynch for years to come, both as a friend and as members within the same academic community.

I would like to acknowledge my committee members, Professor Branko Kerkez, Professor Elizabeth Moje, and Professor Jeff Scruggs, for their insightful input, advice, and encouragement. I would especially like to thank Professor Scruggs for his guidance in both my course work and research, as well as for his unwavering kindness, which I have valued greatly.

I have had the pleasure of working in the Laboratory of Intelligent Systems Technology alongside Rui Hou, Nephi Johnson, Andy Burton, Omid Bahrami, Hao Zhou, Kidus Admassu, Gabe Draughon, Wentao Wang, Patrick Sun, Sean O'Connor, and Yilan Zhang. I am very thankful for their friendship and consider myself extremely lucky to have been able to work with such a hard-working group of scholars.

As an undergraduate student at Princeton University I had the absolute privilege of working with Professor Branko Glišić. I will be forever grateful to Professor Glišić for believing in me and for encouraging me to attend graduate school.

An undertaking of this magnitude would not have been possible without the support of my family and friends (including Milo!). I would like to thank my family for their encouragement throughout the course of my academic journey, even as I have carved my own path so far from home. The pure joy and comfort I derive from my research and academic community is without a doubt rooted in the unrelenting enthusiasm my family expressed towards my fierce love for school growing up. I would also like to give special thanks to Michelle Basham, Malcolm Ammons, Christian Flores, Cassi Champagne, Rachel Vitali, Danny Johnson, Anne Menefee, and Jacquie Handley for their encouragement and friendship.

This research was sponsored and supported by the National Science Foundation (NSF) under grant numbers 1362513, 1446521, 0910765, 1637232, and 1831347, the Michigan Department of Transportation (MDOT) via Contract# 131529 OR 15-114, the United States Department of Transportation (USDOT) via OST-R Award# OASRTRS-14-H-MICH, Union Pacific Railroad, and the Knight Foundation. Additional support was provided by the NSF Graduate Research Fellowship.

Finally, I would like to gratefully acknowledge MDOT, Rui Hou, Nephi Johnson, Curt Wolf, Jacqueline Handley, Todd Martindale, Michael Freeman, Mohammed Ettouney, the Detroit Riverfront Conservancy, and Union Pacific Railroad for their assistance in the field.

TABLE OF CONTENTS

Chapter 6: Community Engagement Using Urban Sensing – Technology Development, Deployment Studies, and CPSS Architecture 198

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

In recent years, rapid urbanization has imposed greater load demands on physical infrastructure while placing stressors (e.g., pollution, congestion, social inequity) on social systems. Despite these challenges, opportunities are emerging from the unprecedented proliferation of information technologies enabling ubiquitous sensing, cloud computing, and full-scale automation. Together, these advancements enable "intelligent" systems that promise to enhance the operation of the built environment. Even with these advancements, the ability of professionals to "sense for decisions" —data-driven decision processes based on sensed data that have quantifiable returns on investment—remains unrealized for an entire class of problems. In response, this dissertation builds a rigorous foundation enabling stakeholders to use sensor data to inform decisions in two applications: infrastructure asset management and community-engaged decision making.

This dissertation aligns sensing strategies with decisions governing infrastructure management by extending the role of reliability methods to quantify system performance. First, the reliability index is used as a scalar measure of the safety (i.e., failure probability) that is extracted from monitoring data to assess structural condition relative to a failure limit state. As an example, long-term data collected from a wireless sensing network (WSN) installed on the Harahan Bridge (Memphis, TN) is used in a reliability framework to track the fatigue life of critical eyebar assemblies. The proposed reliability-based SHM framework is then generalized to formally and more broadly link SHM data with condition

ratings (CRs) because inspector-assigned CRs remain the primary starting point for asset management decisions made in practice today. While reliability methods historically quantify safety with respect to a single failure limit state, this work demonstrates that there exist measurable reliability index values associated with “lower” limit states below failure that more richly characterize structural performance and rationally map to CR scales. Consequently, monitoring data can be used to assign CRs based on quantitative information encompassing the measurable damage domain, as opposed to relying on visual inspection. This work reflects the first-ever SHM framework to explicitly map monitoring data to actionable decisions and is validated using a WSN on the Telegraph Road Bridge (TRB) (Monroe, MI).

A primary challenge faced by solar-powered WSNs is their stringent energy constraints. For decision-making processes relying on statistical estimation of performance, the utility of data should be considered to optimize the data collection process given these constraints. This dissertation proposes a novel stochastic data collection and transmission policy for WSNs that minimizes the variance of a measured process' estimated parameters subject to constraints imposed by energy and data buffer sizes, stochastic models of energy and event arrivals, the value of measured data, and temporal death. Numerical results based on one-year of data collected from the TRB illustrate the gains achieved by implementing the optimal policy to obtain response data used to estimate the reliability index.

Finally, this dissertation extends the work performed in WSN and sense-for-decision frameworks by exploring their role in community-based decision making. This work poses societal engagement as a necessary entry point to urban sensing efforts because

members of under-resourced communities are vulnerable to lack of access to data and information. A novel, low-power WSN architecture is presented that functions as a user-friendly sensing solution that communities can rapidly deploy. Applying this platform, transformative work to "democratize" data is proposed in which members of vulnerable communities collect data and generate insights that inform their decision-making strategies.

CHAPTER 1

Introduction

1.1 The Emerging Physical-Economic-Social System Nexus

Since the beginning of the profession, civil and environmental engineering professionals have served as stewards of the built environment. Although ensuring the common good of society remains at the heart of the field's mission, the challenges faced by the profession have changed rapidly in recent years. Civil and environmental engineers designed, built, and maintained the foundations for modern society from the ground up. Now, unprecedented grand challenges—urbanization, aging infrastructure, and climate change—threaten the integrity of the built environment. In response to these challenges, civil and environmental engineers are being asked to expand the life of aging infrastructure, unlock capacity in infrastructure systems to accommodate population growth, and be adaptive in the face of climate change. These grand challenges, as well as the grand opportunities born out of the proliferation of low-cost sensing technologies, increases in computational capabilities, digitalization, and the development of automation tools, are radically changing the engineering landscape and require civil and environmental engineers to draw upon tools and technologies from across disciplines to devise forward-looking and scalable solutions.

At the forefront of these solutions is a need to design and manage infrastructure to be resilient in the face of these grand challenges. In addition to protecting public safety, infrastructure is a necessary input to every social and economic output (ASCE 2016). Well-managed infrastructure protects populations against natural and man-made hazards that increasingly threaten urban centers, enables the mobility of populations and the transport of goods, water, and energy resources through air, road, rail, and maritime transport systems, and facilitates critical evacuation and external relief support before and after disasters (UNU 2016). However, the neglect of aging infrastructure in the United States—ranging from dams, to wastewater treatment facilities and sewers, to bridges—leaves societies vulnerable as they become increasingly dependent on the transport of vital resources. The rehabilitation of aging dams, wastewater infrastructure, and bridges alone requires an investment of $439 billion (ASCE 2017). As climate change increases the frequency and intensity of natural hazards that threaten aging infrastructure, maintenance, rehabilitation, and replacement may not be enough. For example, as sea levels rise and storms intensify along the nation's coastal regions, civil and environmental engineers need to not only design and maintain more robust infrastructure, but also offer solutions that enable communities to adapt to these changes. While the built environment has historically been the focal point of civil and environmental engineering efforts, the economy, culture, security, and quality of life of populations are becoming increasingly dependent on the performance of civil and environmental infrastructure systems (Jacobs 1961). The aforementioned grand challenges will not only accentuate complex physical problems within the built environment, but also impose stressors on social and economic systems that force society to change the ways it designs, manages, values, and uses infrastructure.

In addition to monitoring and controlling traditional physical assets, rigorous methods to solve and adapt to these grand challenges will require a new quantitative understanding of the relationship between social systems and the built and natural environments, as well as the economic implications of this relationship. The convergence of physical, economic, and social systems is increasingly evident across the domain of civil and environmental engineering. For example, societal resilience requires an understanding of the *interaction* between infrastructure changes and community response when stressors impact one—or both—of these systems (Guidotti 2019). Additionally, technological advancements introduce new components of social interaction when services are embedded within infrastructure systems. For instance, the sharing economy (e.g., ridesourcing) takes advantage of the connectivity between users, the internet, and mobile apps to enable individuals and businesses to monetize mobility services in new ways (USDOT 2017). Despite—or perhaps because of—these interdependencies, social systems are not yet well-defined from a quantitative engineering perspective. Urban cities are complex systems with highly interdependent social, infrastructure (inclusive of physical and natural systems), and economic systems; the inability to quantify human behavior within these systems inhibits the ability of academics, professionals, and communities to confront and mitigate these grand challenges.

1.1.1 Grand Challenges

The rapid pace of urbanization poses a global challenge. In 1950, 30% of the world's population lived in cities. In contrast, 2010 marked the tipping point in which half of the world's population resided in cities; this number is expected to increase to 68% by 2050 (UN 2019). Although today's urbanization contributes to increased productivity and economic growth while facilitating increased system connectivity between economic sectors, urbanization comes at a cost. Indeed, urbanization can lead to negative impacts on public and environmental health (e.g., pollution, fast-spreading disease), congestion, climate change, and constrained resource availability. In addition, as more people flock to cities, they place increasing demand on the built environment and its associated infrastructure (UN 2019).

As urbanization trends impose greater load demands on infrastructure, that infrastructure only continues to naturally deteriorate with age. The deterioration of the nation's infrastructure has potentially catastrophic outcomes since the economy, culture, security, and quality of life of populations rely heavily on the performance of urban infrastructure systems. Dams serve as integral means of protection to surrounding communities as well as a service for the economy. Yet 17% of dams have high-hazard potential—meaning their failure will cause significant economic losses—and require $45 billion in repairs (ASCE 2017). Freight railroads operate over 138 thousand rail miles across the United States; these railroads are forecasted to see a 40% increase in freight shipments by 2040 which exceeds existing capacity (ASCE 2017). Treatment plants process wastewater collected from over 1.3 million miles of public and private sewers that are susceptible to overflows, blockages, and structural failure. Wastewater treatment plants are critical to protecting the environment and public health by reducing toxins, and are expected to support demand from over 56 million new users (i.e., a 23% increase) by 2032—this will require a $271 billion investment (ASCE 2017).

The structural health of bridges is particularly important to the nation's well-being due to the interdependencies between transportation, economic, and social systems. Age-induced damage and deterioration acutely affects bridges nationwide. Nearly 10% of the nation's 614 thousand bridges require significant maintenance, rehabilitation, or replacement efforts (ASCE 2017). According to National Bridge Inventory data, the average bridge in the United States is 43-years old, just shy of the average intended design life of 50 years (ASCE 2017). The Federal Highway Administration (FHWA) continues to rely on the nationally mandated National Bridge Inspection (NBI) program to manage inspection schedules, implement inspection quality control, and establish a common scale for asset management decision making (Frangopol 2001). During regular (e.g., biennial) visual inspections, inspectors rate bridge components and assign condition ratings on a numeric scale to describe the level of deterioration they observe in a bridge component (MDOT 2016). Asset management decisions about infrastructure upkeep and replacement are subsequently made directly from these assigned condition ratings. The frequency of inspections (i.e., months to years, in general) and qualitative nature of assigning discrete condition ratings using visual assessment are insufficient and uneconomical. Based on these inspections, the most recent federal estimate puts a price tag of $123 billion on the nation's necessary bridge rehabilitation efforts. In recent years, annual investments at all government levels have ranged from $11.5 billion in 2006 to $18 billion in 2010 (ASCE 2017). Increasing financial constraints (AASHTO 2007) coupled with increasing travel demands on aging bridge inventories (FHWA 2004) present significant challenges to developing and implementing automated and scalable strategies to continuously quantify damage in bridges and improve methods of asset management.

At the same time, naturally occurring hazards continue to threaten infrastructure and populations. Climate change is expected to increase the frequency and intensity of extreme weather events, coastal flooding, wildfires, droughts, flooding, and landslides, as well as increase air pollution (USGCRP 2017). The global average annual cost of climate-induced disasters (including, storms, droughts, floods, and heat waves) increased from $64 billion between 1985-1994 to $154 between 2005-2014; natural hazards have a more devastating impact with population growth and rapid urbanization. This cost is expected to be much higher once the slow-onset impacts of climate change can be quantified and accounted for (UN 2016). While these hazards have a direct global impact, they disproportionately affect under-resourced communities and are likely to decrease economic growth, increase food insecurity, have a greater impact on the physical environment, reduce the reliability of local water supplies, and cause health problems, among others (UN 2016). Consequently, climate change and inequality are engaged in a vicious cycle: climate change exacerbates existing underlying inequalities present in society that in turn increase vulnerabilities to natural hazards. This suggests that societal resilience is not just a function of the robustness of the physical environment and the ability of a population to return to normal operational levels after a disruptive event. Resilience is also a function of inequality and access (i.e., inequity), employment, education, discrimination (both cultural and institutional), and a community’s ability to have a voice and representation in governance (UN 2016).

1.1.2 Grand Opportunities

Despite these grand challenges, equally grand opportunities are emerging due to the unprecedented proliferation of low-cost sensing technologies, computational capabilities, and advancements in tools enabling automation and actuation. The explosion of computational power beginning in the mid-1900s revolutionized and expanded the role of computation in engineering. This increase in processing power is best embodied by Moore’s Law in which computational speed doubled every 18 months from the 1960s up until the early 2000s (Strawn and Strawn 2015). This led to the development of low-cost and miniaturized, yet high-performing, sensors that enable sensing in the natural and physical environments including in transportation systems, health care (e.g., wearable devices), and environmental systems, to name a few. Sensing and computation have also been aided by wireless communication that has lowered costs, improved communication latency, permitted varying protocols, facilitated sensor fusion between heterogeneous data types, and enabled cloud computing.

Together, these technological trends have led to the emergence of “intelligent” systems, which comprise the synergistic combination of sensing, computing, and action (e.g., actuation). Within engineering, intelligent systems are generally synonymous with cyber-physical systems (CPS). In civil engineering specifically, CPS enables professionals to drive the performance and resilience of infrastructure and cities. For example, earthquake early warning systems process data collected from distributed strong ground motion systems and initiate protocol to shut down and protect critical infrastructure (e.g., high-speed trains) in the event of an earthquake (Nakamura and Saita 2007). “Smart cities” are emerging in which cloud computing platforms access enormous amounts of urban data from a diverse array of sensor networks and information technology to enhance the experiences of citizens and the performance of urban systems (Ismagilova 2019). Water depth sensors are being distributed throughout watersheds and valves are being used as actuators to control storm water systems (Wong and Kerkez 2017). Additionally, sensing and control play an increasingly important role in harnessing renewable sources of energy (Sil et al. 2017).

1.2 Limitations of Existing Intelligent Systems to Sense for Decisions

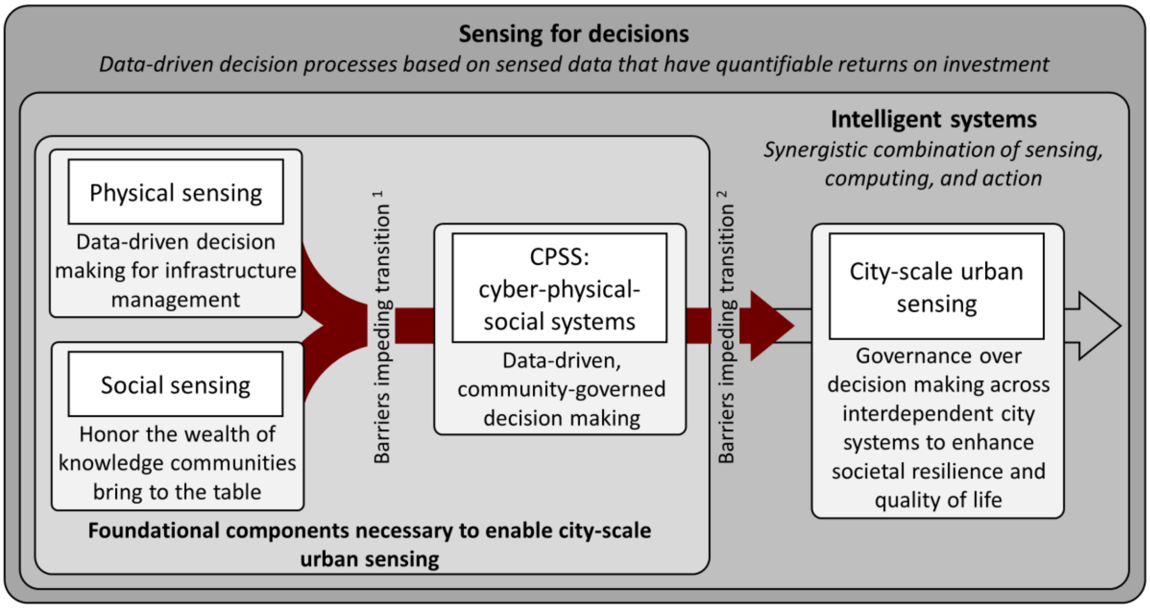
Despite the abundance of technological advances that enable sensing efforts and the implementation of intelligent systems across the civil and environmental engineering domain, the ability to "sense for decisions"—data-driven decision processes based on sensed data that have quantifiable returns on investment—remains unrealized for an entire class of problems. These problems lie on a spectrum of wide-ranging complexity from effectively managing single-asset infrastructure, to managing governance over data in smart cities, to increasing overall objective city-wide societal resilience (it is assumed herein that societal resilience includes rigorous consideration of system dependencies).

Decision-making strategies functioning within the most complex problems on this spectrum—that is, objective societal resilience in urban cities—can only be realized once scalable technological and analytical strategies are able to quantify the complex interdependencies between physical, economic, and social systems. These complex interdependencies cannot be rigorously quantified until the system components are first well understood. Yet, methodologies that truly enable sensing for decisions in many critical physical and social systems have yet to be realized even on a component level. For example, infrastructure management lacks universal structural health monitoring (SHM) strategies and post-processing algorithms that can quantitatively, continuously, and economically inform maintenance, repair, and replacement decision-making strategies used in practice today (Cawley 2018). Despite the proven track record of effective decisions being made by condition ratings assigned during visual inspections, proposed SHM methodologies continue to operate independent of condition ratings. At an urban system scale, there is a reliance on the government sector (in collaboration with the private sector) to select sensors, deploy sensors, and be the first consumers of urban data (Hancke et al. 2013). As a result, data does not benefit communities equally across smart cities as there is limited community-governed decision making before and after deployments. Without a community-based cyber-physical-social system (CPSS) architecture to expand the scope of governance over data to enable members of under-resourced communities to collect data, generate insights, and actuate change on system-wide issues that tie directly to their needs, these communities are likely to be left out of these processes. Already, these communities are being left out of the majority of urban sensing efforts (Hancke et al. 2013).

Figure 1-1 illustrates the necessary progression that must occur to better enable a future of decision making across city-scale urban sensing systems comprised of interdependent physical, economic, and social systems. This progression includes gaining a better understanding of sensing for decisions in physical and social systems, CPSS at the intersection of physical (CPS) and social systems, as well as overcoming the technological and analytical barriers to scaling out. Section 1.2.1 and Section 1.2.2 provide a deeper discussion of the limitations that inhibit sensing for decisions in these foundational domains and identify the primary obstacles that must be overcome.

1.2.1 Decision Making for Infrastructure Asset Management

For structures like bridges, upkeep decisions are based on federally mandated NBI condition ratings assigned during periodic (e.g., biennial) visual inspections. During visual inspections, inspectors rate bridge components and assign condition ratings on a numeric scale (e.g., 0 through 9) to describe the level of deterioration they see in the component (MDOT 2016). This qualitative method of inspection fosters reactive decision making in which repairs are made based on periodic visual examination of existing damage or



[1] Increased resource, community, and user constraints; scalability; usability
[2] Governance; large-scale, interdependent system actuation; fusion of heterogeneous data types

Figure 1-1. Illustration of the progression from enabling sensing for decisions at the level of single-asset infrastructure to the city scale, including primary barriers.

deterioration. In this approach, damage is crudely described by discrete condition states assigned by an inspector. Furthermore, visual inspections may not uncover underlying deterioration (Agrawal et al. 2009). Due to these drawbacks, there is a general consensus within the engineering community that asset management decisions could be aided by quantitative monitoring data in order to reduce uncertainty in assessing structural condition (Frangopol et al. 2001). Monitoring can also be continuous, allowing asset managers to more regularly, if not continuously, track the condition of a structure, especially between biannual inspections.

Consequently, the proverbial "holy grail" of SHM is the quantitative linkage between data and decisions to guide continuous and quantitative asset management decisions in practice. SHM entails installing sensors on civil infrastructure to monitor structural responses to loadings and infer their health (i.e., detect damage). SHM of aging infrastructure has shown continued growth in academia as well as in the private sector, and monitoring systems have been installed on structures for decades (Seo et al. 2016, Lynch et al. 2004, Kim et al. 2007, Kurata et al. 2012, Jang et al. 2010, Feltrin et al. 2010). Despite these efforts, there are very few examples of true SHM systems being used for widespread decision making today (Cawley 2018).

This discontinuity is in part due to the historical SHM paradigm cast as a problem of estimating structural state and detecting damage by monitoring changes in structural properties such as modal properties. The fundamental principle of frequency-based damage detection methods is that modal parameters are functions of physical properties such as mass and stiffness. Vibration-based data collected from SHM systems have been proposed to be used to quantify changes in modal properties associated with damage-induced changes in mass, or more commonly, stiffness. For most operational structures, long-term damage and deterioration may not necessarily correspond to mass or stiffness changes (i.e., structures often operate in their elastic regimes even when deteriorated), making many of these physics-based SHM methods ineffective and potentially misleading. Because existing SHM methods largely operate independent of condition ratings, the state-of-the-art (i.e., SHM) and the state-of-practice (i.e., visual inspection) remain two independent processes, relying on visual inspection to maintain the nation's critical infrastructure. Rather than operating as two separate processes, there is a need to develop an SHM framework that reconciles the state-of-practice and the state-of-the-art by formally linking structural monitoring data with condition ratings; this respects the proven track record of effective decisions being made by condition ratings assigned during visual inspections.

1.2.2 Community-Governed Decision Making

While the scientific foundation and engineering of CPS has rapidly advanced in recent years (NIST 2016), comparatively less research has been devoted to the direct inclusion of the human element within the CPS framework. So termed cyber-physical-social systems (CPSS), these expanded CPS architectures explicitly integrate humans including their ability to observe and take action in relation to physical and cyber CPS elements (Wang 2010). The advancement of CPSS architectures has been explored in a variety of applications including in intelligent transportation systems, building automation systems, and health care (Crowley et al. 2013, Zhang et al. 2017). While CPSS holds tremendous promise in connecting urban citizens to their cities, there remain many technical challenges including how to integrate human-centric data and human-in-the-loop control solutions into CPS frameworks.

Equally—or perhaps more—important are the large number of social challenges that must be surmounted before the implementation of CPSS solutions in urban environments. The proliferation of sensors and the aggregation of urban data has disproportionately focused on improving city cores with less attention paid to residential areas. A related issue is that city governments and the private sector are aggregating large amounts of data that allow them to be more efficient in governing and conducting business within cities. This inherent asymmetry in control over data has limited community engagement in governance over data and decision making both before and after system deployments. One particularly relevant example is nationwide community push-back on data-driven policing. While data-driven (predictive) policing may reduce policing costs, it has been questioned as an indirect means of racial profiling (Patel 2015). Consequently, the data collected in emerging smart city applications is not always benefitting society equally and ultimately failing to increase community resilience and facilitate social mobility (i.e., enable access) in vulnerable, under-resourced communities.

A CPSS framework that empowers community representation in governance over data (both its collection and use) is necessary to enable decision making in any city, but especially in under-resourced communities. Community-governed decision making honors the wealth of knowledge and culture that vulnerable populations bring to the table, and empowers citizens to work in a more meaningful partnership with city governments leading to more resilient modes of governance. There are several technological obstacles that impede the emergence of a successful general-purpose urban sensing architecture for widespread use. These include the development of a sensing platform that supports interoperability among diverse arrays of heterogeneous internet of things (IoT) devices, preserves privacy and trust among citizens, is user friendly and accessible to even the members of society who are least prepared to adapt to new technologies, supports cloud-based analytics, and supports low-power and low-cost sensing and communication, which is particularly difficult to achieve with platforms that require a continuous source of energy (Mehmood et al. 2017). For example, the Array of Things (AoT) deployment in Chicago provides environmental and air quality sensors tied to a Linux-based sensing node (AOT 2015). The high-power demand of the hardware requires access to power sources supplied by street furniture (e.g. light poles) which limits its deployment potential in less populated areas as well as in residential neighborhoods where power sources are severely limited. In addition, Placemeter is a novel camera-based sensing solution designed to track vehicles and pedestrians in city spaces (Placemeter 2015). While Placemeter promises to anonymize data by processing video footage using automated data processing tools, many communities might be resistant to camera-based monitoring due to infringements on privacy and overarching notions of “Big Brother” within the community.

1.3 Research Goals and Objectives

In response to the limitations presented in Section 1.2, this dissertation builds a rigorous technological, analytical, and methodological background necessary to sense for decisions in response to two of the field’s prevailing challenges: infrastructure asset management and community-engaged decision making. An emphasis is placed on addressing the foundational components necessary to enable future city-scale urban sensing efforts (Figure 1-1); this includes consideration of both CPS (i.e., data-driven decisions in the context of physical systems) and CPSS (i.e., data-driven decisions in the context of community needs) applications. The main challenges and goals of the research presented herein are developed in detail in Section 1.3.1 through Section 1.3.5.

Figure 1-2 outlines the primary research objectives of this dissertation and illustrates the deliberateness of the path taken. The first objective of this dissertation is to consider the management of single-asset infrastructure and use long-term monitoring data to assess structural condition using reliability methods with respect to a single failure limit state. Here, outputs trigger maintenance decisions to improve safety. In addition to developing the methodological framework, this work is implemented on the Harahan Bridge using 16 months of long-term response data collected on the bridge using a wireless monitoring system. Second, this framework is generalized to more formally and broadly link SHM data with condition ratings, which are the starting point for bridge management decision-making processes used across the United States today. In this work, “lower” limit

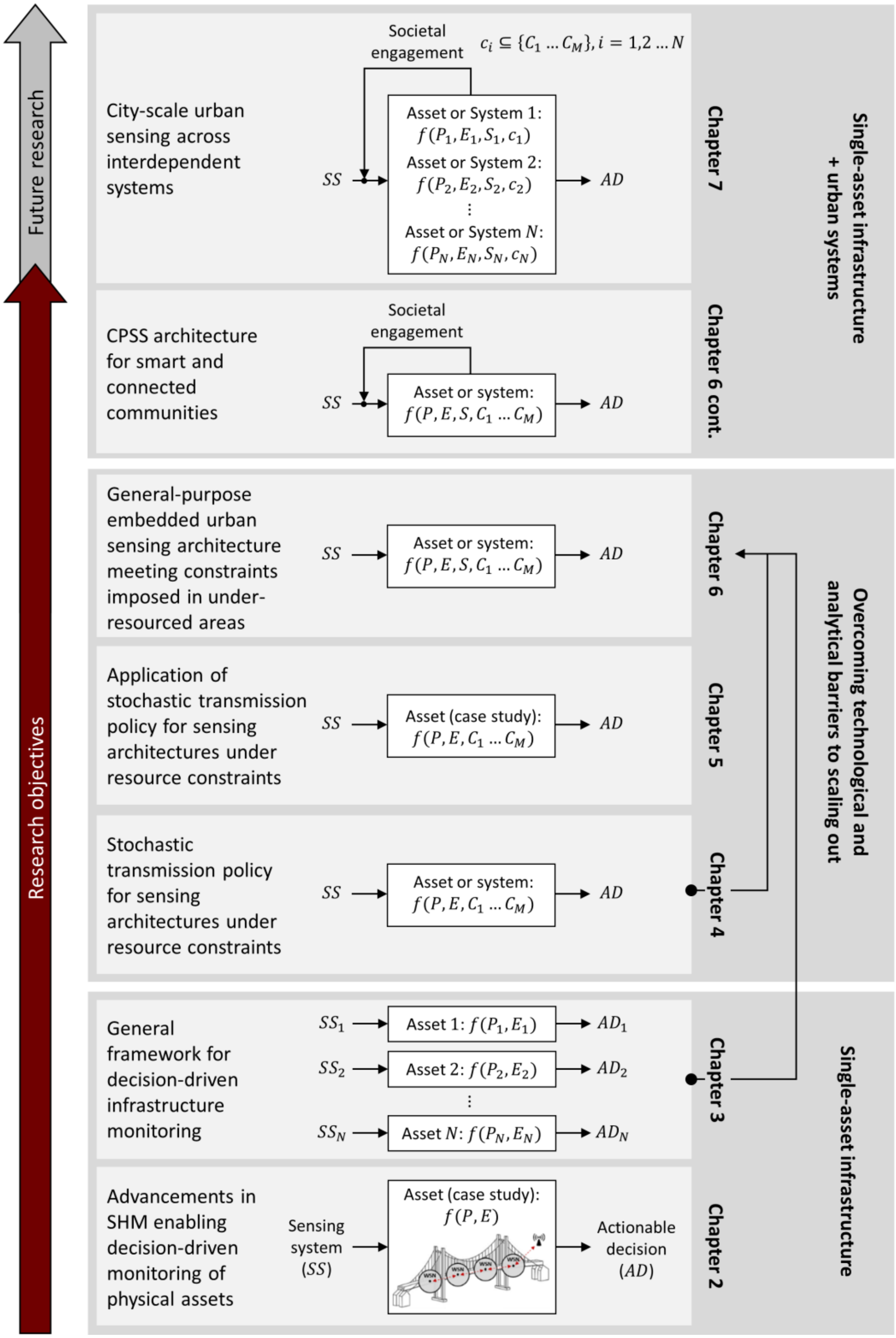


Figure 1-2. Primary research objectives.

states are introduced within the reliability framework to characterize structural performance and rationally map data to condition rating scales. This framework is validated using a wireless sensing system installed on the Telegraph Road Bridge (TRB) (Monroe, MI). Third, this dissertation addresses the difficulties associated with scaling out wireless sensing networks that are subject to stringent energy constraints. Specifically, this dissertation describes the development of a stochastic policy controlling data collection and transmission that optimizes the operation of a sensing node to improve the quality of the parameter estimates of a measured process used in decision-making strategies. Fourth, this optimal policy is implemented using one year of data collected from the Telegraph Road Bridge (Monroe, MI) to illustrate how it improves the quality of the decision-making strategy linking data to condition ratings. Fifth, in addition to addressing the resource constraints that are barriers to scaling out cyber-physical systems, this work addresses user constraints and provides an overview of the development of a low-power embedded wireless sensing architecture that functions as a user-friendly sensing solution accessible to diverse stakeholders. Last, the developed embedded wireless sensing architecture is used to extend the sense-for-decision frameworks developed for cyber-physical systems to those of cyber-physical-social systems by engaging social systems as users and actuators. Through this work, data is "democratized" and feedback mechanisms are enabled by the introduced technology to give citizens full autonomy over sensing and governance over data and decision making in their communities. The following notation is used in Figure 1-2: physical systems are denoted P, economic systems are denoted E, social systems are denoted S, and constraints (either physical, economic, or social) are denoted C.

1.3.1 Advancements in SHM Enabling Decision-Driven Monitoring of Physical Assets

In Chapter 2, this dissertation begins by illustrating how state-of-the-art SHM technologies can be used to sense for decisions with respect to a real-world problem that has far-reaching physical, economic, and social implications for the United States. The 103-year old Harahan Bridge (Memphis, TN) is a critical piece of railroad infrastructure linking Union Pacific Railroad's nation-wide operations on the east side of the Mississippi River in Tennessee to operations on the west side of the river in Arkansas. Regular maintenance of the bridge is a primary concern for the owner because the Harahan Bridge is a fracture-critical, long-span railroad bridge exposed to multiple hazards. Fatigue is a primary concern for the bridge owner because railroad bridges have high live load to dead load ratios and high stress cycle frequencies. However, existing inspection methods and post-inspection analyses are unable to accurately consider the full influence of bridge behavior on the fatigue life of bridge components.

This dissertation develops a reliability-based SHM framework that provides structural owners with actionable information associated with the safety of their assets given the current (potentially deteriorated) condition of their structures. The purely data-driven framework proposed relies on long-term monitoring data to offer a quantitative assessment of structural safety. This framework advances on prior work in the integration of SHM data and reliability methods by tailoring the SHM sensing strategy to offer response data that can reveal the complete behavior of a structural component including behavior at the component boundaries. This minimizes the need to make *a priori* assumptions of ideal component behavior (assuming ideal boundary conditions) that can lead to inaccurate reliability analysis results.

A case study is included to illustrate the proposed probabilistic fatigue assessment methodology for monitored railroad bridge components using only continuous, long-term response data in a purely data-driven reliability framework that is compatible with existing inspection methods. As an illustrative example, this work quantifies the safety profile (corresponding to fatigue life) of a fracture-critical assembly comprising of six parallel eyebars on the Harahan Bridge. The monitored eyebars are prone to fatigue because changes in the boundary conditions can cause some eyebars to carry a greater proportion of the total assembly load than assumed during design and analysis; existing manual inspection practices aim to maintain an equal loading distribution across the eyebars. Consequently, the limit state function derived in this analysis accounts for the coupled behavior between fatigue and relative tautness of the parallel eyebars as well as dependencies between the carried eyebar loads. The reliability index values for both the element (i.e., individual eyebars) and system (i.e., full eyebar assembly) reliability problems are assessed and indicate that, under the conservative assumption that progressive failure is brittle, first failure within the parallel eyebar system is generally equivalent to system failure. The proposed method also serves as an intervention strategy that can quantify the influence of eyebar realignment maintenance efforts on the future evolution of the reliability index.

1.3.2 General Framework for Decision-Driven Infrastructure Monitoring

Recall from Section 1.1.1 that for structures like bridges, upkeep decisions are based on federally mandated condition ratings assigned during visual inspection. During regular (e.g., biennial) visual inspections, inspectors rate bridge components and assign condition ratings on a numeric scale to describe the level of deterioration they observe in a bridge component (MDOT 2016). Asset management decisions about infrastructure upkeep and replacement are subsequently made directly from these assigned condition ratings. The frequency of inspections (i.e., months to years, in general) and qualitative nature of assigning discrete condition ratings using visual assessment are insufficient and uneconomical. Consequently, the ultimate goal of SHM is to augment visual inspection by explicitly quantifying the linkage between monitoring data and decisions. Quantitative monitoring data reduces uncertainty in assessing structural condition and, when collected continuously, allows asset managers to more regularly, if not continuously, track the condition of a structure (Frangopol et al. 2001).

The second objective of this dissertation is to draw on the advancements made—and lessons learned—in developing the decision-driven monitoring system implemented on the Harahan Bridge to develop a more generalized SHM framework to sense for decisions across diverse infrastructure assets. While SHM has shown continued growth over the past several decades, there is a persistent chasm between SHM and the ability of structure owners to make asset management decisions based on SHM data. This is in part due to the historical SHM paradigm cast as a problem of estimating structural state and detecting damage by monitoring changes in structural characteristic properties (namely, reduced stiffness). For most operational structures, deterioration does not necessarily correspond to changes in structural properties with structures operating in their elastic regimes even when deteriorated. This is reflected in the United States' extensive history of codified decision-making practices in which actionable decisions on upkeep are made entirely in the elastic regime based on condition ratings (MDOT 2016).

Since condition ratings are widely accepted in practice and actionable decisions are made directly from assigned condition ratings, this dissertation proposes that, rather than operating as two separate processes, the state-of-practice and SHM be reconciled such that structural monitoring data be explicitly linked to condition ratings. This approach respects that 1) there is a proven track record between effective decisions being made by condition ratings assigned by visual inspections, and 2) quantitative monitoring data can reduce uncertainty in assessing structural condition. In order to link these two processes, condition ratings serve as lower limit states (i.e. limit states below yielding) with long-term monitoring data used to quantify these lower limit states in terms of the reliability index. A primary objective of this work is to develop a method to quantify the reliability index values corresponding to the lower limit states described by condition ratings. Once the reliability index threshold values are established, the data-driven reliability index of the in-service asset can be monitored continuously and explicitly mapped to a condition rating at any time. As an illustrative example, the proposed framework for tracking structural performance is implemented in full using long-term monitoring data collected on a pin-and-hanger assembly on the Telegraph Road Bridge (TRB), a highway bridge located in Monroe, MI.

This chapter illustrates the development and successful implementation of the first ever SHM framework to explicitly map monitoring data to actionable decisions that are made in practice today. This human-independent, data-driven decision-making strategy allows monitoring data to be used to trigger decisions based on quantitative information that encompasses the entire measurable domain of damage that may exist in an asset, as opposed to those only informed by visual inspection. A quantitative and continuous method of assigning condition ratings based on the probability of failure also facilitates a risk-based future of asset management that extends existing decision-making strategies to those that are economically informed.

1.3.3 Stochastic Data Collection and Transmission Policy for Resource-Constrained Sensing Architectures

Recall from Figure 1-1 that limited access to energy presents a number of technical and analytical barriers to scaling out wireless sensing networks (WSNs) to enable city-scale urban sensing. Historically, the government and private sectors have focused on sensing in urban centers, which have access to fixed power sources (e.g., light poles, powered street furniture, outlets installed by the city). Outside of city centers, the lack of access to public power sources in residential communities severely limits the potential to carry out mobile and stationary sensing node deployments. This lack of connection between citizens and smart city initiatives is particularly pronounced in depopulated American cities (e.g., Detroit, Flint, St. Louis, Baltimore) where there has been no notable success in using smart city technologies to connect populations to their larger communities. In the pursuit of democratizing data within a greater societal resilience framework, there is a need to overcome the technical and analytical barriers associated with sensing in energy-constrained environments. Such sensing efforts will require that WSNs rely on energy harvested from the environment and necessitate consideration of energy as an uncertain and limited resource.

Since decision-making practices are often reliant on statistical representations of measured processes as inputs, the limited availability of energy in WSNs is particularly problematic when WSNs are used for remote parameter estimation because only a subset of the measured information can be wirelessly transmitted to the estimator for processing. This leads to a tradeoff between the quality of the parameter estimates and the energy-constrained communication. While much attention has been paid to 1) communication schemes for energy-aware data transmission in WSNs under resource constraints and 2) controlled parameter estimation over WSNs separately, there has yet to emerge a stochastic data collection and transmission policy that minimizes the variance of estimated component parameters of a measured process subject to constraints imposed by a WSN node's energy and data buffer sizes, stochastic models of the incoming energy and event arrivals, the value of data, and temporal death. In order to enable the implementation of WSNs in energy-constrained communities and areas, this dissertation presents the derivation and implementation of just such an optimal policy. By controlling data collection within a transmission subsystem to optimize an objective governed by remote parameter estimation, data collection and transmission is automated to facilitate decision making. This work models a replenishable WSN node as a continuous-time Markov chain and derives a single unique threshold value governing an event-based policy that is independent of the energy buffer's state of charge (SoC) and places no restrictions on the size of the energy and data buffer sizes. The derived optimal threshold value produces the best possible estimate (i.e., minimum component variance) of the process parameters using Maximum Likelihood Estimation (MLE) given the system constraints. Numerical results are presented to reflect the objectives of the theoretical results and to illustrate that the proposed framework is robust against uncertainty in estimates of the process' parameter estimates at the outset of the monitoring period. Implementation of the proposed optimal stochastic policy within SHM and urban systems enables decision makers and stakeholders to make decisions and take action as frequently and accurately as possible given the system constraints.

This work differentiates itself from, and improves upon, existing literature in two notable ways. First, the proposed policy accounts for—and places no restrictions on the size of—a WSN architecture's data storage buffer. Data transmission is often the most significant source of energy consumption in a wireless sensing node; accounting for the storage of (potentially) large amounts of data in a buffer that are communicated to the remote estimator in batch transmissions—as opposed to transmitting each value individually immediately upon collection—leads to significant gains in the transmission rate. Consequently, the proposed policy can be used in diverse applications requiring wide-ranging hardware specifications, from sampling continuous measured processes at high rates (>100 Hz) in SHM applications to low-rate environmental applications. Second, this work represents the first effort to derive a stochastic policy that governs data collection and transmission with the purpose of explicitly linking optimal remote parameter estimation to a rigorous stochastic energy recharging model that accounts for all significant hardware constraints.

1.3.4 Optimal Stochastic Data Collection and Transmission Policy for Self-Sustaining SHM Systems Used for Asset Management

Data acquisition methods for SHM—such as those presented in Section 1.3.1 and Section 1.3.2—historically rely on schedule-based or transmit-all data collection strategies. In the context of this dissertation, a deterministic schedule-based policy refers to a WSN architecture that transmits all data collected periodically on a pre-defined and regular schedule with fixed monitoring periods and fixed periods between monitoring. A transmit-all policy refers to the attempted transmission of all measurements. For sensing systems that are self-sustaining (e.g. those relying on harvested energy as described in Section

1.3.3), these data collection methods are unable to explicitly account for the availability of energy and the value of data. As a result, sensing systems often fail to capture and transmit key data to end users and require an excessive amount of time to characterize statistical parameters describing response data. As structural monitoring data is being increasingly incorporated into decision-making processes for asset management using reliability methods, there is a need for an automated data collection and transmission strategy that facilitates the characterization of the statistical parameters of structural response data with minimum variance so that bridge owners can frequently and accurately track structural condition. This work revisits the stochastic data collection and transmission policy introduced in Section 1.3.3 and extends the optimal data collection and transmission policy to an SHM application on the Telegraph Road Bridge, which is the standard pin-and-hanger steel girder highway bridge located in Monroe, MI that is considered in Section 1.3.2. The TRB's monitoring system utilizes sensing nodes that operate using harvested solar energy and are subject to stringent energy constraints due to the size of the solar panels and availability of incoming energy, geographic location, and battery size. This work advances the general framework for decision-driven infrastructure monitoring described in Section 1.3.2. Since the optimal data collection and transmission policy minimizes the component parameter estimates of a measured process, implementation of the optimal policy on the TRB enables monitoring data to be used to assign condition ratings more frequently and accurately than the existing SHM system installed on the bridge. Numerical results based on one year of monitoring data are presented and illustrate the gains achieved using the optimal policy as compared to the best schedule-based and transmit-all policies.

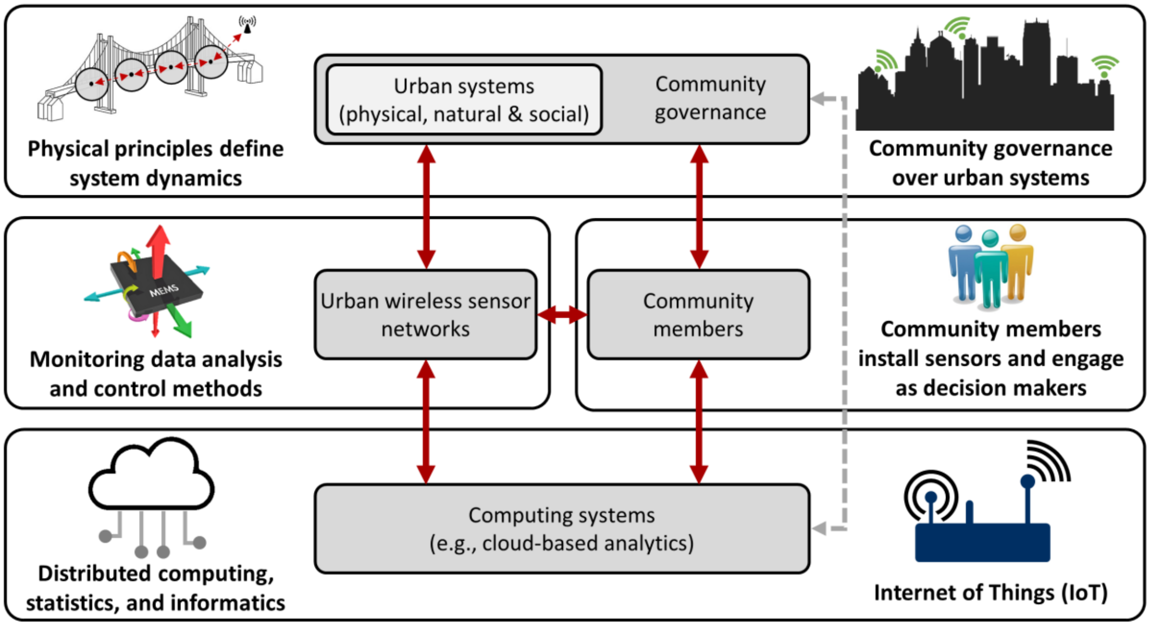


Figure 1-3. CPSS framework for community-governed sensing.

1.3.5 CPSS Architecture for Smart and Connected Communities

To fully and more broadly reap the promise of smart cities, there is a need to expand the use of sensing, especially by the general public, through the development of a more democratized approach to urban data collection and post-collection data uses. This, in part, entails engaging urban youth and citizens with their communities by architecturally embedding them within a smart city's urban CPSS. The third objective of this dissertation is to develop a community-governed CPSS framework (Figure 1-3) that expands a CPS architecture to directly integrate human action by taking into account a citizen's ability to observe and take action in response to CPS elements (Wang 2010).

At the core of this framework is an embedded urban sensing architecture developed by the author. This platform, termed *Urbano*[1], was designed to overcome the technological

---

[1] Urbano: From the Latin urbanus ("of or belonging to a city"), derived from urbs ("city")

challenges that have impeded the emergence of a successful general-purpose urban sensing architecture for widespread use; these challenges emerge from the additional constraints imposed in under-resourced communities. This includes the development of a sensing platform that supports interoperability among diverse arrays of heterogeneous IoT devices, preserves privacy and trust among citizens, supports cloud-based analytics, has a user-friendly design, and supports low-power and low-cost sensing and communication to free nodes from a fixed power source; the latter is especially difficult to achieve with platforms that require a continuous source of energy (Mehmood et al. 2017).

The proposed CPSS framework is implemented at full-scale in Detroit, MI. While most urban cities in the United States have undergone sustained periods of economic growth since the 1940s, a small number of cities have seen dramatic drops in population and economic activity (Ryan 2012). For example, Detroit experienced population reductions from 1.8 million people in 1950 to less than 700,000 in 2015 (The Detroit News 2016). While major revitalization efforts in Detroit's business core are starting to rapidly transform the city, the residential areas and neighborhoods of Detroit remain under-populated with scarce access to important city services. Shrinking cities have resulted in extreme levels of poverty and inequality that result in stressors that disproportionately impact urban youth, who are at risk of losing connectivity to their cities and communities (Luthar 1991). By involving Detroit youth (grades 8-12), the proposed CPSS framework aims to encourage young, vulnerable populations to engage with their communities as citizen scientists. Youth engage in problem-based inquiry, devise sensing solutions, and interpret collected data drawn from their own urban settings. By empowering communities to collect their own data in their neighborhoods and cities using the *Urbano* platform, city governments, local organizations *and* citizens work in a more meaningful partnership with each other, leading to more resilient modes of smart city governance. The program described herein includes an assessment of the effectiveness of the CPSS framework in engaging citizens in governance over their urban systems.

1.4 Organization of the Dissertation

This dissertation is organized as follows:

- Chapter 2 demonstrates how sensing strategies can be aligned with decisions governing infrastructure management by extending the role of reliability methods to quantify system performance. Here, the reliability index is used as a scalar measure of the safety (i.e., probability of failure) that is extracted from monitoring data to assess structural condition relative to a failure limit state. As an example, 16 months of long-term data collected from an SHM system installed on the Harahan Bridge (Memphis, TN) is used in a reliability framework to track the fatigue life of critical eyebar assemblies.
- In Chapter 3, the proposed reliability-based SHM framework from Chapter 2 is generalized to formally and more broadly link SHM data with condition ratings because inspector-assigned condition ratings remain the primary starting point for asset management decisions made in practice today. While reliability methods historically quantify safety with respect to a single failure limit state, this chapter demonstrates that there exist measurable reliability index values associated with "lower" limit states below failure that more richly characterize structural performance and rationally map to condition rating scales. Consequently, monitoring data can be used to assign condition ratings based on quantitative information encompassing the measurable

damage domain, as opposed to relying on visual inspection. This work reflects the first-ever SHM framework to explicitly map monitoring data to actionable decisions and is validated using a wireless SHM system installed on pin-and-hanger assemblies on the Telegraph Road Bridge (TRB) (Monroe, MI).

- Chapter 4 addresses one of the foremost challenges faced by solar-powered WSNs: stringent resource constraints due to the limited and uncertain availability of energy. For decision-making processes relying on statistical estimation of performance, the utility of data should be considered to optimize the data collection process given these constraints. This chapter proposes a novel stochastic data collection and transmission policy for WSNs that minimizes the variance of a measured process' estimated parameters subject to constraints imposed by energy and data buffer sizes, stochastic models of energy and event arrivals, the value of measured data, and temporal death.
- Chapter 5 draws on the theoretical basis presented in Chapter 4 and extends the optimal data collection and transmission policy to an SHM application on the Telegraph Road Bridge, which is introduced in Chapter 3. The monitoring system on the TRB utilizes sensing nodes that operate using harvested solar energy and are subject to stringent energy constraints due to the size of the solar panels and availability of incoming energy, geographic location, and battery size. Numerical results are presented based on one-year of data collected from the TRB and illustrate the gains achieved by implementing the optimal policy to obtain response data used to estimate the reliability index.

Table 1-1. Breakdown of the components necessary to enable objective societal resilience.

	Physical system component	Economic system component	Social system component	Resource constraints	Community and user constraints	City-scale urban sensing
Chapter 2	✓	✓				
Chapter 3	✓	✓				
Chapter 4	✓	✓		✓		
Chapter 5	✓	✓		✓		
Chapter 6	✓	✓	✓	✓	✓	
Chapter 7	✓	✓	✓	✓	✓	✓

- Chapter 6 extends the work performed in wireless sensing and sense-for-decision frameworks by exploring their role in community-based decision making. This chapter poses societal engagement as a necessary entry point to urban sensing efforts in smart cities. Members of under-resourced communities are often stakeholders who are the challenged with technology and those most vulnerable to lack of access to data and information. This chapter offers a novel, low-power wireless sensing architecture that functions as a user-friendly urban sensing solution that communities can rapidly deploy to understand urban processes that are critical to informing their views and guides their decision making. Applying this platform, this chapter presents transformative work to "democratize" data by enabling members of vulnerable communities to easily use these generic urban sensors that can collect data and generate insights on issues that tie directly to their needs and the diverse ways in which they use their cities.
- Chapter 7, the final chapter, presents a dissertation summary, overview of key intellectual contributions achieved, and a discussion focused on future extensions of the research.

Table 1-1 presents a breakdown of the physical, economic, and social components, as well as the barriers to scaling out, that each chapter addresses.

1.5 References

Agrawal A, Kawaguchi A and Qian G. Bridge element deterioration rates: phase I report. Report, New York State Department of Transportation (NYSDOT). Project# C-01-51, 2009. Albany: NYSDOT.

American Association of State Highway and Transportation Officials (AASHTO). *Transportation - invest in our future: future needs of the U.S. surface transportation system*. Washington, DC: AASHTO, 2007.

American Society of Civil Engineers (ASCE). Failure to act: closing the infrastructure investment gap for America's economic future. Report, Economic Development Research Group, 2016.

American Society of Civil Engineers (ASCE). 2017 report card for America's infrastructure. Report, ASCE, 2017.

Array of Things (AOT). Array of things, https://arrayofthings.github.io/ (2020, accessed 1 February 2016)

Cawley, P. Structural health monitoring: closing the gap between research and industrial deployment. *J Struct Health Monit* 2018; 17(5): 1225-1244.

Crowley DN, Curry E and Breslin JG. Closing the loop—from citizen sensing to citizen actuation. In: *7th IEEE international conference on digital ecosystems and technologies*, Menlo Park, CA, 14-16 July 2013, pp. 108-113. IEEE.

The Detroit News. Detroit population rank is lowest since 1850, www.detroitnews.com/story/news/local/detroit-city/2016/05/19/detroit-population-rank-lowest-since/84574198 (2016, accessed January 8 2018).

Feltrin G, Meyer J, Bischoff R, et al. Long-term monitoring of cable stays with a wireless sensor network. *Struct Infrastruct Eng* 2010; 6: 535-548.

Federal Highway Administration (FHWA). 2004 status of the nation's highways, bridges, and transit: conditions and performance. Report, FHWA, 2004.

Frangopol DM, Kong JS and Gharaibeh ES. Reliability-based life-cycle management of highway bridges. *J Comput Civ Eng* 2001; 15(1): 27-34.

Guidotti R, Gardoni P and Rosenheim N. Integration of physical infrastructure and social systems in communities' reliability and resilience analysis. *J Reliab Eng Syst Saf* 2019; 185: 476-492.

Hancke GP, de Carvalho e Silva B and Hancke GP Jr. The role of advanced sensing in smart cities. *Sens* 2013; 13(1): 393-425.

Ismagilova E, Hughes L, Dwivedi YK, et al. Advances in research – an information systems perspective. *Int J Inf Manage* 2019; 47: 88-100.

Jacobs J. *The death and life of great American cities*. New York: Random House, 1961.

Jang S, Jo H, Cho S, et al. Structural health monitoring of a cable-stayed bridge using smart sensor technology: deployment and evaluation. *Smart Struct Syst* 2010; 6: 439-459.

Kim S, Pakzad S, Culler D, Demmel J, et al. Health monitoring of civil infrastructures using wireless sensor networks. In: *6th international symposium on information processing in sensor networks*, Cambridge, MA, 25-27 April 2007, pp. 254-263. IEEE.

Kurata M, Kim J, Lynch JP, et al. Internet-enabled wireless structural monitoring systems: development and permanent deployment at the New Carquinez suspension bridge. *J Struct Eng* 2012; 139: 1688-1702.

Luthar SS. Vulnerability and resilience: a study of high-risk adolescents. *Child Dev* 1991; 62(12): 600-616.

Lynch JP, Law KH, Kiremidjian AS, et al. Design and performance validation of a wireless sensing unit for structural monitoring applications. *Struct Eng Mech* 2004; 17(3-4): pp. 393-408.

MDOT (Michigan Department of Transportation). *Bridge safety inspection NBI rating guidelines*. 2016.

Mehmood Y, Ahmad F, Yaqoob I, et al. Internet-of-things-based smart cities: recent advances and challenges. *IEEE Commun Mag* 2017; 55(9): 16-24.

Nakamura Y and Saita J. UrEDAS, the earthquake warning system: today and tomorrow. In: Gasparini P, Manfredi G and Zschau J (eds) *Earthquake early warning systems*. Berlin: Springer, 2007.

National Institute of Standards and Technology (NIST). Foundations for innovation: strategic R&D opportunities for 21st century cyber-physcial systems. Report, Steering Committee for Foundations in Innovation for Cyber-Physcial Systems, 2016. Gaitherburg: NIST.

Patel F. Be cautious about data-driven policing. New York Times, 3 December 2015.

Placemeter. The placemeter sensor, https://www.placemeter.com/ (2020, accessed 29 February 2016).

Ryan BD. *Design after decline: how America rebuilds shrinking cities*. 1st ed. Philadelphia: University of Pennsylvania Press, 2012.

Seo J, Hu JW and Lee J. Summary review of structural health monitoring applications for highway bridges. *J Perform Constr Facil* 2016; 30(4): 04015072.

Sil I, Mukherjee S and Biswas K. A review of energy harvesting technology and its potential applications. *Environ Earth Sci Res J* 2017; 4(2): 33-38.

Strawn G and Strawn C. Moore’s law at fifty. *IT Prof* 2015; 17(6): 69-72.

United Nations (UN). Climate change resilience: an opportunity for reducing inequalities. Report, UN World Economic and Social Affairs. Report no. ST/ESA/363, 2016. UN.

United Nations (UN). World urbanization prospects: the 2018 revision. Report, UN Department of Economic and Social Affairs, Population Division. Report no. ST/ESA/SER.A/420, 2019. UN.

United Nations University (UNU). World risk report 2016: inadequate infrastructure pushes up the risk of disaster. Report, UNU, August 25 2016.

United States Department of Transportation (USDOT). *Beyond traffic: 2045*. Report, USDOT, 2017.

United States Global Change Research Program (USGCRP). 2017 climate science special report: fourth national climate assessment, volume 1. Report, USGCRP, Washington, DC, 2017.

Wang, FY. The emergence of intelligent enterprises: from CPS to CPSS. *IEEE Intell Syst* 2010; 25(4): 85-88.

Wong BP and Kerkez B. Real-time environmental sensor data: an application to water quality using web services. *J Environ Modell Soft* 2016; 84: 505-517.

Zhang Y, Qiu M, Tsai CW, et al. Health-CPS: healthcare cyber-physical system assisted by cloud and big data. *IEEE Syst J* 2017; 11(1): 88-95.

CHAPTER 2

Probabilistic Fatigue Assessment of Monitored Railroad Bridge Components Using Long-Term Response Data in a Reliability Framework

2.1 Introduction

Fatigue is a prevalent deterioration mechanism influencing the safety profile of aging steel structures. Tensile elements in steel truss railroad bridges are especially susceptible to fatigue-induced damage; the high live load to dead load ratio and repeated high stress cycles induced by long and heavy trains can accelerate fatigue damage and cause failure to occur below the allowable stress level (Imam et al. 2006). Over half of the railroad bridges in the United States were constructed prior to 1950 with many of these bridges displaying signs of deterioration as well as approaching their designed fatigue life (Rakoczy et al. 2016). For many of these aging structures, loading estimates have increased compared to the assumed loading when the bridges were first designed (Moreu and Spencer 2015). Consequently, fatigue is a primary concern for railroad owners. Many older steel truss railroad bridges designed before World War II rely on the use of parallel eyebar elements as tensile components in their truss systems to introduce redundancy in critical components. The popularity of eyebar assemblies was due to their structural efficiencies and easy installation during construction. However, there have been a number of notable fatigue-induced eyebar failures such as the catastrophic failure of the Silver Bridge in 1967 (WVDOT 2019) and the fracture of an eyebar element on the San Francisco-Oakland Bay bridge in 2009 (Gostautas and Tamutus 2015, Reid 2010). Despite the threat of fatigue-induced damage on railroad bridge components, existing inspection-based maintenance strategies and remaining fatigue life estimations based *solely* on deterministic *S*-*N* curves and assumptions of total stress cycles and equivalent stress range are unable to accurately consider the full influence of bridge behavior on the fatigue life of railroad bridge components (AREMA 2016).

As railroad infrastructure continues to age, there is a need for asset management decision-making strategies that accurately consider the influence of a wider domain of deterioration mechanisms, such as fatigue. An accurate assessment of the in-service parameters influencing the estimation of remaining fatigue life is necessary to inform upkeep, repair, and replacement decision-making strategies in practice. Reliability-based probabilistic fatigue analysis methods have gained favor in the civil engineering community because uncertainty in parameters such as environmental conditions, material properties, and variable-amplitude stress cycles influence remaining fatigue life estimation (Kunz and Hirt 1994). The use of monitoring data to characterize the distributions of these uncertain inputs can further increase the accuracy of fatigue life estimates (Zhou 2006). Monitoring data accurately reflects the distribution of the live load response of a structure exposed to varying environmental conditions and traffic load patterns. The most accurate estimation of remaining fatigue life requires long-term monitoring over the lifetime of the asset. The field of structural health monitoring (SHM) has advanced structural monitoring technologies to a point where they are an affordable option for asset owners. The research community has embraced the use of long-term monitoring data within reliability frameworks; a number of past studies have proposed data-driven probabilistic fatigue assessment methods for steel bridges using SHM data. However, these prior studies introduce some uncertainties into their analysis methods. For example, Chen et al. (Chen and Xu 2012), Kim et al. (Kim et al. 2001), and Li et al. (Li et al. 2016) use assumed probabilistic models of traffic loading (e.g., train length, weight, and response) and finite element (FE) models to estimate fatigue life. In the case of Li et al. (Li et al. 2016), the assumed loading model is based on short-term controlled loading tests. These fatigue life analyses are not actually based on continuous, long-term measured loads intended to be permanent fixtures on the structure. In addition to introducing uncertainty, the use of approximate methods requiring detailed modeling or controlled experimentation can be a bottleneck that hinders method scalability and inhibits their translation to general, widespread use in practice. Other works, such as Frangopol et al. (Frangopol et al. 2008) and Kwon et al. (Kwon and Frangopol 2010), make assumptions of ideal component behavior and ignore dead load effects. This introduces additional uncertainties when assets have secondary stresses, such as those resulting from unexpected behavior at the boundary conditions or considerable static mean stress.

This chapter presents a probabilistic fatigue assessment framework that can be applied to continuously monitored steel structures. The goal of the work is to develop a reliability-based SHM framework that provides structural owners with actionable information associated with the safety of their assets given the current—potentially deteriorated—condition of their structure. The purely data-driven framework relies on long-term monitoring data to offer a quantitative assessment of structural safety. This framework advances on prior work in the integration of SHM data and reliability methods

(a) (b)

Figure 2-1. (a) View of the Harahan Bridge's truss and parallel tracks spanning the Mississippi River; (b) monitored US0-LS0 tensile truss member consisting of six pin-connected vertical eyebars.

by tailoring the SHM sensing strategy to offer response data that can reveal the complete behavior of a structural component (due to both dead and live loads) including behavior at the component boundaries. This minimizes the need to make *a priori* assumptions of ideal component behavior (i.e., assuming ideal boundary conditions) that can lead to inaccurate reliability analysis results.

As an illustrative example, this work quantifies the safety profile of a fracture-critical set of tensile parallel eyebars on the Harahan Bridge—herein referred to as the US0-

LS0 eyebar assembly—with respect to a limit state function governed by fatigue. The Harahan Bridge opened for operation in July 1916 and is an in-service, long-span, steel truss railroad bridge spanning the Mississippi River near Memphis, TN (Figure 2-1). The US0-LS0 assembly was selected to be monitored continuously because it is fracture critical and changes in the boundary conditions at the pin-eyebar connections can cause some of the eyebars to carry a greater proportion of the total assembly load than originally designed for; this leads to accelerated fatigue damage accumulation. This work leverages the fact that these assemblies are already a primary concern for the bridge owner and tailors a sensing strategy to measure the safety of the assemblies in a way that is consistent with existing decision-making strategies. Consequently, the probabilistic fatigue analysis presented in this chapter accounts for the coupled behavior between fatigue and relative tautness of the eyebars as well as the practical system-level constraints imposed by the mechanics of the system. This probabilistic framework is purely data driven and carried out without the need for approximate methods (e.g., train parameter estimates, FE model simulations, controlled loading tests) during the monitoring period. While an FE model is used to confirm the bridge owner's concern that the US0-LS0 assemblies are fracture critical and a high priority, the FE model presented herein is *not* used within the proposed probabilistic methodology. The study illustrates the use of the proposed data-driven probabilistic fatigue analysis using 16 months of long-term monitoring data collected using a wireless sensing system installed by the author on the Harahan Bridge. While the first 16 months of data are used for the illustrative purposes of this chapter, the installed long-term monitoring system is intended to be a permanent fixture on the bridge and was turned

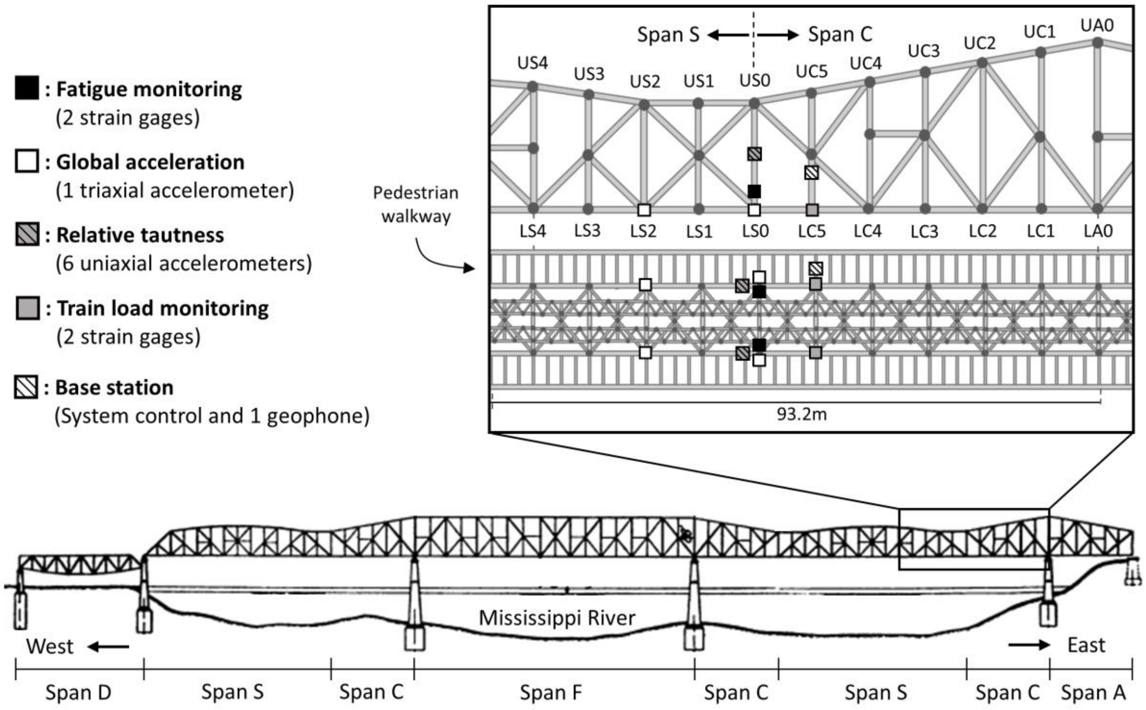


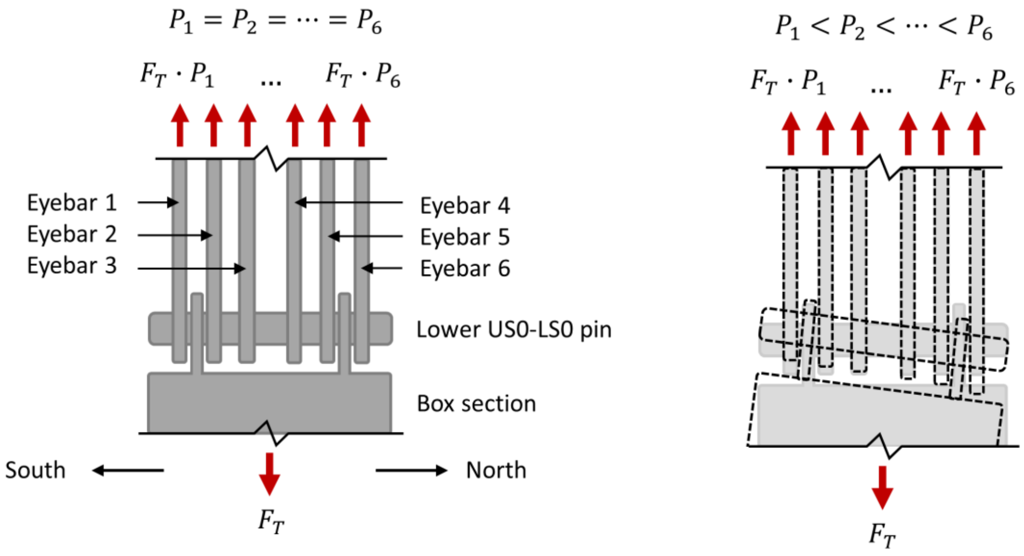
Figure 2-2. Span design and instrumentation plan for the long-term monitoring system installed on the Harahan Bridge (Memphis, TN).

over to the bridge owner for use in their management processes upon completion of the project.

2.2 Harahan Bridge Monitoring System

2.2.1 Harahan Bridge Description

The Harahan Bridge opened for operation in 1916 and is a steel truss bridge that carries railroad traffic across the Mississippi River near Memphis, TN. The bridge has a total span of 1497.5m and comprises of five spans including an anchor arm (A), three cantilever arms (C), two spans (S), one fixed span (F), and one deck span (D) (Figure 2-2). The bridge carries two railroad tracks that each operate in both directions as well as a single pedestrian walkway along the north truss. The Harahan Bridge is part of Union Pacific’s Memphis



(a) (b)

Figure 2-3. (a) Labeling scheme for the Harahan Bridge north truss' US0-LS0 assembly's six parallel eyebars; (b) illustration of a change in boundary conditions causing an uneven distribution of the total load, F_T, across the six eyebars.

subjunction and serves as a critical part of the railroad's transportation infrastructure because it links railroad operations on the east and west sides of the Mississippi River. The bridge has a unique risk profile because its location exposes it to a number of natural and man-made hazards; at the same time, it serves as a critical network component within a regional supply chain network. The Harahan Bridge is the only point of connection between the railroad's operations on the east side and west side of the river thereby requiring the rerouting of rail traffic in the case of bridge closure (e.g., for maintenance or due to bridge failure). Consequently, bridge managers must be able to make quantitatively informed asset management decisions in the face of hazards. The bridge is potentially exposed to a number of external hazards including seismic activity (due to its location within the New Madrid Fault zone), barge collisions, and scour, in addition to load- and weather-induced "aging hazards" such as fatigue. External and aging hazards can alter the safety profile of the structure at a component or system level at any point in time between scheduled maintenance inspections.

2.2.2 Fracture-Critical Eyebar Assemblies

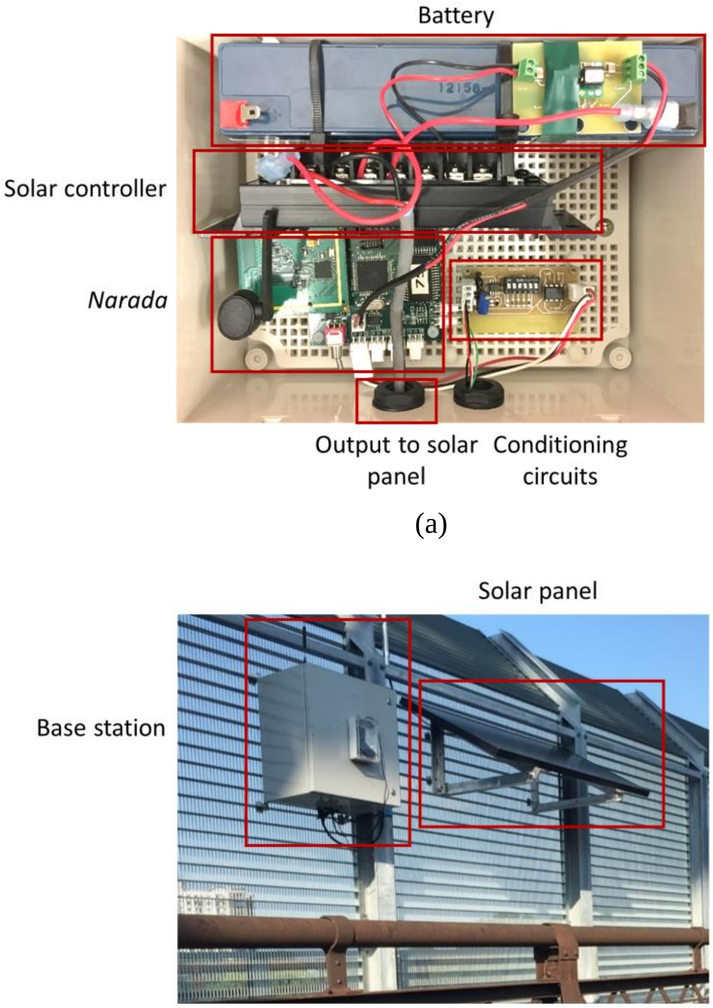
The Harahan Bridge's tensile truss members—including its tensile eyebar assemblies—undergo repeated high live load cycles due to long and heavy freight trains crossing the bridge daily. Consequently, the eyebars are especially susceptible to failure resulting from accelerated fatigue damage (Gostautas and Tamutus 2015, Reid 2010). The Harahan Bridge consists of 67 eyebar assemblies, each comprising of four or six parallel eyebar plates (Figure 2-3). These eyebar assemblies serve as upper chord, lower chord, diagonal truss, and vertical truss assemblies throughout the bridge. The configuration of the north truss' US0-LS0 eyebar assembly is detailed in Figure 2-3 and its location within the greater truss system is identified in Figure 2-2. The north and south US0-LS0 eyebar assemblies connect the easternmost cantilever arm (Span C in Figure 2-2) to the main span (Span S in Figure 2-2) and are fracture-critical structural assemblies that can lead to span failure. This is confirmed using the FE model presented in Section 2.2.3. Each US0-LS0 assembly has one upper pin connection and one lower pin connection. As illustrated in Figure 2-3, changes in the boundary conditions at the pin connections can cause some eyebars to carry a greater proportion of the total axial assembly load distributed across all six eyebars than expected (AREMA 2016). Here, $P_1, P_2, \ldots, P_6 \in [0, 1]$ is the proportion of the total assembly load, F_T, carried by each eyebar, subject to the constraint $\sum_{i=1}^{6} P_i = 1$. An increase in the proportion of the total assembly load carried by an eyebar is directly correlated to a decrease in remaining fatigue life (i.e., accelerated fatigue accumulation) (AASHTO 2012, Flanigan et al. 2017). Since this behavior is not explicitly accounted for in the design process, the alignment of parallel eyebar plates is a priority for the bridge owner. This is reflected in the bridge owner's existing inspection methods, which require inspectors to estimate eyebar tautness manually during periodic inspections. While the proposed framework is applied to a specific assembly type, the methodology can be naturally extended to any truss element of concern.

Standard protocol for eyebar inspection on any truss bridge—including the Harahan Bridge—dictates that inspectors estimate eyebar tautness manually during periodic inspections in order to identify imbalances leading to an uneven distribution of the assembly load across the parallel eyebar plates (Flanigan et al. 2017). Inspectors are trained to manually induce a lateral out-of-plane vibration in each eyebar and count the number of vibration cycles over a defined period. Eyebar plates vibrating with a higher frequency compared to the other parallel plates implies they are carrying a greater portion of the assembly's total axial tensile load. Eyebar plates are categorized into defect classes based on their measured frequencies. Any imbalances, when significant, can prompt a decision to rebalance the eyebar pins to redistribute the axial load equally across all of the eyebar plates.

However, *quantitatively* informed prognostics are difficult to achieve based on these current management strategies. Infrequent manual inspections allow for accelerated fatigue damage to occur before changes in the pin boundary conditions are detected and corrected. Additionally, manual inspection lacks a quantitative measure of how much fatigue has actually occurred. This means decisions made based on visual inspection remain decoupled from codified estimation of remaining fatigue life based on *S*-*N* curves or probabilistic fatigue assessment. Due to these drawbacks, the primary goal of the implemented sensing program is to quantitatively measure the coupled behavior between fatigue and relative tautness *continuously* using automatic sensing in a probabilistic analytical framework.

2.2.3 Harahan Bridge Data Collection Program and Instrumentation

The author deployed a long-term wireless structural monitoring system on the Harahan Bridge in 2016. Data collected from July 2016 through October 2017 is used in this chapter; after October 2017, the system was turned over to the bridge owner for use in their management processes. The primary purpose of the wireless structural monitoring system is to collect and transmit structural response information that is necessary to quantify bridge performance under routine train loads, thereby offering the bridge owner pertinent information on bridge behavior that can augment their visual inspection processes and aid in their decision making. The structural monitoring system is designed to accomplish three objectives. First, the monitoring system is designed to observe the global behavior of the bridge under routine loading (e.g., train traffic, wind loads). Specifically, vibrations of the bridge are sought to extract modal parameters for updating an FE model of the bridge. Second, the monitoring system is designed to observe the local behavior of critical bridge members. In consultation with the bridge owner, the US0-LS0 eyebar assemblies are identified as some of the bridge's most critical tensile truss members and are monitored for condition assessment by tracking fatigue accumulation under routine live loading. Given that fatigue damage will be primarily caused by repeated stress cycles induced by train loads, the third monitoring system objective is to capture the assembly behavior under all train loads. Towards this end, the monitoring system is designed to turn on and initiate data collection when measured vertical vibrations of the bridge's main span exceed a



(a)

(b)

Figure 2-4. (a) *Narada* wireless sensing node; (b) base station installed on the pedestrian walkway fence.

predefined threshold. In addition, the monitoring system is designed to extract train load parameters including train speed, direction (i.e., eastbound versus westbound), and a crude estimate of train and car bogie weights.

2.2.3.1 Data Collection and Wireless Communication Program

The wireless monitoring system installed on the Harahan Bridge uses the *Narada* wireless sensing technology developed at the University of Michigan (Swartz et al. 2005). This high-fidelity monitoring technology has been used to monitor a wide variety of bridges in permanent deployments over the past decade including on the New Carquinez Bridge (Kurata et al. 2013), Telegraph Road Bridge (O'Connor et al. 2017), and Newburg Road Bridge (Hou et al. 2019). The *Narada* wireless sensing node is used as the primary data collection unit to collect and wirelessly transmit time-synchronized data to a base station installed on the pedestrian walkway along the north truss of the Harahan Bridge. The *Narada's* 16-bit analog-to-digital converter supports data collection on four channels and its 8-bit embedded processor enables node-based data interrogation. Each node also has a 2.4 GHz IEEE802.15.4 transceiver used to transmit and receive data with the network base station. Each *Narada* node—pictured in Figure 2-4(a)—contains a solar controller connected to a 10W solar panel and a 12V rechargeable battery for solar harvesting and energy storage, respectively.

The base station—pictured in Figure 2-4(b)—contains a PC-104 single-board computer, a CC2420 RB transceiver that communicates with *Narada* nodes over the 2.4 GHz frequency band using the IEEE 802.15.4 standard, and an LTE cellular modem that offers the base station Internet access. The purpose of the base station is to initiate time-synchronized data collection among the network's *Narada* units installed on the bridge. The base station uses the cellular modem to transmit collected data via the IEEE802.15.4 transceivers to an off-site SQL server for storage and processing. Since the base station and sensing nodes operate using harvested solar energy, it is necessary to minimize power consumption of the monitoring system. Rather than collecting data on a schedule—which does not ensure that all train loading events are captured—a trigger system was designed to initiate data collection only during active loading events (such as when trains are

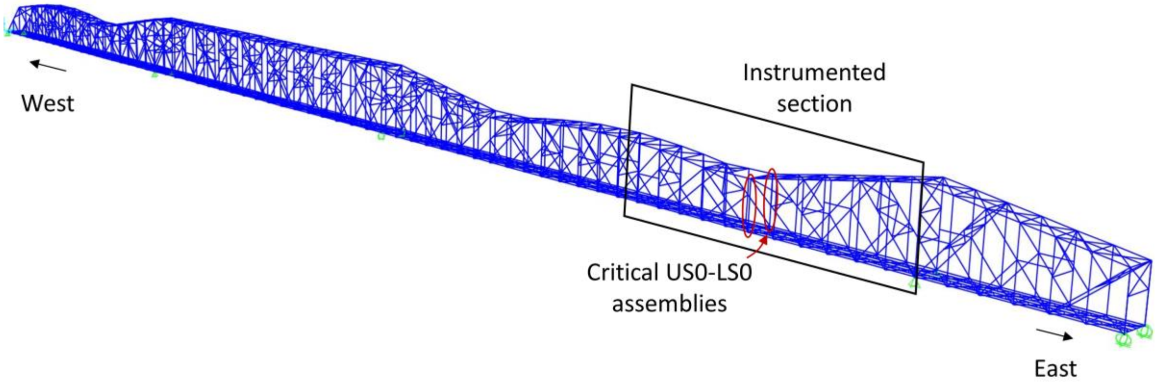


Figure 2-5. FE model of the Harahan Bridge (modeled in CSiBridge).

Table 2-1. Summary of the sensing transducers used in the Harahan Bridge's monitoring system.

Equipment type	Model	Type of measurement	Primary function within framework	Number deployed	Location
Weldable strain gage	Hitec HBWF-35-125-6-10GP-TR	Local strain	Measure stress cycles[1]; monitor axle loads	8	US0-LS0, UC5-LC5
Uniaxial accelerometer	Silicon Design 2012-002	Local acceleration	Measure total load carried by each eyebar (relative tautness)[1]	12	US0-LS0
Triaxial accelerometer	Silicon Design 2422-005	Global acceleration	FE model development[2]	4	US0-LS0, US2-LS2
Geophone	GeoSpace Geo-11D 4.5-380 VT	Global velocity	Initiate data collection[1]	1	UC5-LC5 Pedestrian walkway fence

[1] Used in proposed probabilistic fatigue analysis

[2] Not used in proposed probabilistic fatigue analysis

Table 2-2. Comparison of the modal frequencies extracted from the finite element (FE) model and measured data.

Mode #	Mode type	Simulated modal freq. (Hz)	Measured modal freq. (Hz)	Percent error (%)
1	Transverse[1]	0.679	0.683	0.59
2	Transverse	0.878	0.917	4.3
3	Transverse	1.08	–	–

[1] Modal assurance criterion (MAC) value = 0.83

crossing the bridge, seismic events occur, or a barge collides with a bridge pier) with the nodes remaining in a low-power state otherwise. This enables the wireless sensing system to collect and transmit data in 30-second polling cycles for every train that passes. The trigger system operates using a geophone installed within the base station to measure vertical velocity continuously. As trains approach the bridge, their ground vibrations cause the bridge to vibrate. When the geophone measurement exceeds a predefined threshold velocity of 3.15cm/s, the base station is triggered to send a "wake-up" command to the *Narada* wireless units and initiates the data collection process. The threshold is set to be low enough such that the structural response due to every train passing over the bridge is recorded, regardless of the configuration and size of the train.

2.2.3.2 Sensing System Design and Instrumentation

The installed wireless structural monitoring system measures a number of local and global structural properties. Figure 2-2 presents the system instrumentation plan which includes four triaxial accelerometers (sampled at 200Hz), twelve uniaxial accelerometers (sampled at 200Hz), eight strain gages (sampled at 200Hz with x500 signal gain), one geophone, and one base station (Table 2-1). Flanigan et al. (Flanigan et al. 2017) presents the development of an FE model of the Harahan Bridge using CSiBridge (CSI 2016). The FE model discussed herein is *not* used to carry out the probabilistic fatigue analysis. Rather, it is used in this application to validate the sensing instrumentation plan and confirm that the selected eyebars are in fact fracture critical and a primary concern for the bridge owner. The FE model—shown in Figure 2-5—was constructed based on detailed engineering design drawings provided by the bridge owner. Experimental data collected from the triaxial accelerometers is used to calculate the modal properties (i.e., modal frequencies, modal damping, and mode shapes) of the in-service structure and confirm that the FE model's boundary conditions are modeled properly by comparing experimental modes to those analytically estimated by the FE model. The FE model's boundary conditions are modeled based on the original engineering drawing descriptions. Table 2-2 indicates that the FE model's modal characteristics are well aligned with those empirically extracted from bridge response data. Strain gages are installed on the UC5-LC5 vertical hanger assemblies on the north and south truss systems to gain insight into the train car and locomotive axle loads (Figure 2-2). These vertical hangers are suspended by upper and lower longitudinal chords and theoretically carry the entire axle loads of each passing train. Since the UC5-LC5 members are adjacent to the US0-LS0 assemblies, time-synchronized vertical strain measurements in both members combined with information on standard train car geometries allows each train's direction and speed to be accurately estimated. While beyond the scope of this chapter, these measurements are also capable of estimating the vertical load of each passing car bogie.

The primary purpose of the instrumented sensing program is to measure the coupled behavior between relative tautness and fatigue of the north truss' US0-LS0 eyebar assembly. This requires: 1) measurement of all stress cycles; 2) measurement of the proportion of the total assembly load carried by each of the six eyebars. Rather than monitor fatigue and relative tautness as two separate limit states using strain, an alternative approach of using acceleration and strain sensing modalities in parallel is used to inform a single limit state function (as will be fully discussed in Section 2.3). First, since stress cycles govern fatigue, strain gages are installed on the lower box section of the US0-LS0 assembly directly below the lower eyebar pin to measure the US0-LS0 assembly's axial strain cycles (Figure 2-6). Second, uniaxial accelerometers are installed on each of the US0-LS0 assembly's six eyebar plates near mid-height (Figure 2-6). The measured lateral

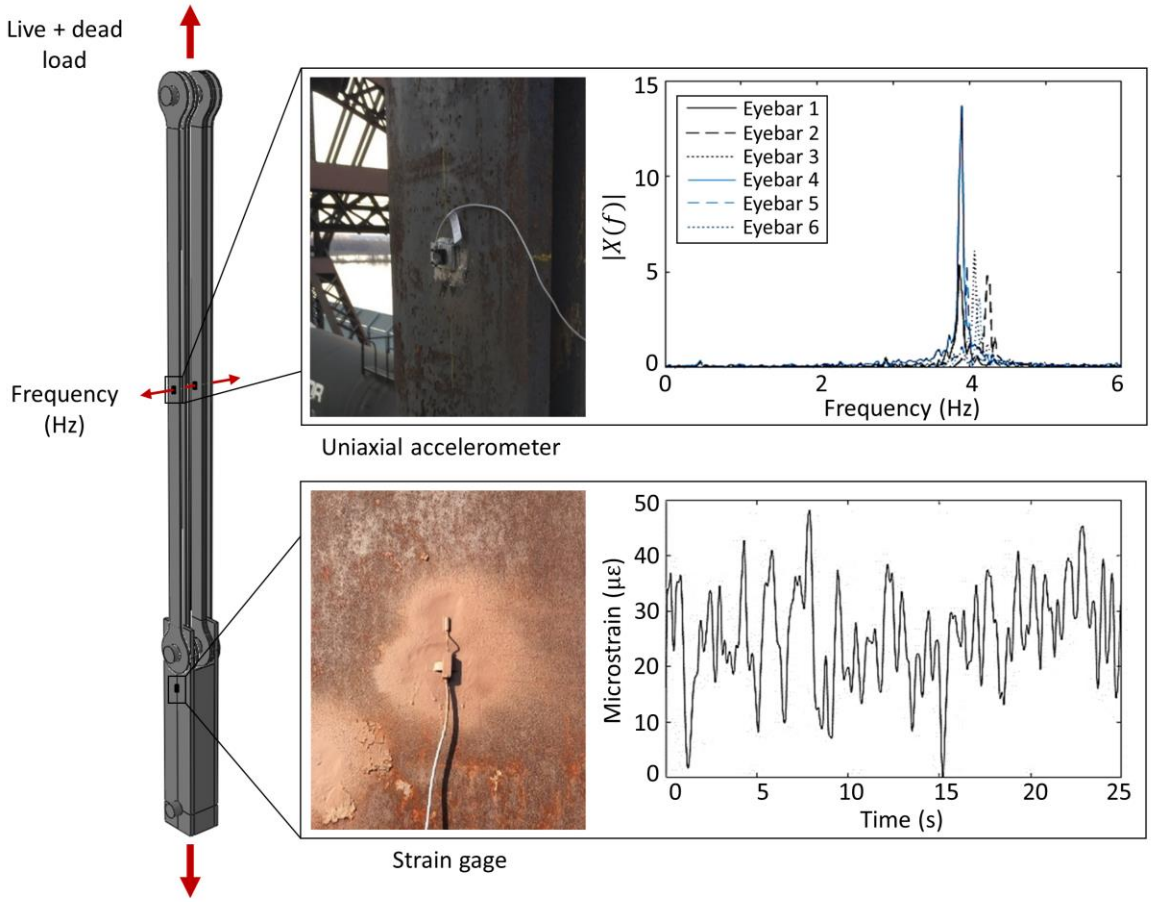


Figure 2-6. Uniaxial accelerometers and strain gages installed on the US0-LS0 eyebars. Also shown are sample frequency spectra (derived from lateral acceleration data, $\ddot{x}(t)$) and sample axial strain response data collected at these locations over a 25-second period.

acceleration of each vibrating eyebar is used to calculate each eyebar's loaded natural frequency. Given the known geometric and material properties of each eyebar, the natural frequencies are used in a physics-based model to calculate the total load in each eyebar. Consequently, the uniaxial acceleration data explicitly informs the proportion of the total US0-LS0 assembly load carried by each of the parallel eyebar elements. This approach is commonly adopted to estimate the tautness of in-service axially loaded steel cables and hangers (Fang and Wang 2012, Nugroho et al. 2014, Priyosulistyo and Ferdina 2014). Acceleration-based relative tautness assessment of each eyebar is used instead of strain measurements because measured frequencies correlate to the *total* combined axial dead and live load, whereas strain measurements only capture the live load response. Further, acceleration-based measurements are consistent with visual inspection practices in which inspectors visually count vibration cycles of manually excited eyebar plates to estimate axial tautness. This makes the adoption of an automated and quantitative asset management framework more feasible.

2.3 Long-Term Reliability-Based Condition Monitoring

The methodology proposed herein uses long-term monitoring data to quantify the remaining fatigue life of monitored railroad bridge components in a probabilistic analytical framework that continuously and automatically processes data. Unlike existing literature, this framework is purely data-driven and is carried out using 16 months of continuous, long-term structural response data collected for each passing train on the Harahan Bridge. Additionally, this methodology implements a novel holistic approach to tracking fatigue accumulation in which the sensing design empowers full consideration of the dead and live loads and associated changes in boundary conditions. This is critical for monitoring assets with secondary stresses due to high static mean stress levels or unexpected behavior at the boundary conditions that are not considered in idealized mechanical systems. The proposed analytical framework accounts for the coupled behavior between fatigue and relative tautness of a set of critical parallel eyebars subject to both dead and live loads.

2.3.1 Fatigue Reliability Method

The goal of reliability methods is to calculate a component or system's reliability index, β^r, which is an objective measure of the level of safety corresponding to a specific limit state function (Nikolaidis et al. 2004). Given an n-dimensional vector of random variables, $\boldsymbol{X} = [X_1, X_2, \dots, X_n]^T$, a limit state function, $G(\boldsymbol{X})$, is defined as $G(\boldsymbol{X}) = R - S$, where R denotes the resistance and S denotes the load effects. In general, the probability of failure, $P_f = P(\boldsymbol{X} \in \Omega)$, is defined as

$$P_f = \int_{\Omega} f_{\boldsymbol{X}}(\boldsymbol{X}) d\boldsymbol{X} \tag{2.1}$$

where $f_{\boldsymbol{X}}(\boldsymbol{X})$ is the joint probability density function (PDF) of $\boldsymbol{X}$ and the failure domain, Ω, is defined as $\Omega \equiv G(\boldsymbol{X}) \leq 0$ for the single component reliability problem (Nikolaidis et al. 2004). For the series system reliability problem in which the failure domain is described by the union of c component failure events, the failure domain is defined as $\Omega \equiv \bigcup_{i=1}^{c} G_i(\boldsymbol{X}) \leq 0$. Similarly, for the parallel system problem (i.e., one with redundancy) in which the failure domain is described by the intersection of c component failure events, the failure domain is defined as $\Omega \equiv \bigcap_{i=1}^{c} G_i(\boldsymbol{X}) \leq 0$ (Nikolaidis et al. 2004). The level of safety associated with a limit state function is characterized by the reliability index, β^r. When the limit state function is linear in the independent standard normal space (ISNS), the reliability index is defined as the minimum distance between the origin and the limit state function's failure domain in the independent standard normal space. When the limit state function is not linear in the ISNS, first-order approximations can be used to approximate the reliability index, β^r, in which case the reliability index is defined as the approximate, rather than absolute, minimum distance. The reliability index, β^r, is a powerful scalar metric because the probability of failure is fully described by the reliability index as

$$P_f \cong \Phi(-\beta^r) \tag{2.2}$$

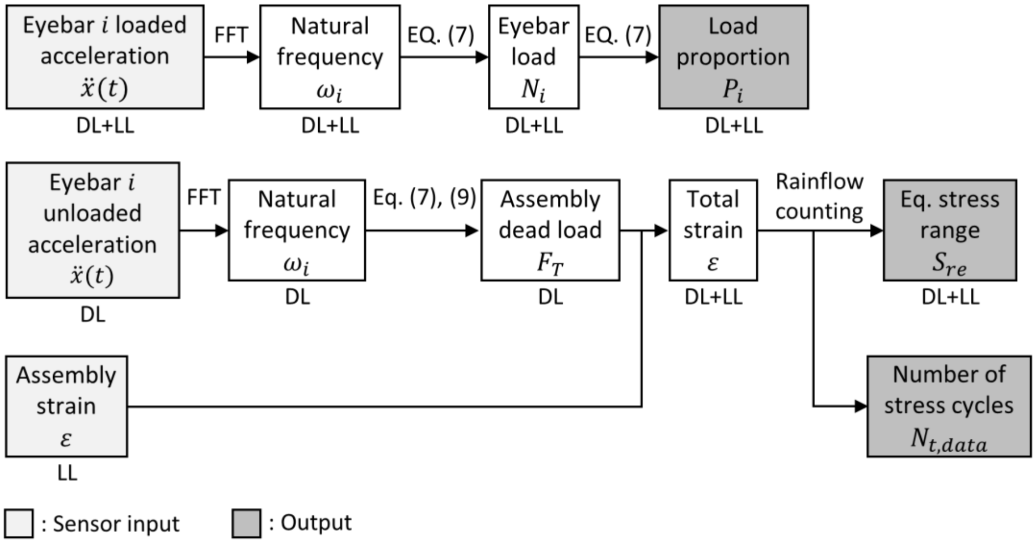


Figure 2-7. Outline of the mapping between sensor inputs (i.e., strain and uniaxial acceleration) and measured deterministic and random variable outputs. Here, DL denotes dead load and LL denotes live load.

where $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF). The equality in Equation (2.2) holds if the limit state function is linear in the ISNS.

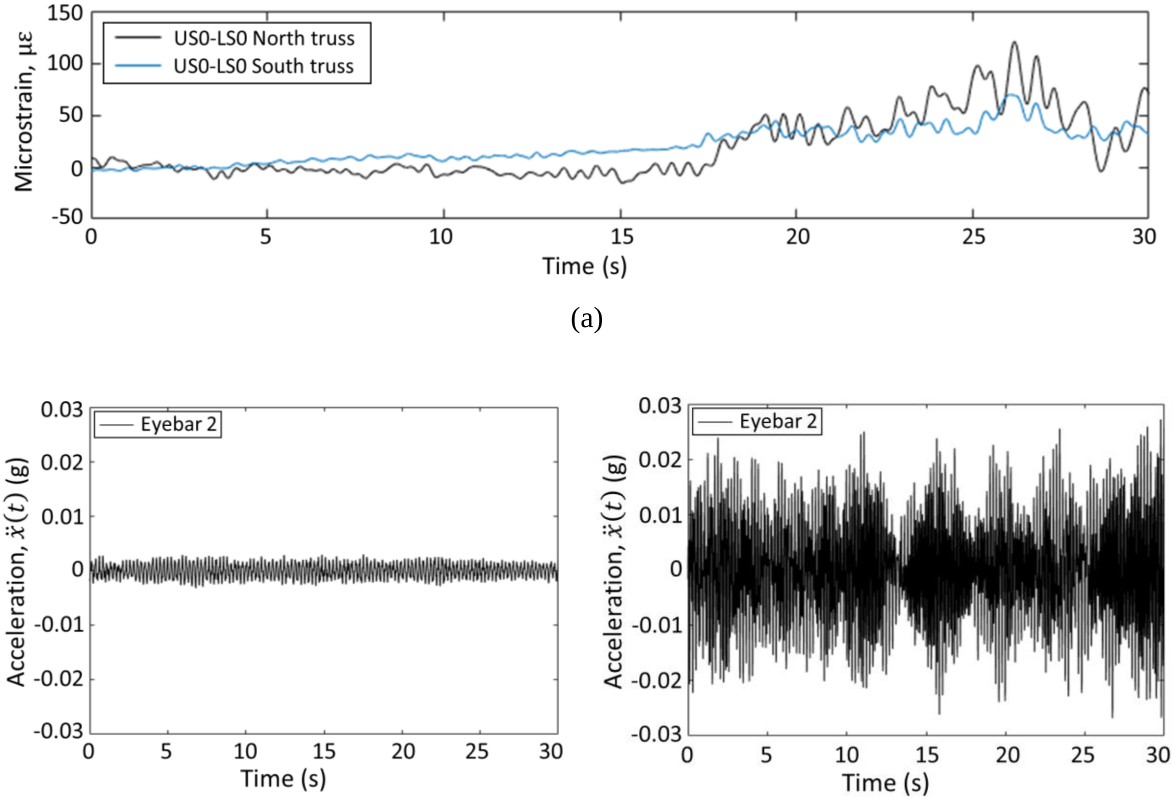
Fatigue-induced failure can be framed as a reliability problem where uncertainty is modeled in the resistance and load effects (Rakoczy et al. 2016, Kim and Mha 2001, Li et al. 2016, Tobias and Foutch 1997). The goal of the proposed method is to formulate a single limit state function that assesses the probability of failure considering *both* the operational condition of the eyebar (i.e., relative tautness) and fatigue accumulation. First, Section 2.3.2 formulates the component reliability problem where the eyebars are considered independent of each other and the objective is to calculate the reliability index for each eyebar individually. However, in reality, the loads carried by each of the eyebars are correlated and all six of the eyebar plates in the US0-LS0 assembly must fail for the entire parallel system to fail. Consequently, Section 2.3.3 formulates the parallel system reliability problem in which the objective is to calculate the reliability index for the entire US0-LS0 assembly system. This section also outlines the data-driven method to calculate the distribution parameters of random variables derived from monitoring data that are necessary inputs to the two reliability problems.

2.3.2 Component Reliability

The failure domain for fatigue is characterized by $D \geq \Delta$, where Δ is Miner's critical fatigue damage index and D is the fatigue damage accumulation index (Miner 1945). According to Miner's Rule, failure due to fatigue occurs when $D \geq 1$. However, Sobczyk and Spencer (Sobczyk and Spencer 1992) found that the fatigue damage accumulation index, D, at failure can range from 0.5 to 2.0 and Wirsching (Wirsching 1984) extended the definition of failure due to fatigue to $D \geq \Delta$, where Miner's critical fatigue damage index, Δ, is modeled as a lognormal random variable with a mean of 1.0 and a coefficient of variation (COV) of 0.3. Based on this failure domain, the limit state function governing each eyebar of the US0-LS0 assembly is

$$G_i(\boldsymbol{X}) = R - S = \Delta - D = \Delta - \frac{N_t \cdot (S_{re} \cdot P_i)^m}{A}, i = 1,2,\dots,6 \tag{2.3}$$

where $i = 1,2,\dots,6$ designates the limit state function corresponding to each of the six parallel eyebar plates as labeled in Figure 2-3(a). Here, S_{re} is the equivalent stress range, P_i is the proportion of the total assembly load carried by each eyebar, N_t is the total number of stress cycles, A is the fatigue-strength coefficient defined by the American Railway Engineering and Maintenance-of-Way Association (AREMA) (AREMA 2016), and m is a material constant assigned by AREMA (AREMA 2016). The equivalent stress range, S_{re}, load proportion carried by each eyebar, P_i, and total number of stress cycles, N_t, are



(a)

(b)

(c)

Figure 2-8. (a) Time series strain response of the north and south US0-LS0 eyebar assemblies as a train enters the Harahan Bridge on the north track and loads the bridge; (b) transverse eyebar acceleration when the bridge is unloaded; (c) transverse eyebar acceleration when the bridge is loaded and a train is crossing.

calculated using monitoring data, whereas the fatigue-strength coefficient, A, and material constant, m, are adopted from *S-N* curves within AREMA design specifications and are introduced here as random and deterministic variables, respectively (AREMA 2016). The remainder of Section 2.3.2 derives a closed-form solution for the reliability index and discusses how to characterize the statistical parameters of the deterministic and random variables based on measurement data (Sections 2.3.2.1 through 2.3.2.3) and design specifications (Section 2.3.2.4). Figure 2-7 illustrates the methodology used to derive the data-driven deterministic and random variables in Sections 2.3.2.1 through 2.3.2.3. The resulting distribution parameters are presented in Table 2-3.

Table 2-3. Deterministic and random variable parameters.

Parameter	Distribution	Value	Mean	Standard deviation	Source
Miner's critical damage accumulation index, Δ	Lognormal	–	1.0	0.3	(Wirsching 1984)
Fatigue detail coefficient, A MPa^3	Lognormal	–	3.61(10^{12})	1.62(10^{12})	(AREMA 2016)
Equivalent stress range, S_{re} MPa	Lognormal	–	15.3	31.2	Monitoring data
Material constant, m	Deterministic	3.0	–	–	(AREMA 2016)
Total number of stress cycles[1], $N_{t,data}$	Deterministic	1,725,727	–	–	Monitoring data
Load proportion in eyebar 1, P_1	Lognormal	–	0.171	0.0152	Monitoring data
Load proportion in eyebar 2, P_2	Lognormal	–	0.180	0.0060	Monitoring data
Load proportion in eyebar 3, P_3	Lognormal	–	0.170	0.0051	Monitoring data
Load proportion in eyebar 4, P_4	Lognormal	–	0.150	0.0037	Monitoring data
Load proportion in eyebar 5, P_5	Lognormal	–	0.160	0.0042	Monitoring data
Load proportion in eyebar 6, P_6	Lognormal	–	0.172	0.0054	Monitoring data

[1] Number of stress cycles collected over 16-month monitoring period

There are two notable observations that help to simplify the reliability index calculation with respect to the limit state function defined in Equation (2.3). First, the random variables outlined in Table 2-3 are all lognormally distributed. Second, two limit state functions are equivalent if they share the same failure surface, namely, $G(\boldsymbol{X}) = 0$ for the purposes of this work (Barker and Puckett 2013). Consequently, the limit state function governing fatigue presented in Equation (2.3) is reformulated using an equivalent representation in order to drastically simplify the reliability analysis and produce a closed-form solution for the reliability index, β^r. Letting the resistance and load effects derived in Equation (2.3) be governed by $G(\boldsymbol{X}) = \ln\left(\frac{R}{S}\right) = \ln(R) - \ln(S)$, the original limit state function can be equivalently represented as

$$G_i(\boldsymbol{X}) = \ln(\Delta \cdot A) - \ln[N_t \cdot (S_{re} \cdot P_i)^m] = \ln(\Delta) + \ln(A) - [\ln(N_t) + m \cdot \ln(S_{re}) + m \cdot \ln(P_i)] \tag{2.4}$$

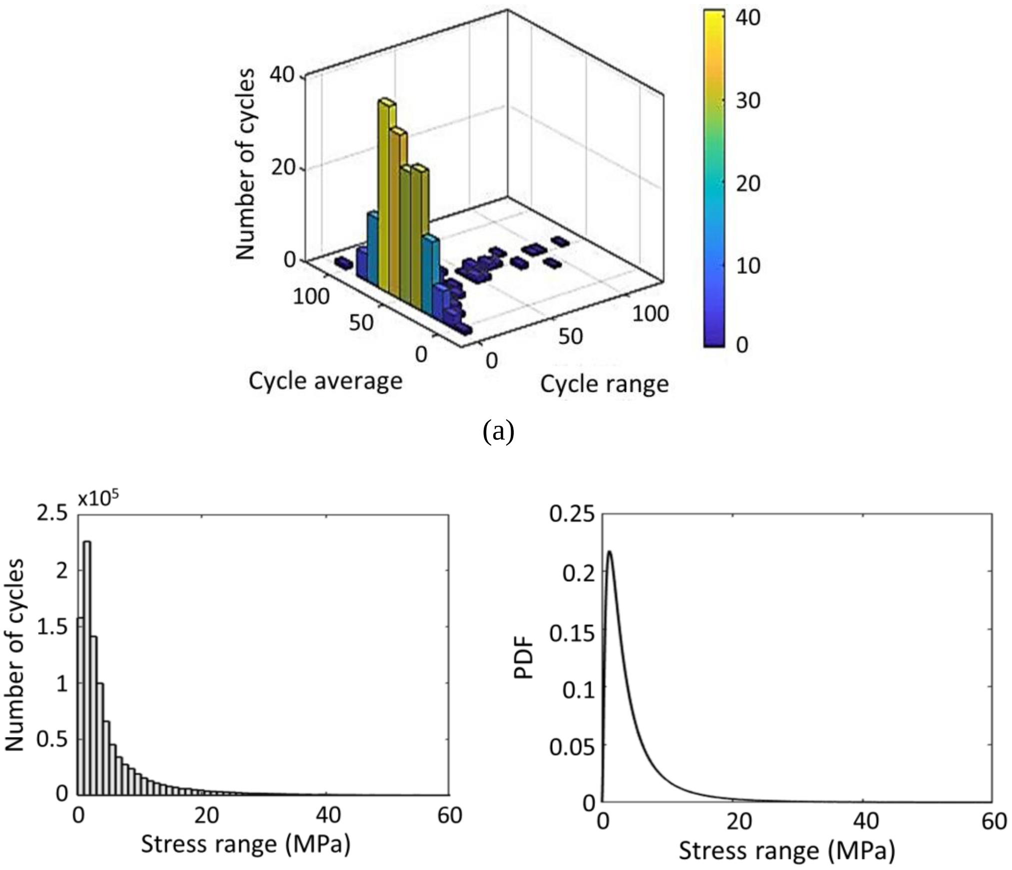
An exact closed-form solution of the reliability index, β_i^r, is derived directly from the mean, λ_{G_i}, and standard deviation, ζ_{G_i}, of the reformulated limit state function in Equation (2.4) as

$$\beta_i^r = \frac{\lambda_{G_i}}{\zeta_{G_i}} = \frac{\lambda_\Delta + \lambda_A - [\ln(N_t) + m \cdot \lambda_{S_{re}} + m \cdot \lambda_{P_i}]}{\left[(\zeta_\Delta)^2 + (\zeta_A)^2 + (m \cdot \zeta_{S_{re}})^2 + (m \cdot \zeta_{P_i})^2\right]^{\frac{1}{2}}} \tag{2.5}$$

This simplification eliminates any need for approximate reliability methods such as first-order reliability methods (FORM), second-order reliability methods (SORM), or Monte Carlo methods.

2.3.2.1 Equivalent Stress Range

The random variable S_{re} is the equivalent stress range assuming the entire assembly load, F_T, is applied to a single eyebar. Here, F_T, is the total axial load on the box section of the US0-LS0 assembly below the lower pin (see Figure 2-3). The equivalent stress range, S_{re}, is derived from the measured strain response at the box section and is modeled as a lognormal random variable to account for variable-amplitude stresses. A 30-second sample of strain time series data collected at the US0-LS0 north and south eyebar box sections is shown in Figure 2-8. The live load strain data in Figure 2-8 reflects the assembly's structural response to a train—beginning with the heavy leading locomotives—entering the Harahan Bridge on the north track and subsequently passing the US0-LS0 strain gage location at approximately 25 seconds. The train's leading locomotives enter Span C at around 17 seconds, causing an increase in the box section's live load strain response. The US0-LS0 box sections typically experience the highest amplitude strain cycles when the locomotives pass directly adjacent to the monitored assemblies; as



(a)

(b)

(c)

Figure 2-9. (a) Rainflow counting results for a 30-second sample monitoring period; (b) equivalent stress-range histogram for the 16-month monitoring period (July 2016 through October 2017); (c) the corresponding lognormal PDF characterizing the equivalent stress-range.

expected, the strain cycles for the US0-LS0 north truss assembly undergo higher amplitude cycles since the train is crossing along the northernmost track. The box section experiences a short period of slight *relative* compression during the first 15 seconds, meaning the assemblies are still in tension due to the dead load but the compressive live load is decreasing the overall tension because the train is entering the bridge on Span A. When the train enters Span A, the adjacent cantilever span (Span C) comprising of the US0-LS0 assemblies experiences slightly reduced tensile loads. Once the leading locomotives enter Span C at around 17 seconds, the box section remains in relative tension for the remainder of the train crossing. A rainflow counting algorithm is implemented based on the ASTM E 1049 Standard (ASTM 2017) in order to calculate the amplitudes of fully reversible and half cycles from the strain data each time a train crosses the bridge (Figure 2-9(a)). The number of applied stress cycles and equivalent stress range are the primary mechanisms contributing to fatigue damage. Although a secondary factor, mean stress is also correlated to fatigue damage (Lee et al. 2005) and is considered in this study because steel railroad bridges experience a high live load to dead load ratio as trains often fully load the bridge during crossings, resulting in a high mean live load component (Unsworth 2010). This is evident in Figure 2-8(a) in which the mean strain response in the US0-LS0 assemblies increases as the train continues to load the bridge. By 25 seconds, the stress cycles induced by train axle loads appear as harmonic signals superimposed on the mean. The Goodman Method (Lee et al. 2005) is used as a total mean stress correction technique within the rainflow counting algorithm to compensate for the effects of normal tensile mean stress on high-cycle fatigue strength. The Goodman Method defines the relationship

$$\frac{S_a}{S_e} + \frac{S_m}{S_u} = 1 \tag{2.6}$$

where S_u is the ultimate tensile strength and the equivalent completely reversed stress amplitude, S_e, is the stress that is expected to result in the same fatigue life as the stress amplitude S_a at mean stress S_m. Figure 2-9(b) shows the resulting stress-range bin histogram for strain data collected over the 16-month monitoring period as well as the associated lognormal PDF characterizing the random variable parameters (Figure 2-9(c)).

2.3.2.2 Relative Tautness Across the Parallel Eyebar System

Recall from Figure 2-3 that F_T is the total load carried by the US0-LS0 assembly and that $P_i \in [0, 1]$ is the proportion of the total assembly load carried by each eyebar, subject to the constraint $\sum_{i=1}^{6} P_i = 1$. In other words, the axial load carried by the i^{th} eyebar is $P_i \cdot F_T$. As is the case of the equivalent stress, S_{re}, the proportion of the total assembly load in each eyebar, P_i, is a lognormal random variable with distribution parameters derived from monitoring data; the proportion of the total assembly load carried by an eyebar will never be below zero. While the six parallel eyebar plates are designed to carry equal proportions of the total assembly load, changes in the boundary conditions at the pins can lead to unequal load proportions across the eyebars. As discussed in Section 2.2.3, uniaxial accelerometers are installed near the mid-span of each eyebar plate in order to measure the out-of-plane (lateral) acceleration. These vibration measurements are used to inform a frequency-based analysis to calculate the relative tautness of each of the six eyebars during loading events.

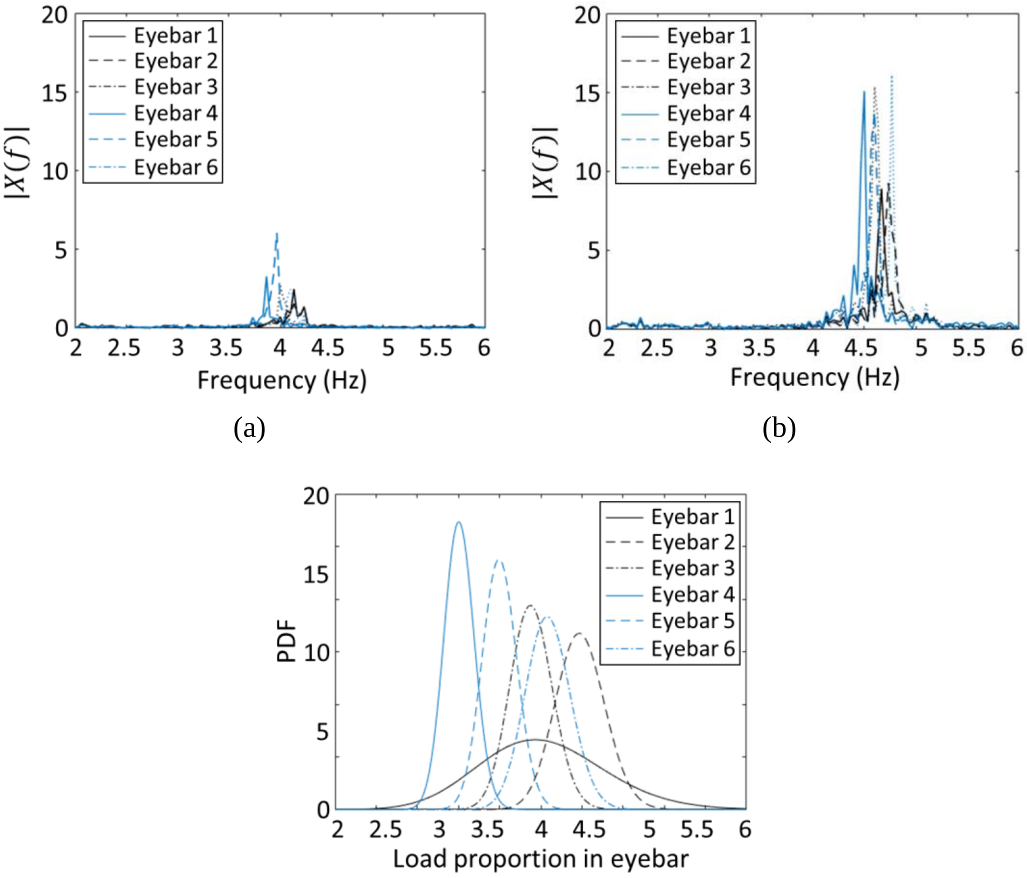
The following methodology outlines how monitoring data is used in a physics-based framework to calculate the proportion of the assembly load carried by each eyebar. First, the total load carried by each eyebar, N_i, is calculated as (Kollár 2003)

$$N_i = N_E \cdot \left(\frac{\omega_i{}^2}{\omega_0{}^2} - 1 \right) \tag{2.7}$$

A positive value of N_i denotes tension and a negative value denotes compression. Here, ω_i is the transverse natural frequency of each loaded eyebar calculated from acceleration measurements. The natural frequency is derived by transforming acceleration time series data for each 30-second polling cycle—beginning when the train enters Span C and ending when it exits Span C—into the frequency domain using the Fast Fourier Transform (FFT) and identifying the first dominant frequency. It is analytically confirmed that the first natural frequency corresponds to transverse vibrations. This is done by representing the mechanics of each eyebar as a continuous plate fixed at both ends with an applied axial tensile load (as estimated in the original engineering design drawings) and calculating the corresponding natural frequencies in each direction. To illustrate the influence of increased loading on the measured frequencies, Figure 2-10(a) and Figure 2-10(b) present sample frequency response spectra for the eyebars when the bridge is unloaded (i.e., no train event) and loaded, respectively. As expected, there is a clear increase in the transverse natural frequency when a train loads the bridge. Here, ω_0 is the natural frequency of the unloaded eyebar. The natural frequency, ω_0, is calculated based on the geometry of each eyebar under the assumption that the eyebars have fixed-fixed boundary conditions in the out-of-plane direction. This assumption is justified based on direct observation of the as-built structure and structural design drawings, which reveal that plates inserted between the eyebars near the lower and upper box sections hold the eyebar ends in place with rotational and out-of-plane lateral fixity (Figure 2-1(a)). Since all six of the eyebars have identical geometric and material properties, the unloaded natural frequency for each eyebar is calculated as (Blevins 2016)

$$\omega_0 = k^2 \cdot \left(\frac{E \cdot I}{A_g \cdot \rho \cdot L^4}\right)^{\frac{1}{2}} \tag{2.8}$$

where k is a constant based on the boundary conditions ($k = 4.730$ for fixed-fixed supports), $L = 12.8\text{m}$ is the pin-to-pin length of the eyebars, $A_g = 0.0181\text{m}^2$ is the gross-section area of an eyebar (i.e., the uniform width of the eyebar between the upper and lower



(a)

(b)

(c)

Figure 2-10. Sample frequency response (derived from acceleration data, $\ddot{x}(t)$ during one polling cycle) when the bridge is (a) unloaded and (b) loaded; (c) the PDF characterizing the load proportion carried by the six eyebars over the monitoring period.

pins), and $I = 2.97(10^{-6})\text{m}^4$ is the moment of inertia. The Harahan Bridge's truss system comprises of both carbon steel and alloy steel; the eyebars are constructed from alloy steel that was made especially for the project. The alloy steel eyebars were shipped and stored separately and testing occurred both on-site and during the pouring of each melt to ensure consistent mineral, material, and strength properties. The assumed material and geometric constants reported by the engineers during construction are taken to be deterministic. This is supported by the fact that the eyebars are made of alloy steel which reduces the threat of corrosion-induced section loss. Additionally, visual inspection confirms the absence of corrosion. That being said, one should be thoughtful of the influence of corrosion on the geometric properties of the structural members in applications when this assumption cannot be made. Hence, in Equation (2.8), the density used is $\rho = 0.284\text{kg/m}^3$ and the modulus of elasticity used is $E = 200\text{GPa}$.

Since the total load carried by the US0-LS0 assembly is $F_T = \sum_{i=1}^{6} N_i$, the proportion of the total assembly load carried by each eyebar is

$$P_i = \frac{N_i}{F_T} = \frac{N_i}{\sum_{i=1}^{6} N_i} \tag{2.9}$$

It should be noted that N_i is the total load carried by each eyebar including dead and live loads. Additionally, since the Euler buckling load, N_E, in Equation (2.7) cancels out in Equation (2.9), it will not be defined in this chapter. The process for calculating P_i is repeated for each polling cycle over the entire duration of each train loading. The resulting lognormal PDFs describing the load proportion, P_i, for each eyebar over the 16-month data collection period are presented in Figure 2-10(c).

2.3.2.3 Total Stress Cycles

Recall from Section 2.2.3 that a trigger system is configured to wake the sensing system up from a low-power sleep state and initiate data collection each time a train approaches. Consequently, the total number of stress cycles measured during the 16-month monitoring period, $N_{t,data}$, is considered to be deterministic over the monitoring period and is generated directly from the rainflow counting algorithm presented in Section 2.3.2.1. The deterministic value, $N_{t,data}$, is a subset of N_t, which is the total number of stress cycles that occurs over the asset's entire loading history (beginning when it opened for operation in July 1916).

2.3.2.4 Variables Derived from Design Specifications

The AREMA (AREMA 2016) and American Association of State Highway and Transportation Officials (AASHTO) (AASHTO 2012) codes outline provisions for railroad and steel highway bridge fatigue design, respectively. Both specifications use the same set of experimental data to derive *S*-*N* curves for steel, which correlates the number of cycles to failure for a given stress-range magnitude (Zhou 2006, Keating and Fisher 1986). Within these provisions, eyebars are categorized as Category E elements, meaning the corresponding value of m in Equation (2.3) is deterministic and equals 3.0. A_d is the design value of A and equals 3.61(10^{11})MPa^3. Since there is a 95% confidence level associated with A_d, the mean value, μ_A, standard deviation, σ_A, and COV, δ_A, of the lognormal random variable parameters characterizing A are derived from $A_d = \mu_A - 2 \cdot \sigma_A = \mu_A \cdot (1 - 2 \cdot \delta_A)$, where the COV is assumed to equal 0.45 (Wirsching et al. 1987, Zhao et al. 1994).

2.3.3 System Reliability of the Eyebar Assembly

Recall from Section 2.3.2 that the failure domain under consideration is governed by $D \geq \Delta$ and the corresponding limit state function for each eyebar is formulated in Equation (2.3). Equation (2.5) derives the closed-form solution for the reliability index, β_i^r, for each eyebar component separately. In reality, the US0-LS0 assembly's parallel plate configuration leads to a parallel system reliability problem with a number of factors complicating the analysis. First, all six of the eyebars have to fail for the entire US0-LS0 parallel system to fail (Constraint 1). Second, after an eyebar fails, the load carried by the failed eyebar is transferred to the remaining eyebars (Constraint 2). Third, the loads carried by each of the eyebars are correlated and the sum of the proportion of the load carried by all intact eyebars must always equal one (Constraint 3). That is, even as eyebars fail and loads are redistributed, the sum of the load proportion across all remaining eyebars must equal one.

Failure due to fatigue typically undergoes three stages: 1) cracks initiate around stress concentrations; 2) cracks propagate incrementally; 3) cracks reach a critical size and propagate rapidly (TRB 2014). The stress levels leading to fatigue-induced failure are typically significantly lower than those stresses causing failure under static loading conditions (TRB 2014). The limit state function considered in this chapter (Equation (2.3)) is based on remaining fatigue life which is the sum of the cycles in the first and second stages (i.e., crack initiation and incremental propagation). While failure with respect to the limit state function occurs after the first two stages of fatigue-induced failure, there may be reserve capacity in the cracked eyebar(s) leading to ductile, rather than brittle, system failure (TRB 2014). The redundant parallel eyebar plates comprising the US0-LS0 assembly allow for the excess stresses in a cracked plate to be redistributed (TRB 2014). This means that there is likely a period of time before the eyebar(s) with fatigue cracks undergo the third stage of fatigue-induced failure. When this time period is measured in at least weeks or months, the time lag would permit the owners to identify the local damage and perform maintenance. To remain conservative though, the following analysis assumes that the failure is brittle and that when an eyebar violates the limit state function (i.e., $G(\boldsymbol{X}) \leq 0$), the eyebar is no longer in service and the loads are redistributed to the remaining eyebars immediately.

For a parallel system, safety estimation under multiple limit states and failure modes must be considered. The failure domain is defined as $\Omega \equiv \bigcap_{i=1}^{c} G_i(\boldsymbol{X}) \leq 0$ where $c = 6$ for the case of the US0-LS0 member's six parallel eyebar plates. The reliability index calculation for the proposed parallel system problem is reformulated as follows. For the limit state function presented in Equation (2.3)

$$\beta_{sys}^{r} = -\Phi^{-1}\left(\int \dots \int I_p\,(\boldsymbol{X}) f_{\boldsymbol{X}}(\boldsymbol{X}) d\boldsymbol{X}\right) \text{ where } I_P = \begin{cases} 1 \text{ if} \bigcap_{i=1}^{6} G_i(\boldsymbol{X}) \leq 0 \\ 0 \text{ if} \bigcap_{i=1}^{6} G_i(\boldsymbol{X}) > 0 \end{cases} \tag{2.10}$$

The system described by Equation (2.10) is subject to Constraint 3, meaning that

$$\sum_{i=1}^{j} P_i = 1 \tag{2.11}$$

where $j \in [1,2,\dots,6]$ is the number of remaining eyebars that have not failed. That is, the sum of the proportion of the loads carried by all eyebars that have not yet failed must always be equal to one.

Because there is no closed-form solution for Equation (2.10) subject to Constraint 3, Monte Carlo methods are used to approximate the reliability index. The estimator, J, of this system-level failure probability is defined as

$$P_f \approx J = \frac{1}{N} \cdot \sum_{w=1}^{N} I_P \left(\bigcap_{i=1}^{6} G_i(\widehat{\boldsymbol{X}}_w)\right) \tag{2.12}$$

where N is the total number of trials and samples $w = 1,2,\dots,N$ are drawn from $\widehat{\boldsymbol{X}}_w = [\widehat{\Delta}_w, \hat{A}_w, \hat{S}_{re_w}, \widehat{P}_{1_w}, \widehat{P}_{2_w}, \dots, \widehat{P}_{6_w}]$ using the statistical model parameters defined in Table 2-3. Samples drawn from the distributions of $\widehat{P}_i$ are subject to Constraint 3. Given that u is a uniformly distributed random variable between zero and one, $\hat{A}_w = F_A^{-1}(\hat{u}_w)$, $\widehat{\Delta}_w = F_\Delta^{-1}(\hat{u}_w)$, and $\hat{S}_{re_w} = F_{S_{re}}^{-1}(\hat{u}_w)$, where $F(\cdot)$ is the CDF characterized by A, Δ, and S_{re}, respectively. In order to satisfy Constraint 3 (or equivalently, Equation (2.11)), let $f_{\boldsymbol{P}}(\cdot)$ be the PDF of a 6-dimensional multivariate normal distribution where

$$f_{\boldsymbol{P}}(\boldsymbol{P}|\boldsymbol{\lambda},\boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{6}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} exp\left[-\frac{1}{2}\cdot(\boldsymbol{P}-\boldsymbol{\lambda})^{T}\cdot\boldsymbol{\Sigma}^{-1}\cdot(\boldsymbol{P}-\boldsymbol{\lambda})\right] \tag{2.13}$$

The multivariate distribution is characterized by the mean, $\boldsymbol{\lambda} = [\lambda_{P_1}, \lambda_{P_2}, \dots, \lambda_{P_6}]^T$, and covariance matrix, $\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma} = \mathbf{DRD}$. The correlation coefficient matrix, **R**, quantifies the correlation between the proportion of the total assembly load carried by each of the six eyebars. Each component of **R**, denoted $\mathbf{R}_{s\ell}$ for $s, \ell = 1,2,\dots,6$, describes the linear dependence between monitoring data observations for eyebars s and ℓ (Figure 2-11). The diagonal matrix, **D**, has diagonal elements ζ_{P_i}. Let $\hat{P}_{i\ w}^{N}$ be the elements of a vector, $\widehat{\boldsymbol{P}}_{w}^{N}$, chosen from the multivariate distribution described in Equation (2.13). Then,

$$\hat{P}_{i_w} = \exp\,(\hat{P}_{i\ w}^{N}) \tag{2.14}$$

Given the aforementioned process for drawing random samples from $\widehat{\boldsymbol{X}}_w$ subject to Constraint 1, Figure 2-12 outlines the procedure for conducting the system-level Monte Carlo reliability analysis. The path that ultimately results in system failure and increments the number of failure events, n_f, in the Monte Carlo simulation is highlighted in bold. Following Constraints 2 and 3, system failure only occurs if failure of one or more eyebars ultimately results in the cascading failure of all six eyebars where the loads carried by failed eyebars are redistributed to the remaining eyebars upon failure.

2.4 Results

Section 2.3.2 and Section 2.3.3 formulate the reliability method for the component- (i.e., single eyebar) and system- (i.e., parallel eyebar assembly) level analysis for the US0-LS0

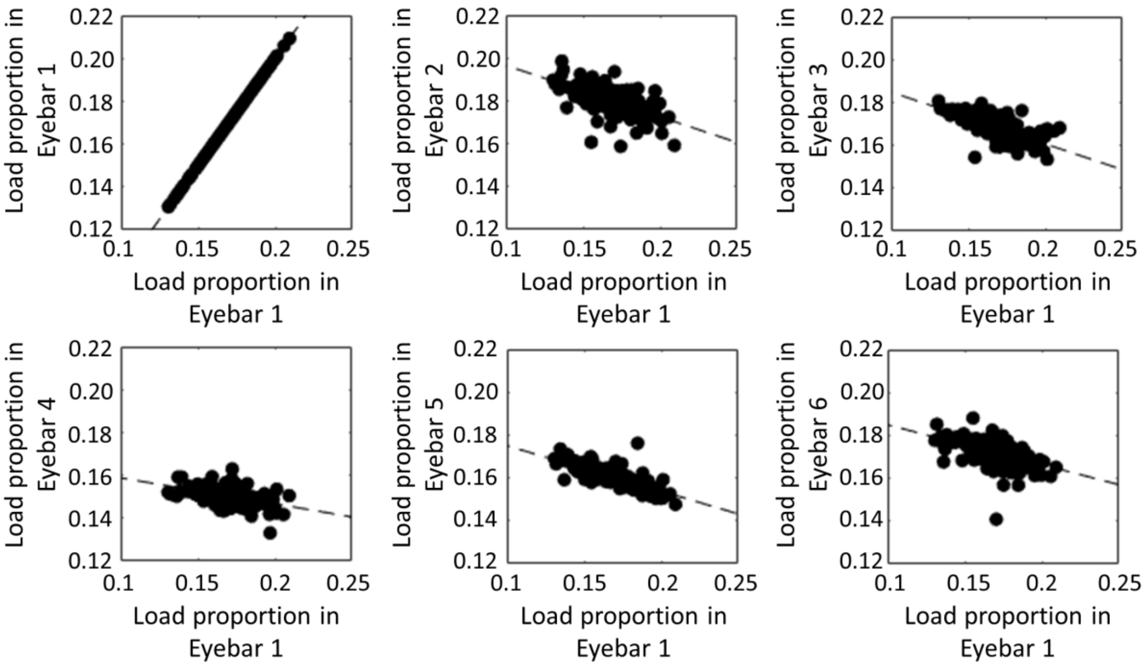


Figure 2-11. Sample data illustrating the correlation between the proportion of the total assembly load carried by Eyebar 1 and the other five eyebars during train events.

assembly's set of parallel eyebar plates with respect to a failure limit state governed by fatigue. For the component reliability problem, the reliability index for each of the US0-LS0 assembly's eyebars is evaluated separately, where each eyebar is considered to be independent of the others. For the system reliability problem, the loads carried by each of the six eyebars are correlated and all of the parallel eyebar plates in the US0-LS0 assembly must fail under the system-level constraints imposed by the mechanics of the US0-LS0 assembly for the entire system to fail. The proposed method develops and evaluates a limit state function that couples fatigue and relative tautness; this methodology is implemented with a scalable long-term monitoring program instrumented on the Harahan Bridge in which the sensing design empowers full consideration of the dead and live loads and associated changes in boundary conditions. Strain and acceleration response data collected by the monitoring system are used to characterize the deterministic and random variable

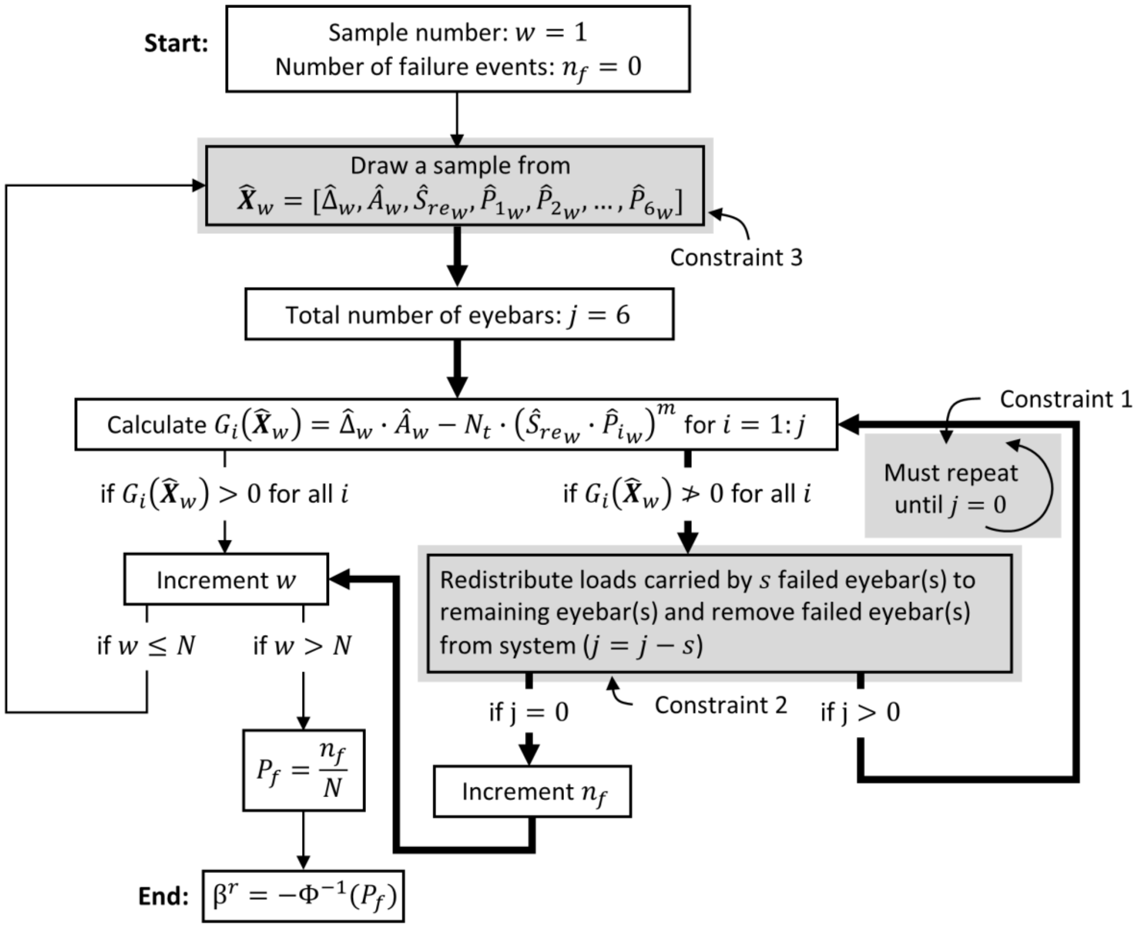


Figure 2-12. Monte Carlo simulation process satisfying Constraints 1 through 3. The failure path that increments the number of failure events is highlighted.

parameters necessary to evaluate the limit state functions. This includes the equivalent stress range, proportion of the total assembly load carried by each of the six eyebars, and the total number of stress cycles. The proposed methodology and long-term monitoring program are translatable across steel truss railroad bridges. If a structural monitoring system is installed at the beginning of a structure's life, the total number of stress cycles during the monitoring period, $N_{t,data}$ (see Table 2-3), equals the total number of stress cycles that occurs over the bridge's lifetime, N_t (see Equation (2.3)), and this work serves as an accurate way to continuously assess the safety of a structure.

For the case study presented in this chapter, the structural monitoring system installed on the Harahan Bridge began collecting data in July 2016, exactly one hundred years after the opening of the bridge. While the monitoring system is used to accurately determine the equivalent stress range bin histogram and the total number of stress cycles that occur within the 16-month monitoring period between July 2016 and October 2017 (i.e., S_{re} and $N_{t,data}$, respectively), the loading history and the total number of stress cycles that occur over the bridge's entire lifespan beginning in July 1916 is unknown. The cumulative number of lifetime stress cycles and equivalent stress range are governed by the frequency of trains (including their length) and their loads. Consequently, the known number of stress cycles during the monitoring period, $N_{t,data}$, is a subset of N_t,

Despite the lack of detailed historical loading information, a conservative assumption can be made about the demand and corresponding equivalent stress range. The first national structural design specification for steel railroad bridges was published by AREMA in 1905 (Unsworth 2010). Since the establishment of these regulations, the design and rating capacity for railroad bridges based on Cooper E series loading doubled over the subsequent decades. The increase in capacity was accompanied by a rapid increase in the demand (i.e., weight per car) on railroad bridges that quickly approached the doubled capacity specifications, resulting in capacities being exceeded in older bridges (Moreu and Spencer 2015, Unsworth 2010). Loads have generally not increased over the past several decades and are not likely to increase in the foreseeable future as railroads have realized that the gap between capacity and demand has already been closed or even exceeded (Moreu and Spencer 2015). Consequently, it is justified to assume that the demand (i.e., weight per car) on the Harahan Bridge before the monitoring period was less

Table 2-4. Total stress cycles accumulating each year and cumulative cycles occurring since the bridge opening.

Year	Monitoring duration (years)	Total cycles during year	Cumulative cycles
July 1916 – June 1917	1	N_1	N_1
July 1917 – June 1918	2	$N_1 \cdot (1+r)$	$N_1 + N_1 \cdot (1+r)$
⋮	⋮	⋮	⋮
July 2016 – June 2017	101	$N_1 \cdot (1+r)^{100} = N_{t,data} \cdot \frac{12}{16}$	$N_1 \cdot \left[\frac{(1+r)^{100}-1}{r}\right]$
⋮	⋮	⋮	⋮
July 1916+y – June 1917+y	$y+1$	$N_1 \cdot (1+r)^{y}$	$N_1 \cdot \left[\frac{(1+r)^{y}-1}{r}\right]$

than or equal to the demand during the monitoring period. A conservative assumption can be made that the equivalent stress range characterized by the 16-month monitoring period is representative of the past and future loading response.

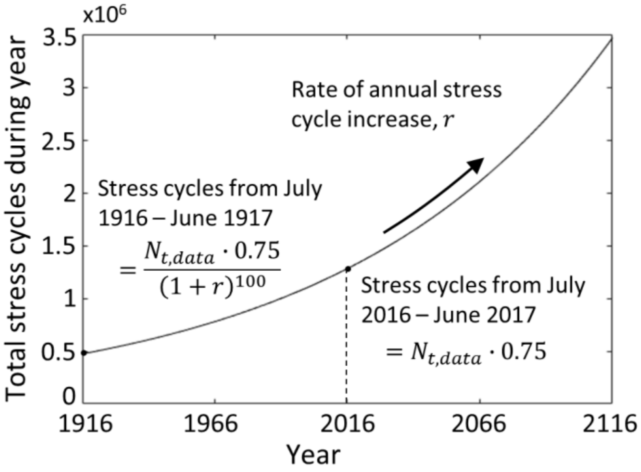
For the purpose of illustrating the proposed methodology, it is assumed that the number of stress cycles occurring each year since the opening of the Harahan Bridge has increased at a rate, r (i.e., $r = 0.01$ for a 1% annual increase), due to increasing use of the bridge, which is a critical network component within a growing regional supply chain network. Given the annual increase rate, r (Constraint 4), and knowledge that the number of stress cycles during the 16-month monitoring period from July 2016 through October 2017 is $N_{t,data}$ (Constraint 5), then the number of stress cycles that occur during each year and the cumulative stress cycles y years after the opening of the bridge in July 1916 are shown in Table 2-4. Given Constraint 4 and Constraint 5, the number of cycles that occurred during the first year the Harahan Bridge was operational is $N_1 = N_{t,data} \cdot$

$\left(\frac{12}{16}\right)/(1+r)^{100}$. From Table 2-4, the cumulative number of stress cycles, N_t, that occurs through y years after the bridge opening in July 1916 is,

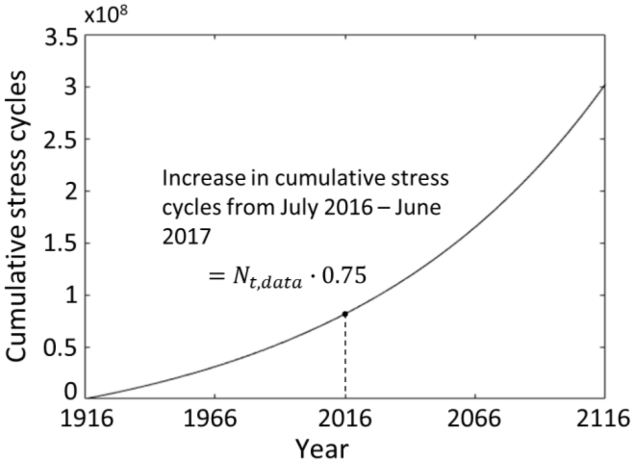
$$N_t(y) = \frac{N_{t,data} \cdot \left(\frac{12}{16}\right)}{(1+r)^{100}} \cdot \left[\frac{(1+r)^y - 1}{r}\right] \tag{2.15}$$

Equation (2.15) ensures that the number of stress cycles that occurs each year increases at the assumed rate of annual increase, r, and that given r, the number of stress cycles that occur between July 2016 and June 2017 is equal to the number of stress cycles measured by the monitoring system (i.e., $N_{t,data} \cdot \frac{12}{16}$). Figure 2-13(a) and Figure 2-13(b) show the increase in total stress cycles during each year for the first 200 years as well as the cumulative stress cycles over the first 200 years, respectively, assuming $r = 0.01$.

Carrying out the methodology proposed in this chapter, Figure 2-14 shows the evolution of the component reliability index for each of the six eyebars of the US0-LS0 assembly given that the number of stress cycles occurring each year increases at a rate, r. These results are based on the conservative assumption that the distribution parameters of the equivalent stress range (Figure 2-9) and relative tautness (Figure 2-10(c)) accurately characterize the past loading behavior and response of trains prior to the installation of the structural monitoring system. Equation (2.15) is used to calculate the total number of stress cycles over the lifetime of the Harahan Bridge. Figure 2-14(a)-(c) show the evolution of the component reliability index values for assumed yearly traffic increase rates of $r = 0\%$, 1%, and 2% respectively. Figure 2-14(d)-(f) highlight the reliability index values near the year 2019 to illustrate that eyebars carrying a greater proportion of the load than expected undergo accelerated fatigue accumulation which decreases the safety of the element (i.e., reduces the reliability index, β_i^r). That is, Eyebar 4 carries the lowest proportion of the



(a)



(b)

Figure 2-13. (a) Total number of stress cycles that occur during each year following the opening of the bridge in July 1916 (e.g., during the year July 2016 – June 2017); (b) cumulative number of stress cycles that occur since the opening of the Harahan Bridge in July 1916.

total assembly load (i.e., it is the least taut) and Eyebar 2 carries the greatest proportion of the total assembly load (i.e., it is the tautest).

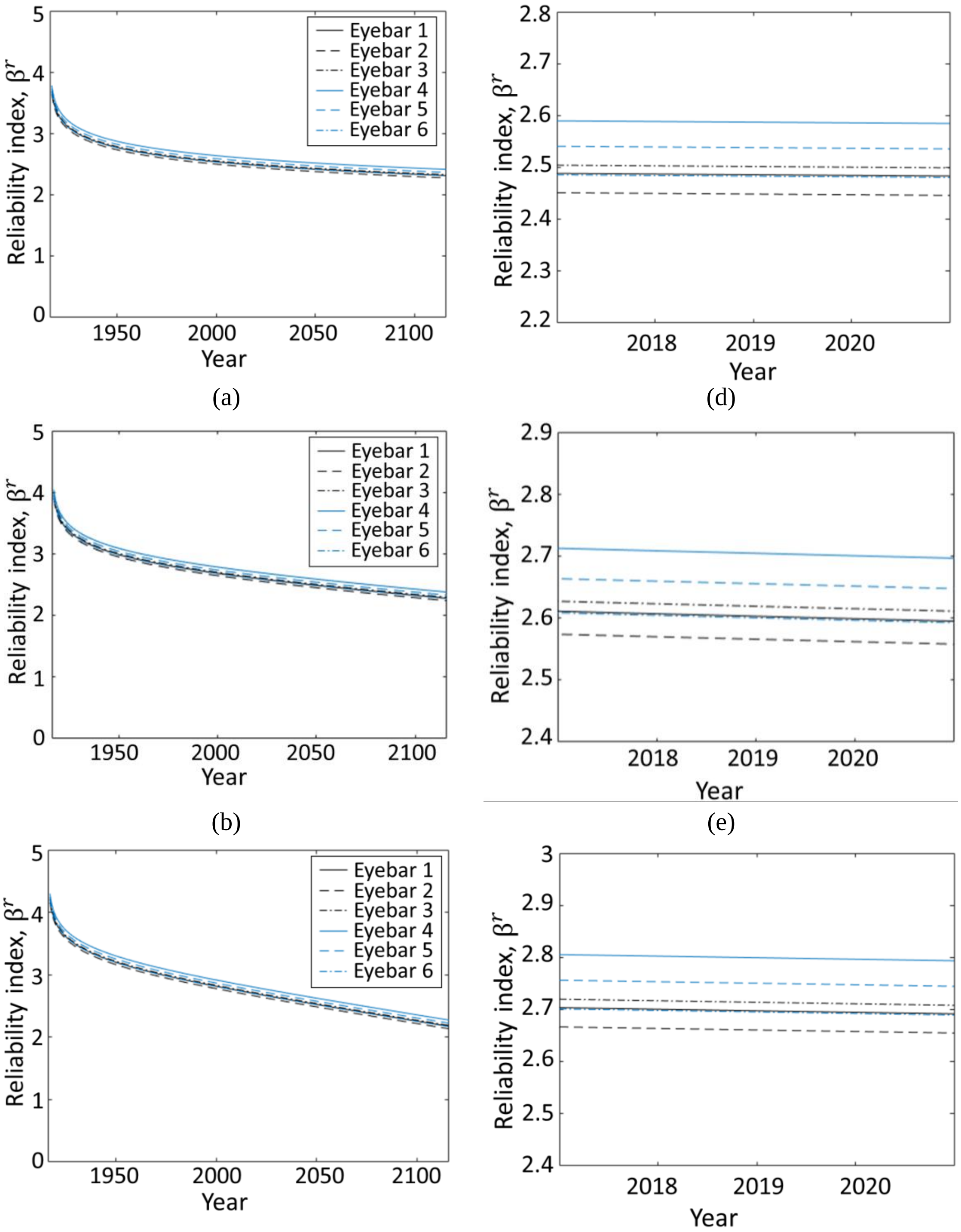
Similarly, Figure 2-15(a) shows the evolution of the system reliability index values for assumed yearly traffic increase rates of $r = 0\%$, 1%, and 2%. The reliability index values for the system reliability problem shown in Figure 2-15(a) closely resemble the results for the component reliability problem shown in Figure 2-14. Close comparison of these two results reveals that when the conservative assumption is made that the US0-LS0 assembly's parallel eyebar system is brittle—meaning that when an eyebar violates the limit state function (i.e., $G(\boldsymbol{X}) \leq 0$) the eyebar is no longer in service and the loads are redistributed to the remaining eyebars immediately—then first failure is tantamount to system failure. That is, extreme load events in the tail of the equivalent stress range, S_{re}, ultimately govern the failure limit state function and progressive failure initiates when the eyebar carrying the greatest proportion of the total assembly load fails. Even though the parallel set of eyebars is intended to increase system resilience by introducing redundancy into this fracture-critical assembly, the failure of a single eyebar leads to failure of the entire parallel system. This result exposes an unintended vulnerability in the system's structural design—under the condition that it has brittle behavior—that reduces its resilience. Consequently, the described system-level probabilistic fatigue assessment method can be used to enable bridge owners and designers to better design eyebar assemblies subject to their most conservative behavior to ensure redundancy in the parallel system given that parallel eyebar assemblies are prone to uneven loading distributions that can result in accelerated fatigue rates in some members. This includes determining the number of parallel eyebar plates necessary to introduce redundancy in the parallel system under extreme loading conditions and quantifying a threshold on the proportion of the total assembly load carried by a single eyebar that initiates eyebar assembly realignment. In reality, it is expected that system failure would occur in a much more ductile manner in which there is reserve capacity in eyebar(s) with fatigue cracks offering a period of time before the eyebar(s) experience rapid, rather than incremental, crack growth and redistribute loads to the remaining eyebars. When this time period is measured in at least weeks or months, the time lag would permit the owners to identify the local behavior based on monitoring data or see cracks during visual inspection allowing maintenance to be prescribed before system failure could occur.

To illustrate the impact of existing asset management practices (i.e., eyebar assembly realignment) on the evolution of the system-level safety profile of the eyebars, consider the system reliability results in Figure 2-15(a). The bridge owner can initiate standard maintenance practices to realign the six US0-LS0 eyebars at any point in time in order to redistribute the assembly load uniformly across the eyebars. Let y_m denote the number of years that occur between the bridge opening and the realignment. Based on Equation (2.3) and Equation (2.15) the limit state function governing the two time periods (i.e., before and after maintenance) is

$$G_i(\boldsymbol{X}) = \Delta - \frac{N_t(y)\cdot\left(S_{re}\cdot P_i{}^1\right)^m}{A}, i = 1,2,\dots,6 \text{ for } y \le y_m \tag{2.16}$$

$$G_i(\boldsymbol{X}) = \Delta - \frac{N_t(y_m)\cdot\left(S_{re}\cdot P_i{}^1\right)^m}{A} - \frac{[N_t(y)-N_t(y_m)]\cdot\left(S_{re}\cdot P_i{}^2\right)^m}{A}, i = 1,2,\dots,6 \text{ for } y_m < y \tag{2.17}$$

Here, y is the number of years that occur between the bridge opening and the time at which the reliability index is being evaluated. $P_i{}^1$ and $P_i{}^2$ are random variables characterizing the proportion of the total assembly load carried by each eyebar before and after the realignment, respectively. Following the procedure outlined in Section 2.3.3 to calculate the system reliability index, Figure 2-15(b)-(c) illustrate the impact of realignment on the safety profile of the system assuming $y_m =$ 34, 64, 84, and 104—meaning that



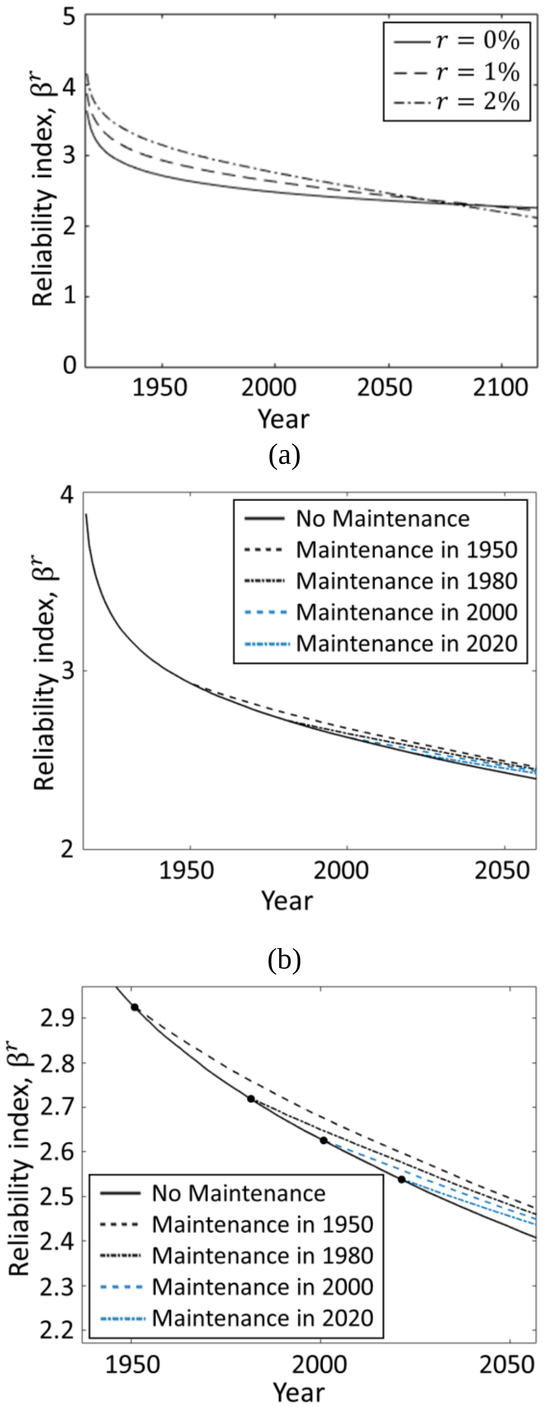
(a) (d)

(b) (e)

(c) (f)

Figure 2-14. Evolution of component reliability index for an increase in annual traffic of (a) $r = 0\%$, (b) $r = 1\%$, and (c) $r = 2\%$. Current reliability index (zoomed in) for (d) $r = 0\%$, (e) $r = 1\%$, and (f) $r = 2\%$.

maintenance takes place in 1950, 1980, 2000, or 2020. Because the limit state function



(a)

(b)

(c)

Figure 2-15. Evolution of the system reliability index (a) without maintenance, (b) with maintenance for $r = 1\%$, and (c) with maintenance for $r = 1\%$ (zoomed in).

derived in this chapter captures the dependency between fatigue and relative tautness, the slope of the system reliability index decreases after the realignment, thus increasing the amount of time until a bridge owner's lower safety threshold is reached. As expected, the reliability index continues to decrease even after realignment—albeit at a slower rate—because fatigue damage continues to accumulate under the evenly distributed expected loads. Let $\Delta\beta_M^r$ denote the slope of the reliability index trajectory for the year immediately following the realignment in Figure 2-15(b) and let $\Delta\beta_{NM}^r$ denote the slope of the reliability index trajectory for that same year assuming no maintenance occurs. Even though each of the eyebars carries a proportion of the total assembly load that deviates within just a few percent of what is expected under ideal conditions (see Figure 2-10(c)), the immediate ratio of the change in slope of the safety profile, $\Delta\beta_{NM}^r/\Delta\beta_M^r$, is approximately 0.75 for maintenance conducted in 1950, 1980, 2000, or 2020. This result suggests that the proposed method can be used as in intervention strategy to inform asset management decisions in the future in which the owner seeks to keep the reliability index above a safety threshold over a desired time horizon.

2.5 Conclusions

Despite the fact that fatigue damage in fracture-critical tensile members is a primary concern for railroad bridge owners, existing code- and inspection-based maintenance techniques do not sufficiently assess remaining fatigue life. As railroads continue to age, there is a need for data-driven methods to reduce uncertainty and continuously quantify the safety profile of structural elements. Due to the history of codified decision making in practice, the most feasible strategies are those that complement existing management strategies.

This chapter presents the application of a scalable and fully automated wireless sensing system on the Harahan Bridge that is used to assess probabilistic fatigue life coupled with relative tautness of the US0-LS0 tensile eyebar assembly. A significant contribution of this work is that it implements a novel holistic approach to tracking fatigue accumulation in which the sensing design empowers full consideration of the dead and live loads and associated changes in boundary conditions. The contribution of changes in boundary conditions and secondary deterioration mechanisms on structural safety is critically important, yet often ignored due to the emphasis of existing sensing strategies on monitoring global structural behavior. There is a need to strategically implement sensing systems that are tailored to the mechanics of the monitored structural component in order to capture the full behavior of structural components. The proposed method is purely data-driven and is carried out using 16 months of continuous, long-term structural response data collected for each passing train. Consequently, the proposed probabilistic methodology operates without the need for train parameter estimates, FE models, or controlled loading tests. This method is synergistic with existing maintenance strategies because it accounts for the influence of changes in the boundary conditions of the elements on the safety of the asset with respect to fatigue; the acceleration data used to inform the relative tautness analysis is consistent with existing visual inspection techniques.

Both the component and system reliability problems are formulated and evaluated using the proposed data-driven approach. The system-level fatigue reliability analysis requires careful consideration of a number of complicating constraints that are induced by the physical system configuration and load path. The resulting evolutions of the component and system reliability index values indicate that under the conservative assumption that progressive failure is brittle, first failure within the parallel eyebar system is generally equivalent to system failure. This analysis also illustrates that the relative tautness among the parallel eyebars influences the evolution of the component or system's safety profile with respect to the fatigue-induced failure limit state. Specifically, the proposed methodology can quantify the impact of existing asset management practices (i.e., eyebar assembly realignment) on the current and future evolution of the system-level safety profile. Consequently, this work provides the foundation for informing data-driven intervention strategies for decision making.

An exact calculation of the reliability index requires knowledge of the full loading history and the accumulated lifetime stress cycles. Since the sensing system presented in this chapter was installed on the Harahan Bridge nearly a century after the bridge's opening, future work will include obtaining historical track usage and loading estimates from the bridge owner. While beyond the scope of this chapter, loading information such as the frequency and length of trains collected from alternative data sources such as the bridge owner's manifest data and wheel impact load detector data can inform the total number of stress cycles that occur over the bridge's lifetime, N_t. Despite the lack of measured monitoring data prior to the sensing system installation, Section 2.4 justifies conservative assumptions for estimating the total stress cycles and loads occurring over the assembly's lifetime. Also, the chapter presents a conservative lower bound on the parallel eyebar system's reliability index assuming that if one eyebar exceeds its limit state, then its load would immediately redistribute to the other eyebars leading to a seemingly brittle failure mechanism. As previously mentioned, there remains axial load capacity in the cracked eyebar past the limit state resulting in a ductile failure mechanism. Future work should explore the parallel eyebar system after crack initiation to more accurately identify the reliability of the parallel eyebar component. Such future work would lead to an even more realistic approach to assessing the component reliability and would also offer some insight to the time scales over which progressive failure of the component would occur, allowing for inspection cycles to be optimally timed. Additionally, future work aimed at modelling repeated eyebar realignment will help to inform optimally timed maintenance efforts that maximize the remaining fatigue life of the bridge.

2.6 References

American Association of State Highway and Transportation Officials (AASHTO). *LRFD bridge design specifications*. Washington, D.C: AASHTO, 2012.

American Railway Engineering and Maintenance-of-Way Association (AREMA). *Manual for railway engineering*. Landover: AREMA, 2016.

ASTM International. ASTM standard e 1049-85 standard practices for cycle counting in fatigue analysis. In: *Annual book of ASTM standards*. Vol. 03.01. Philadelphia: ASTM, 2017.

Azizinamini A, Power EH, Myers GF, et al. Fatigue and fracture of steel structures. In: D Trackman (ed) *Design guide for bridges for service life*. Washington, DC: Transportation Research Board, 2014, pp. 316-342.

Barker RM and Puckett JA. *Design of highway bridges*. Hoboken: John Wiley & Sons, 2013.

Blevins RD. *Formulas for dynamics, acoustics and vibration*. 1st ed. Somerset: John Wiley & Sons, 2016.

Chen ZW and Xu YL. SHMS-based fatigue reliability analysis of multiloading suspension bridges. *J Struct Eng* 2012; 138: 299-307.

Computers and Structures Inc. (CSI). CsiBridge 2016: introduction to CsiBridge, 2016.

Fang Z and Wang J. Practical formula for cable tension estimation by vibration method. *J Bridge Eng* 2012; 17: 161-164.

Flanigan KA, Johnson N, Hou R, et al. Utilization of wireless structural health monitoring as decision making tools for a condition and reliability-based assessment of railroad bridges. In: *SPIE sensors and smart structures technologies for civil, mechanical, and aerospace systems* (ed JP Lynch), Portland, OR, 26 March – 29 March 2017, paper no. 101681X, pp. 1-17. Bellingham: SPIE.

Frangopol DM, Strauss A and Kim S. Bridge reliability assessment based on monitoring. *J Bridge Eng* 2008; 13: 258-270.

Gostautas RS and Tamutus TA. SHM of the eyebars of the old San Francisco Oakland bay bridge. In: *Proceedings of the 10th international workshop on structural health monitoring* (eds F Chang and F Kopsaftopoulos), Menlo Park, CA, 1 September – 3 September 2015. Lancaster: DEStech Publications, Inc.

Hou R, Jeong S, Law K, et al. Reidentification of trucks in highway corridors using convolutional neural networks to link truck weights to bridge responses. In: *SPIE sensors and smart structures technologies for civil, mechanical, and aerospace systems* (eds JP Lynch, H Huang, H Sohn, et al.), Denver, CO, 4 March – 7 March 2019, paper no. 109700P, pp. 1-14. Bellingham: SPIE.

Imam B, Righiniotis TD, Chryssanthopoulos MK, et al. Analytical fatigue assessment of riveted rail bridges. *Proc Inst Civ Eng – Bridge Eng* 2006; 159: 105-116.

Keating PB and Fisher JW. Evaluation of fatigue tests and design criteria on welded details. Report, National Cooperative Highway Research Program (NCHRP). Report no. 286, 1986. Washington, D.C.: Transportation Research Board.

Kim S, Lee S and Mha H. Fatigue reliability assessment of an existing steel railroad bridge. *Eng Struct* 2001; 23: 1203-1211.

Kollár L. *Structural stability in engineering practice*. Boca Raton: CRC Press, 2003.

Kunz PM, Bruehwiler E and Hirt MA. Evaluation of the remaining fatigue life of steel bridges. In: *CSCE proceedings of 4th international conference on short and medium span bridges*, Halifax, Nova Scotia, Canada, 8 August – 11 August 1994, pp. 1219-1230. Montreal: CSCE.

Kurata M, Kim J and Lynch JP. Internet-enabled wireless structural monitoring systems: development and permanent deployment at the New Carquinez suspension bridge. *J Struct Eng* 2013; 139: 1688-1702.

Kwon K and Frangopol DM. Bridge fatigue reliability assessment using probability density functions of equivalent stress range based on field monitoring data. *Int J Fatigue* 2010; 32: 1221-1232.

Lee Y, Pan J, Hathaway R and Barkey M. *Fatigue testing and analysis - theory and practice*. 1st ed. Burlington: Elsevier, 2005.

Li H, Frangopol DM, Soliman M, et al. Fatigue reliability assessment of railway bridges based on probabilistic dynamic analysis of a coupled train-bridge system. *J Struct Eng* 2016; 142: 04015158.

Miner MA. Cumulative damage in fatigue. *J Appl Mech* 1945; 12: 159-164.

Moreu F and Spencer BF Jr. Framework for consequence-based management and safety of railroad bridge infrastructure using wireless smart sensors (WSS). Report, Newmark Structural Engineering Laboratory Report Series 041, University of Illinois at Urbana-Champaign, June 2015.

Nikolaidis E, Ghiocel DM and Singhal S. *Engineering design reliability*. 1st ed. Boca Raton: CRC Press, 2004.

Nugroho G, Priyosulistyo H and Suhendro B. Evaluation of tension force using vibration technique related to string and beam theory to ratio of moment of inertia to span. *Procedia Eng* 2014; 95: 225-231.

O'Connor SM, Zhang Y, Lynch JP, et al. Long-term performance assessment of the Telegraph Road bridge using a permanent wireless monitoring system and automated statistical process control analytics. *Struct Infrast Eng* 2017; 13: 604-624.

Priyosulistyo H and Ferdina M. Force prediction of cylindrical steel beams on pinned-pinned supports under axial load by vibration technique. *Procedia Eng* 2014; 95: 178-184.

Rakoczy AM, Nowak AS and Dick S. Fatigue reliability model for steel railway bridges. *Struct Infrast Eng* 2016; 12: 1602-1613.

Reid RL. Damaged eyebar section replaced on San Francisco-Oakland bay bridge. *Civil Eng* 2010; 80: 18-22.

Sobczyk K and Spencer BF Jr. *Random fatigue: from data to theory*. San Diego: Academic Press, Inc, 1992.

Swartz RA, Jung D, Lynch JP, et al. Design of a wireless sensor for scalable distributed in-network computation in a structural health monitoring system. In: *Proceedings of the 5th international workshop on structural health monitoring* (eds F Chang and F Kopsaftopoulos), Menlo Park, CA, 12 September – 14 September 2005, pp. 1570-1577. Lancaster: DEStech Publications, Inc.

Tobias DH and Foutch DA. Reliability-based method for fatigue evaluation of railway bridges. *J Bridge Eng* 1997; 2: 53-60.

Union Pacific Railroad. Bridge assessment inspection reference cards. Report no. PB-22046, 4 October 2013. Omaha: Union Pacific Railroad.

Unsworth JF. *Design of modern steel railway bridges*. Boca Raton: CRC Press, 2010.

West Virginia Department of Transportation. Silver bridge, https://transportation.wv.gov/highways/bridge_facts/Modern-Bridges/Pages/Silver.aspx. (2019, accessed 31 March 2019).

Wirsching PH. Fatigue reliability for offshore structures. *J Struct Eng* 1984; 110: 2340–2356.

Wirsching PH, Ortiz K and Chen YN. Fracture mechanics fatigue model in a reliability format. In: *OMAE proceedings of the sixth international symposium on offshore mechanics and arctic engineering* (eds JS Chung, S Berg, SK Chakrabarti et al.), Houston, TX, 1987.

Zhao Z, Haldar A and Breen F Jr. Fatigue-reliability evaluation of steel bridges. *J Struct Eng* 1994; 120: 1606-1623.

Zhou EY. Assessment of bridge remaining fatigue life through field strain measurement. *J Bridge Eng* 2006; 11: 737-744.

CHAPTER 3

Quantitatively Linking Long-Term Monitoring Data to Condition Ratings Through a Reliability-Based Framework

3.1 Introduction

Structural health monitoring (SHM) of aging highway infrastructure has shown continued growth over the past several decades in an effort to guide data-driven asset management decisions (Seo et al. 2016). Despite this growth, there is a persistent discontinuity between algorithmic strategies used in SHM and the ability of bridge managers to make immediate and long-term asset management decisions in practice (Figure 3-1). This is in part due to the focus of physics-based SHM methods on estimating structural state and detecting damage by monitoring changes in structural properties such as global modal parameters (i.e. modal frequencies and mode shapes) (Cawley and Adams 1979, Doebling et al. 1998, Peeters and De Roeck 2001, Kim et al. 2003, Whelan and Janoyan 2010). The fundamental principle of frequency-based damage detection methods is that modal parameters are functions of physical properties such as mass and stiffness. Vibration-based data collected from SHM systems have been proposed to be used to quantify changes in modal properties associated with damage-induced changes in mass, or more commonly, stiffness. However, for many operational structures, long-term damage and deterioration may not necessarily correspond to mass or stiffness changes, making many of these physics-based SHM methods ineffective. Damage and deterioration can lead to changes in an element's

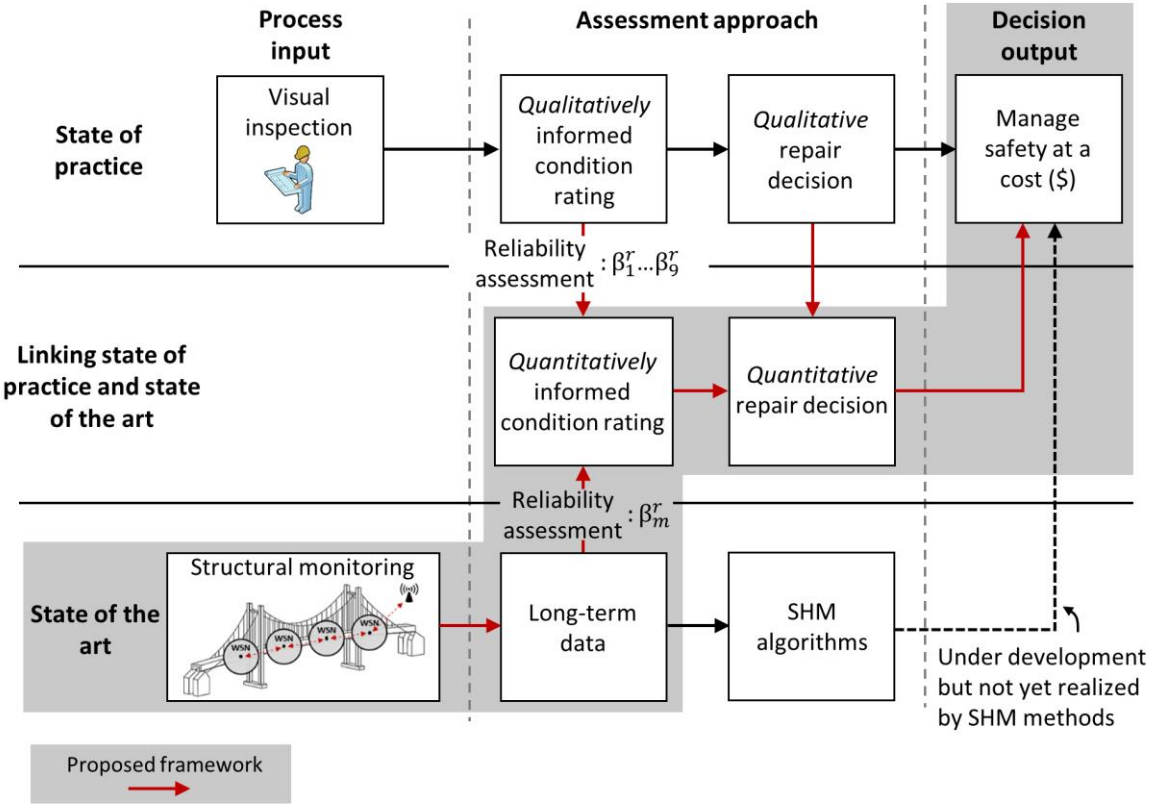
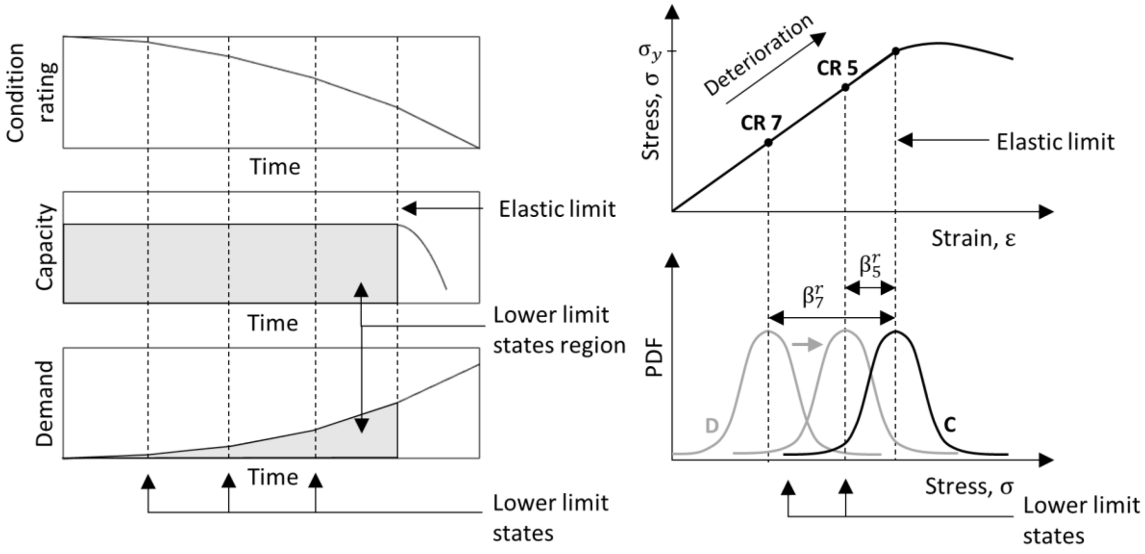


Figure 3-1. Proposed framework (highlighted by the shaded region) linking the state of practice and state of the art in bridge asset management.

boundary conditions that can change the very nature of a system or component's safety profile without that system or component leaving the linear response regime. For example, corrosion, scour, fatigue, and spalling can change a system or component's safety without causing any measurable change in stiffness (MDOT 2016). This is reflected in the United States' extensive history of codified decision-making practices in which actionable upkeep decisions are made based on condition ratings (CR) with the bridge behaving entirely in the elastic regime (Ettouney and Alampalli 2016).

The Federal Highway Administration (FHWA) has relied on the nationally mandated National Bridge Inventory (NBI) program to manage inspection schedules, implement inspection quality control, and establish a common scale for asset management decision making (Frangopol et al. 2001). During regular (e.g., biennial) visual inspections, inspectors rate bridge components and assign condition ratings on a numeric scale (e.g., 0 through 9, where the condition rating number decreases with increasing damage) to describe the level of deterioration they see visually in a bridge component (MDOT 2016). Asset management decisions surrounding upkeep and maintenance are subsequently made directly from the assigned condition ratings. This process is illustrated as the "State of practice" in Figure 3-1. This qualitative method of inspection fosters reactive decision making where repairs are made based on periodic visual examination of existing damage or deterioration. Following this approach, damage is crudely described by discrete condition states assigned by an inspector. Furthermore, visual inspections may not uncover underlying deterioration (Agrawal et al. 2009). Due to these drawbacks, there is a general consensus within the engineering community that asset management decisions could be aided by quantitative monitoring data in order to reduce uncertainty in assessing structural condition (i.e. "State of the art" in Figure 3-1) (Frangopol et al. 2001). Monitoring can also be continuous, allowing asset managers to more regularly, if not continuously, track the condition of a structure, especially between biennial inspections. However, the state of practice (i.e. visual inspection) and state of the art (i.e. SHM) continue to remain as two independent processes. Because existing SHM methods largely operate independent of condition ratings, upkeep decisions based on monitoring data are not generally made in practice, leading to limited commercial adoption of SHM for bridges.

The author proposes in this study that rather than operating as two separate processes, the current state of practice and SHM be reconciled (i.e. "Linking state of

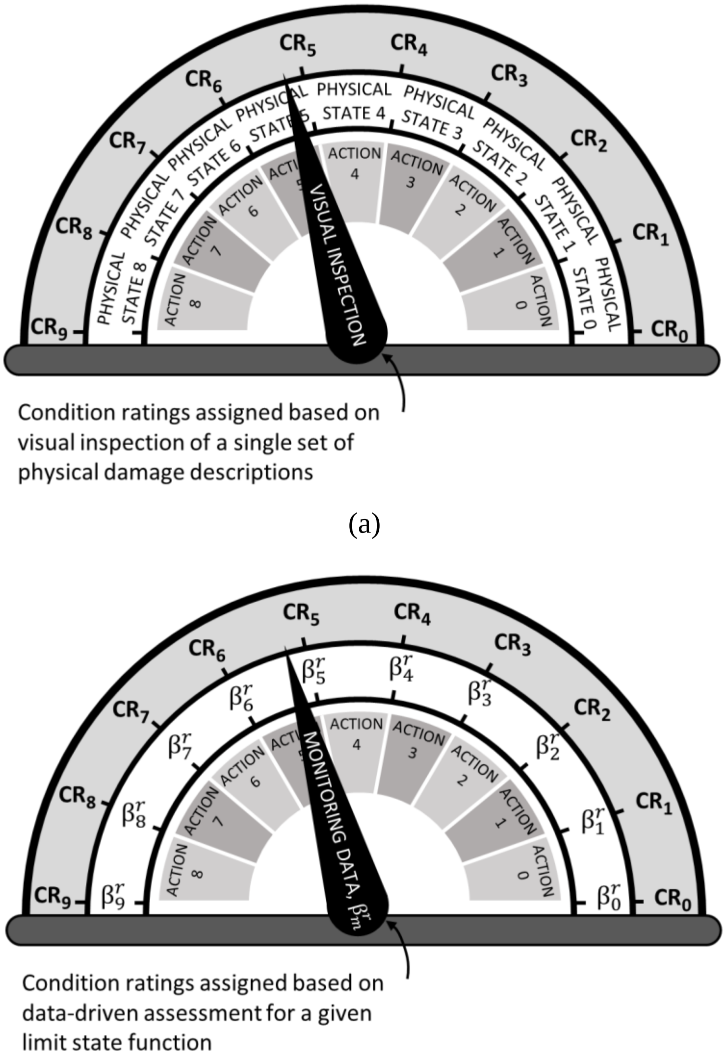


(a) (b)

Figure 3-2. (a) Relationship between condition ratings, capacity, and demand for the same load profile but with increasing structural deterioration over time; (b) an equivalent representation that reflects how the reliability index (which corresponds to specific condition ratings) decreases as deterioration increases.

practice and state of the art" in Figure 3-1). The strategy proposed herein aims to formally link structural monitoring data with condition ratings. The benefit of such a mapping is that it allows monitoring data to be used to make decisions based on quantitative information that encompasses the entire measurable domain of damage that may exist in an asset, as opposed to those only informed by visual inspection. This approach respects that 1) there is a proven track record between effective decisions being made by condition ratings assigned by visual inspections, and 2) quantitative monitoring data can reduce uncertainty in assessing structural condition. As shown in Figure 3-1, the proposed framework explicitly links visual inspection practice and SHM approaches using reliability methods, which have matured in civil engineering applications over the past several decades and are used to define structural design codes (Elishakoff 1983, Estes and Frangopol 1991, Frangopol and Estes 1997, Frangopol et al. 2008, Nowak and Zhou 1990, Rausand 2004, Tabsh and Nowak 1991). Reliability methods rigorously consider the effects of uncertainty in an asset's capacity and demand, and provide a scalar measure of the safety of a system or component as measured by the reliability index, β^r (Nikolaidis et al. 2004). A higher reliability index value, β^r, represents a higher level of safety (i.e., lower probability of failure) than a lower value. The ability of reliability methods to quantify structural safety based on changes in the structure as it operates in its elastic range is necessary to link state-of-practice and state-of-the-art approaches. Condition ratings can then be thought of as decision points for the asset manager (as the structure operates in the elastic regime) that correlate to the reliability of the structure. In this manner, lower condition ratings implicitly suggest a higher probability of failure.

Ettouney and Alampalli (2016) defines the theoretical limit states within the elastic range for which bridge managers make asset management decisions based on codified condition ratings as "lower limit states" (Figure 3-2(a)). Since condition ratings are often specified by physical deterioration descriptions (e.g., percent section loss), these quantitative descriptions can be used to define the loss of structural integrity associated with each condition rating leading to the definition of lower limit states in terms of the capacity and demand. Figure 3-2(a) illustrates that under the same load profile, an asset's condition rating decreases over time as deterioration accumulates, thereby increasing demand on the element without changing the structural capacity. This is equivalently represented by Figure 3-2(b), which illustrates that as deterioration increases there is a corresponding decrease in condition rating and increase in demand resulting in the reliability index, β^r, decreasing. *Consequently, there theoretically exist quantitative reliability index values,* β_i^r, *associated with lower limit states described by each existing*



Condition ratings assigned based on visual inspection of a single set of physical damage descriptions

(a)

Condition ratings assigned based on data-driven assessment for a given limit state function

(b)

Figure 3-3. (a) The state of practice assigns condition ratings and repair actions based on qualitative visual assessment of an asset's physical condition, while (b) the proposed framework assigns equivalent condition ratings corresponding to reliability index thresholds, β_i^r, and repair actions based on an asset's current safety level, β_m^r.

condition rating, i, for $i = 0,1,\ldots,9$ (assuming a 10-level rating scale). Given the asset's load profile, the reliability index values, β_i^r, are quantitative measures of the safety of an asset with a physical condition state described by each condition rating. These values, β_i^r, serve as thresholds that quantify the accepted probability of failure (as defined by the limit state) for which maintenance, repair, and replacement decisions are made for each condition rating in practice today.

The method of linking structural monitoring data with existing condition ratings using reliability methods discussed herein comprises two primary stages. In the first stage, the author presents a method to objectify lower limit states by quantifying the reliability index thresholds, β_i^r, corresponding to existing condition ratings. This is conceptually illustrated by the dashboards in Figure 3-3, where the physical descriptions of damage in Figure 3-3(a) are replaced by associated quantitative reliability index thresholds in Figure 3-3(b). In the second stage, we introduce a method to continuously calculate the reliability index of an in-service asset based only on long-term monitoring data (we refer to this as a "data-driven" approach because the only input is bridge response data); this is conceptually illustrated by the moving dial in Figure 3-3(b) which replaces visual inspection in Figure 3-3(a). Since reliability index thresholds are already established for establishing condition ratings as lower limit states in the first stage, the measured reliability index, β_m^r, calculated in the second stage serves a direct indicator of the in-service asset's condition rating. By replacing qualitative condition assessment practice with a quantitative (and automated) assessment metric, bridge owners would be able to continuously track the structural condition of the asset at any point in time, and the same decision output that is derived from the state of practice today is achieved, albeit with the repair action guided by quantitative rather than qualitative assessment methods. Additionally, since the reliability index threshold values, β_i^r, correspond to the probability of failure governing each condition rating, the data-driven performance, β_m^r, corresponding to *any* measurable limit state function can be assigned to a condition rating, not just a limit state corresponding to the deterioration and failure mechanisms traditionally assessed using visual inspection. This is justified because the reliability index, and corresponding probability of failure, is a universal measure across any failure limit state. After detailing the theoretical background, this chapter implements the full two-stage data-driven reliability framework for tracking structural performance on an instrumented pin-and-hanger assembly on the Telegraph Road Bridge (TRB), which is a steel girder highway bridge located in Monroe, MI.

3.2 Methodology

This section outlines the methodology for quantitatively linking long-term monitoring data and condition ratings through a reliability-based framework. The following section comprises three primary components. First, Section 3.2.1 presents the necessary data-driven first-order reliability method (FORM) used to calculate the reliability index throughout this chapter. An emphasis is placed on accounting for correlation between measured, non-normal random variables (RV) within the FORM analysis, which is a condition that is imposed when there is correlation between random variables measured from monitoring data due to dependencies in the structural response at the sensor locations. Section 3.2.2.1 then outlines the first stage of the proposed framework in which the reliability index threshold values, β_i^r, associated with condition rating lower limit states are determined. Section 3.2.2.2 introduces the second stage of the proposed framework and describes the procedure for measuring the in-service reliability index of the asset, β_m^r, using long-term monitoring data; the measured safety of the asset, β_m^r, maps directly to the condition ratings defined in Section 3.2.2.1.

3.2.1 Overview of Reliability Methods

The goal of structural reliability is to calculate the reliability index, β^r, which is a scalar measure of safety of a system or component with respect to a failure limit state (Nikolaidis et al. 2004). A limit state function, $G(\boldsymbol{X}) = C - D$, is described by an *n*-dimensional vector of random variables, $\boldsymbol{X} = [X_1, X_2, \ldots, X_n]^T$, the load effect, D, and the resistance, C. For the single component reliability problem, the probability of failure, $P_f = P(\boldsymbol{X} \in \Omega)$, is defined as

$$P_f = \int_{\Omega} f_{\boldsymbol{X}}(\boldsymbol{X}) d\boldsymbol{X} \tag{3.1}$$

where $f_{\boldsymbol{X}}(\boldsymbol{X})$ is the joint probability density function (PDF) of $\boldsymbol{X}$, and $\Omega \equiv G(\boldsymbol{X}) \leq 0$ is the failure domain. When the limit state function is linear in the independent standard normal space (ISNS), the reliability index, β^r, is defined as the minimum distance from the origin to the failure domain of the limit state function, Ω, in the ISNS. When the limit state function is not linear in the ISNS, first-order approximations can be used to approximate the reliability index, β^r, in which case the reliability index is defined as the approximate, rather than absolute, minimum distance. The reliability index, β^r, is a direct indicator of the probability of failure, where

$$P_f \cong \Phi(-\beta^r) \tag{3.2}$$

and $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF) (Nikolaidis et al. 2004). The equality in Equation (3.2) holds if the limit state function is linear in the ISNS.

When no closed-form solution exists for the integral in Equation (3.1), FORM can be used to approximate the probability of failure by linearizing the limit state function at an optimal point in the ISNS (Nikolaidis et al. 2004). While the use of FORM to evaluate structural reliability has become ubiquitous across the field of civil engineering, there are very few examples of data-driven reliability assessments that use long-term monitoring data, especially ones that account for correlation between non-normal random variables measured from monitoring data (Frangopol et al. 2008, Catbas et al. 2008). The iterative FORM procedure utilized in this chapter is presented in Algorithm 3-1 in order to introduce notation and to elaborate on how correlation between non-normal random variables is accounted for within the analysis (such as is the case in Section 3.5). For limit state functions that are highly nonlinear in the standard normal space, second-order reliability methods are necessary to approximate the probability of failure with higher accuracy. The FORM results presented herein are compared to those generated by Monte Carlo simulations to confirm that first-order methods (namely, FORM) are sufficient.

Algorithm 3-1. FORM algorithm procedure for calculating the reliability index assuming 1) independent random variables, 2) dependent (possibly) non-normal random variables.

Step 1:

Identify the limit state function, $G(\boldsymbol{X})$, and transform $G(\boldsymbol{X})$ into the standard normal space, $G_U(\boldsymbol{U})$,

$$G(\boldsymbol{X}) = G(X_1, X_2, \dots, X_n) \rightarrow \boldsymbol{U} = T(\boldsymbol{X}) \rightarrow G_U(\boldsymbol{U}) = G_U(U_1, U_2, \dots, U_n)$$

where $\boldsymbol{U} = T(\boldsymbol{X})$ is a one-to-one transformation of $\boldsymbol{X}$ to the standard normal space.

Case 1.1

For the case of independent random variables,

$$\boldsymbol{U} = \Phi^{-1}\big(F_{\boldsymbol{X}}(\boldsymbol{X})\big) \tag{3.3}$$

where $F_{X_i}(X_i)$ is the CDF.

Case 1.2

For the case of dependent (possibly) non-normal random variables, there are two primary methods that can be used to uncorrelate the random variables: the Rosenblatt Transformation and the Nataf Transformation (Nikolaidis et al. 2004). The Nataf Transformation is well suited for the analysis in this chapter because the dependent random variables considered herein (Table 3-5), with prescribed marginal CDFs and correlation coefficients, are jointly normal when marginally transformed. To transform the correlated non-normal random variables into the uncorrelated standard normal space using the Nataf Transformation, the lower Cholesky decomposition, $\boldsymbol{L}$, of $\boldsymbol{R}$ is utilized such that,

$$\boldsymbol{U} = \boldsymbol{L}^{-1}\Phi^{-1}\big(F_{\boldsymbol{X}}(\boldsymbol{X})\big) \tag{3.4}$$

where $\boldsymbol{R}$ is the matrix of correlation coefficients and $\boldsymbol{R} = \boldsymbol{L}\boldsymbol{L}^T$.

Step 2:

Evaluate the gradient $\nabla G_U(\boldsymbol{U}) = \left[\frac{\partial G_U(\boldsymbol{U})}{\partial U_1}, \frac{\partial G_U(\boldsymbol{U})}{\partial U_2}, \dots, \frac{\partial G_U(\boldsymbol{U})}{\partial U_n}\right]$.

Step 3:

Begin an iterative process with $i = 0$.

Step 4:

Assume the design point's initial value, $\boldsymbol{X}^{*(i)} = \left[X_1^{*(i)}, X_2^{*(i)}, \dots, X_n^{*(i)}\right]$ to obtain $\boldsymbol{U}^{*(i)} = T\big(\boldsymbol{X}^{*(i)}\big)$.

Step 5:

Estimate the direction cosines, $\boldsymbol{\alpha}^{(i)}$, of the normalized gradient $\nabla G_U(\boldsymbol{U})$ evaluated at $\boldsymbol{U}^{*(i)}$. That is, $\boldsymbol{\alpha}^{(i)} = \frac{\nabla G_U(\boldsymbol{U}^{*(i)})}{\|\nabla G_U(\boldsymbol{U}^{*(i)})\|}$, where $\|\cdot\|$ denotes the l^2–norm.

Step 6:

Let $\boldsymbol{U}^{*(i)} = -\boldsymbol{\alpha}^{(i)}\beta^{r(i)}$ and solve $G_U\left(\boldsymbol{U}^{*(i)}\right) = 0$ for $\beta^{r(i)}$.

Step 7:

If $\beta^{r(i)}$ has converged, then $P_f \cong \Phi\left(-\beta^{r(i)}\right)$. If not, update $\boldsymbol{U}^{*(i+1)} = -\boldsymbol{\alpha}^{(i)}\beta^{r(i)}$, increment i (such that $i = i + 1$), and return to Step 5 until convergence.

3.2.2 Proposed Methodology

The methodology proposed in this study consists of two primary steps, which are illustrated in Figure 3-4. The first step, termed "Step 1" in Figure 3-4, is outlined in Section 3.2.2.1 and serves as the framework for quantifying reliability index threshold values, β_i^r, associated with lower limit states described by condition ratings. Section 3.2.2.2 outlines the second step, termed "Step 2" in Figure 3-4, which is the procedure for measuring the reliability index, β_m^r, of the in-service asset using long-term monitoring data and assigning its condition rating to track the structural condition of the asset. The proposed methodology applies regardless of whether or not the monitoring system is installed at the beginning of the structure's operational life (as opposed to being installed mid-life).

3.2.2.1 Step 1: Establish Reliability Index Thresholds on Condition Rating Lower Limit States

The following discussion of the methodology for Step 1 parallels the block diagram in Figure 3-4. Figure 3-4 illustrates that the inputs of the proposed process are long-term monitoring data and physical deterioration descriptions of an asset's condition ratings; both inputs correspond to a given failure limit state. The output is a set of reliability index thresholds, β_i^r, that measure the level of safety of the condition rating lower limit states. In order to create an explicit linkage between the inputs and outputs, a finite element (FE) model of the asset must be developed and calibrated using monitoring data to reflect the conditions of the asset in its in-service state. The demand effect for each condition rating, i, is then simulated by modeling the physical deterioration described by each condition rating in the FE model and subjecting the modeled asset to load distributions derived from the long-term monitoring data. The output data collected at the sensor locations of the FE model inform the random variable parameters of the demand and are used to calculate the reliability index values, β_i^r, describing each condition rating lower limit state. A detailed discussion of this process is described below.

Condition ratings are defined by descriptions of physical deterioration that can progressively lead to well-documented modes of failure. For example, net-section yielding is a threat to pin-connected hanger plates (AASHTO 2012) because corrosion can lead to reduced net-section area at the pin-plate connection. Consequently, condition ratings for such elements are described by varying levels of percent net-section loss (MDOT 2016). The first step towards establishing reliability index thresholds on condition rating lower limit states is to identify the failure limit state governing an asset's condition rating parameters and to let the physical deterioration described by each condition rating be denoted $D_G = \{D_0, D_1, \ldots, D_9\}$. Deterioration is a physical parameter that can be quantified, and hence D_G can be quantitatively modeled.

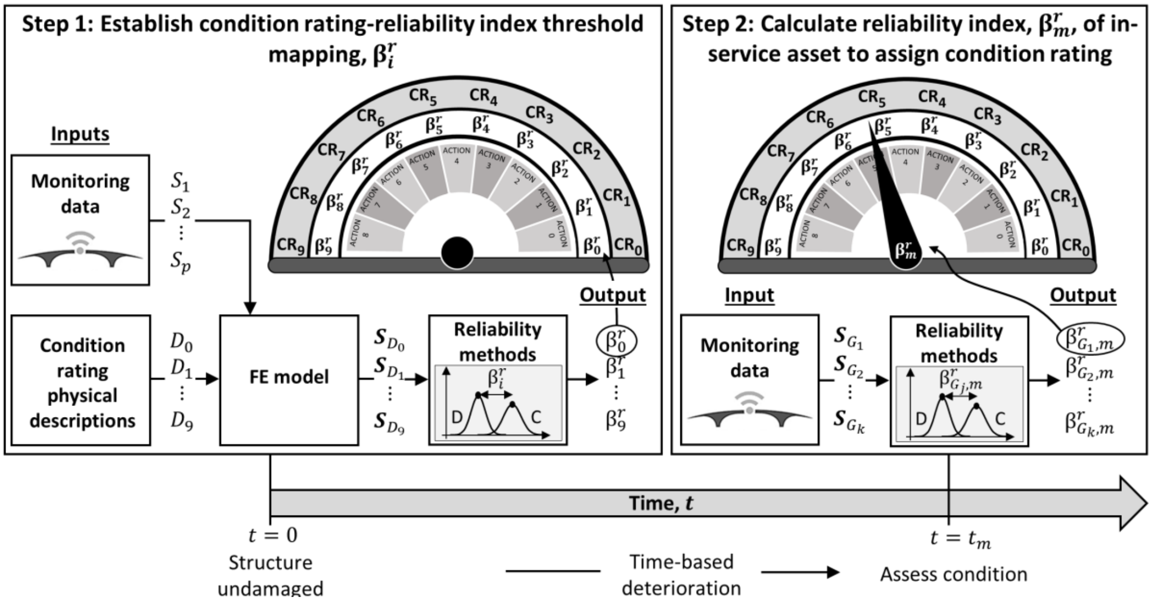


Figure 3-4. Two-step process for quantitatively linking long-term monitoring data to decisions through condition ratings. Here, t_m is the time at which β_m^r is being evaluated.

Given the quantitative physical deterioration, D_G, governed by the limit state function, $G(\boldsymbol{X}) = C - D$, a sensing program can be devised to collect long-term monitoring data capable of characterizing the limit state function, $G(\boldsymbol{X})$. Given the sensing instrumentation strategy, a subset of the n-dimensional vector of random variables, $\boldsymbol{X}$, will be random variables characterized by monitoring data collected at the sensor locations, denoted $S_1, S_2, \dots, S_p$ for p sensors. The remaining random variables in $\boldsymbol{X}$ will be those that are not characterized by long-term monitoring data, such as uncertainty in material properties and the capacity (e.g., modulus of elasticity, yield strength), which are assumed to be independent random variables. These random variables are taken from literature based on experimental test data.

The next step is to develop a FE model of the asset and to calibrate the boundary conditions (e.g., coefficient of friction at interfaces with surface-to-surface contact) using monitoring data to reflect the conditions of the asset in its in-service state. Use the long-term monitoring response data of the asset in its in-service state to infer load history inputs to the FE model. If the asset is not in its undamaged condition (i.e. the condition rating is less than 9), use nondestructive evaluation techniques to rigorously assess the condition rating of the in-service asset at the time that Step 1 is performed. This is necessary to facilitate the accurate calibration of the FE model's boundary conditions for its given physical state and condition rating. When the FE model reflects the asset in its in-service state, the input loads should result in response data at the sensor locations of the model that match the response data collected from monitoring data. The purpose of the FE model is to represent the asset at its current condition rating subject to loading histories derived from monitoring data, and then simulate physical deterioration, D_G, corresponding to each condition rating in order to quantify how the demand effect changes (as manifested by the reliability index, β_i^r). To simulate the demand effect for each condition rating, create ten FE models subjected to the same load inputs derived from long-term monitoring data, where each model reflects the physical deterioration described by one of the ten condition ratings, $D_G = \{D_0, D_1, \dots, D_9\}$. For each of the FE models, $i = 0,1,\dots,9$, characterize the output measurements at the sensor locations within the model as random variables, denoted $\boldsymbol{S}_{D_0}, \boldsymbol{S}_{D_1}, \dots, \boldsymbol{S}_{D_9}$, where $\boldsymbol{S}_{D_i}$, is an *p*-dimensional vector of random variables characterizing the outputs at the sensor locations associated with $S_1, S_2, \dots, S_p$ and corresponding to deterioration described by condition rating i.

The response data characterized by the random variable distributions for each model, $\boldsymbol{S}_{D_i}$, $i = 0,1,\dots,9$, can then be used as inputs to a FORM analysis (Algorithm 3-1) to calculate the reliability index values, β_i^r, corresponding to each model of condition rating deterioration state $D_G = \{D_0, D_1, \dots, D_9\}$. The resulting reliability index thresholds, β_i^r, are measures of the level of safety of the lower limit states corresponding to each condition rating, $i = 0,1,\ldots,9$. The reliability index values, β_i^r, serve as the bounds on each condition rating such that,

$$\mathrm{CR}_9\colon \beta_m^r \geq \beta_9^r, \mathrm{CR}_8\colon \beta_9^r > \beta_m^r \geq \beta_8^r, \ldots, \mathrm{CR}_1\colon \beta_2^r > \beta_m^r \geq \beta_1^r, \mathrm{CR}_0\colon \beta_m^r < \beta_1^r \tag{3.5}$$

If the physical deterioration description for a condition rating is predominantly qualitative and not sufficiently quantitative to inform the FE model, a shape-preserving spline can be fit to the neighboring reliability index values, β_i^r, in order to interpolate the missing threshold value not informed by quantitative descriptors (e.g. section loss). The reliability index values, β_i^r, are quantitative measures of the effect of deterioration described by each existing condition rating on the demand and an objective measure of the accepted safety threshold at which preexisting decisions are made. The temporal representation of Step 1 in Figure 3-4 illustrates that Step 1 need only be carried out a single time. This is due to the fact that the objective of Step 1 is to quantify bounds on the safety of each condition rating for an asset in terms of the reliability index, β_i^r, and serve as an objective metric and scale to continuously assess the condition of the asset into the future. The author denotes the time that Step 1 is carried out as $t = 0$.

3.2.2.2 Step 2: Calculate the Reliability Index of an In-Service Asset to Assign a Condition Rating

The following discussion of the methodology for Step 2 parallels the block diagram in Figure 3-4. Figure 3-4 illustrates that the input of the proposed process is long-term monitoring data and the output is the current reliability index of the in-service asset, $\beta_{G_j,m}^r$, corresponding to limit state function $G_j(\boldsymbol{X}_j)$ for k limit states, $j = 1,2,\ldots,k$. Reliability methods (i.e. FORM introduced in Algorithm 3-1) are used to create an explicit linkage between the inputs and outputs. A condition rating is assigned to each limit state function $G_j(\boldsymbol{X}_j)$, through the reliability index value of the in-service asset, $\beta^r_{G_j,m}$ (Figure 3-3(b)).

Section 3.2.2.1 requires implementing a monitoring program to collect long-term data capable of characterizing the limit state function governing the deterioration and failure mechanism described by an asset's set of condition ratings. In reality, deterioration typically reduces safety with respect to a number of failure limit states, not just the failure limit state and deterioration mechanisms considered by condition ratings. Consequently, it is desirable to measure the safety of an asset with respect to any failure limit state of concern in order to assess a wider domain of damage (both expected and unexpected) that may influence the performance and safety of an asset. The reliability index threshold values, β^r_i, derived in Step 1 inform the probability of failure governing each condition rating for a particular asset and reflect the level of safety that decision makers are comfortable with for each condition rating. Consequently, the measured, data-driven structural performance, β^r_m, corresponding to any measurable limit state function can be assigned a condition rating based on the defined lower limit states, β^r_i. This is justified because the reliability index, and corresponding probability of failure, is a universal measure across any failure limit state. If Step 1 is implemented when the structure first becomes operational, then any limit state function can be assessed in Step 2. This includes limit state functions governing modes of failure such as fatigue, which requires measuring cumulative parameters such as total lifetime stress cycles. If Step 2 is implemented mid-life, only limit state functions whose random variable are not characterized by data collected over the asset's entire lifespan can be accurately assessed.

After collecting long-term monitoring data, the in-service reliability index is calculated for any number of the measurable failure limit states, $G_1(\boldsymbol{X}_1), G_2(\boldsymbol{X}_2), \ldots, G_k(\boldsymbol{X}_k)$, where k limit states are monitored. Given the total array of sensors installed on the asset to measure the k limit states (where the response data collected from some sensors may be used to inform multiple failure limit states), let $\boldsymbol{S}_{G_j}$ be a vector of the random variables characterizing the response for each of the sensor locations necessary to evaluate limit state function $G_j(\boldsymbol{X}_j)$, where $j = 1,2, \ldots, k$. $\boldsymbol{S}_{G_j}$ represents all of the random variables derived from monitoring data that are used to evaluate the limit state function $G_j(\boldsymbol{X}_j)$, and is a subset of $\boldsymbol{X}_j$. Given the vector of random variables, $\boldsymbol{X}_j$, Algorithm 3-1 is used to calculate the measured reliability index values of the in-service bridge, $\beta^r_{G_{1,m}}, \beta^r_{G_{2,m}}, \ldots, \beta^r_{G_{k,m}}$, corresponding to each failure limit state, $G_1(\boldsymbol{X}_1), G_2(\boldsymbol{X}_2), \ldots, G_k(\boldsymbol{X}_k)$. Each reliability index value of the in-service asset, $\beta^r_{G_j,m}$, is then explicitly assigned a condition rating using Equation (3.5), where $\mathrm{CR}_9\colon \beta^r_{G_j,m} \geq \beta^r_9, \mathrm{CR}_8\colon \beta^r_9 > \beta^r_{G_j,m} \geq \beta^r_8, \ldots, \mathrm{CR}_1\colon \beta^r_2 > \beta^r_{G_j,m} \geq \beta^r_1, \mathrm{CR}_0\colon \beta_{G_j,m} < \beta^r_1$. The temporal representation of Step 2 in Figure 3-4 illustrates that Step 2 can be carried out and evaluated at any point in time after Step 1, as many times as desired. Since the measured reliability index, $\beta^r_{G_j,m}$, corresponding to limit state function $G_j(\boldsymbol{X}_j)$ is a metric that can be tracked using continuous, long-term data, asset managers can more regularly, if not continuously, track the condition of a structure, especially between biennial inspections.

3.3 Illustrative Example: Telegraph Road Bridge

The proposed reliability-based framework for tracking structural performance using long-term monitoring data is applied to the Telegraph Road Bridge (TRB). The TRB (Monroe,

Table 3-1. Summary of FHWA condition ratings for pin-and-hanger assemblies (MDOT 2016).

Code	Condition	Description
9	New	No deficiencies that affect long term performance.
8	Good	Protective coatings are sound and functioning but with minor weathering.
7	Good	Full section properties and functions as designed with limited deterioration.
6	Fair	Full section properties and functions as designed with minor deterioration or superficial impact damage.
5	Fair	Moderate deterioration, minor section loss, and functions as designed.
4	Poor	Considerable deterioration affecting members with up to 10% section loss. Substantial impact damage may be present.
3	Serious	Considerable deterioration affecting members with up to 25% section loss. Threat to design capacity.
2	Critical	Member will not support design loads.
1	Imminent failure	Potential for superstructure failure.
0	Failed	Failed condition.

MI) is a standard steel girder highway bridge with pin-and-hanger assemblies, included as integral structural elements in its design. The bridge has been monitored continuously since 2011. The American Association of State Highway and Transportation Officials load-and-resistance factor design (AASHTO-LRFD) of pin-connected assemblies assumes pin-and-hanger assemblies to be purely tensile elements whose design is governed by net-section yielding (AASHTO 2007). As such, the FHWA stipulates the condition ratings summarized in Table 3-1 for Michigan's pin-and-hanger assemblies, which primarily consider deterioration resulting in section loss. While pin-and-hanger assemblies are designed to be purely tensile elements (and evaluated as such using condition ratings), the safety and performance of pin-and-hanger assemblies on the bridge is a primary concern for bridge owners due to the well-documented history of bridge failures associated with corroding pin-and-hanger assemblies in the past (Fisher and Yuceoglu 1981, Juntunen 1998, NTSB 1984, South et al. 1992). These deterioration mechanisms are described in detail in Section 3.3.1 and include damage due to unexpected in-plane and out-of-plane bending, as well as damage due the expected axial deformation. Consequently, this assembly stands to benefit greatly from the proposed framework because the structural

(a) (b)

Figure 3-5. (a) An in-service pin-and-hanger assembly; (b) the inside face of a hanger plate with severe corrosion-induced section loss.

performance corresponding to the well-documented failure limit states not currently assessed by existing visual inspection-based condition ratings can be quantitatively tracked. The remainder of Section 3.3 presents an overview of the pin-and-hanger detail, TRB instrumentation plan, and long-term data collection program. The structural condition assessment of the in-service pin-and-hanger assembly on the TRB is presented and follows the methodology proposed in Section 3.2. Section 3.4 and Section 3.5 apply Step 1 and Step 2 of the methodology to the illustrative example, respectively.

3.3.1 Pin-and-Hanger Detail

Several design advantages of pin-and-hanger bridges led to an increase in the frequency of their use in bridge design in the United States from 1940 to 1980. Pin-and-hanger connections (Figure 3-5(a)) accommodate thermal expansion, reduce moments at supports, and reduce corrosion at bearing elements by locating expansion joints away from piers. However, over time, aging pin-and-hanger assemblies demonstrated a number of deficiencies that jeopardize structural safety. While the design of hanger plates assumes that such connections are purely tensile elements (AASHTO 2007), corrosion-induced changes in the assembly boundary conditions can violate this assumption by introducing stresses around the pins, leading to failure (Juntunen 1998, Nowak and Zhou 1990). For example, the Mianus River Bridge (Greenwich, CT) collapsed in 1983 due to the failure of a non-redundant pin-and-hanger assembly that had corrosion-induced lateral displacement of the hanger plate resulting in pin shear (NTSB 1984). In 1978 the deck of the St. Clair Avenue Bridge (French Village, IL) dropped by 19mm because several hanger plates fractured due to high in-plane bending stresses caused by corrosion-induced pin-plate fixity (Fisher and Yuceoglu 1981). While the design and construction of pin-and-hanger bridges was discontinued in Michigan around 1983, many in-service bridges remain operational. Specifically, of the 2,914 steel beam bridges owned by the Michigan Department of Transportation (MDOT), around 25 percent have pin-and-hanger assemblies in their design (Jansson 2008).

Deterioration of pin-and-hanger assemblies typically initiates when salt and water leak through expansion joints and corrode the pin-hanger connection. Dirt and sand behind the plate can absorb moisture, thereby accelerating corrosion, especially at the bottom pin of the plate. It can be difficult to identify the onset of corrosion induced pitting behind hanger plates at the plate-girder interface through visual inspection. For example, the back side of a hanger plate taken out of service is shown in Figure 3-5(b). It is evident that a significant amount of corrosion-induced section loss has occurred near the lower pin at the net-section area and propagates toward the gross-section area. As a result of these challenges, nondestructive ultrasonic inspection often supplements visual inspection in order to better detect cracks and section loss in hanger plates (Clark et al. 1999, Graybeal et al. 2007, Moore et al. 2004). Corrosion-induced damage typically yields three primary responses that can lead to failure: 1) partial, or complete fixity at the pin-plate interface which increases both stress at the hanger plate net-section area and shear stress in the pin, 2) displacement of the hanger plate towards the open end of the pin when coupled with transverse swaying of the span (common in skewed bridges), and 3) section loss that typically initiates at the net-section's pin-plate interface (Juntunen 1998).

3.3.2 Telegraph Road Bridge Instrumentation Plan and Data Collection Program

The TRB is a skewed (33 degrees) steel and concrete highway bridge with pin-and-hanger assemblies that was constructed in 1973. The bridge has three spans with a 20 cm concrete deck, seven girders, and carries three lanes (two lanes and one merging lane) of traffic on Michigan's northbound Interstate 275. Figure 3-6 provides information about the primary span dimensions and structural configuration. The bridge consists of three spans, where the center span is suspended between the two end spans and is entirely supported by pin-and-hanger assemblies. The bridge has two pin-and-hanger assemblies at the two ends of each of the seven center span girders (i.e. 28 hanger plates in total). Based on the documented degradation mechanisms listed in Section 3.3.1, strain gages are instrumented on a hanger plate located at the east end of the center span's Girder 2 (Figure 3-6) to monitor stresses that could be induced by damage due to long-term deterioration and corresponding changes in plate boundary conditions. Six Tokyo Sokki 120Ω uni-axial strain gages are installed on the hanger plate (Figure 3-7) in order to monitor the plate's response corresponding to limit state functions associated with axial, in-plane bending, and out-of-plane bending stresses, and torsional forces applied to the hanger plate pins (O'Connor et al. 2017). Strain gages are sampled at 100 Hz for 100 seconds every two hours using the *Narada* wireless sensor (Swartz et al. 2005) from June 2016 through May 2017. Monitoring data provides quantitative information about the loading history and

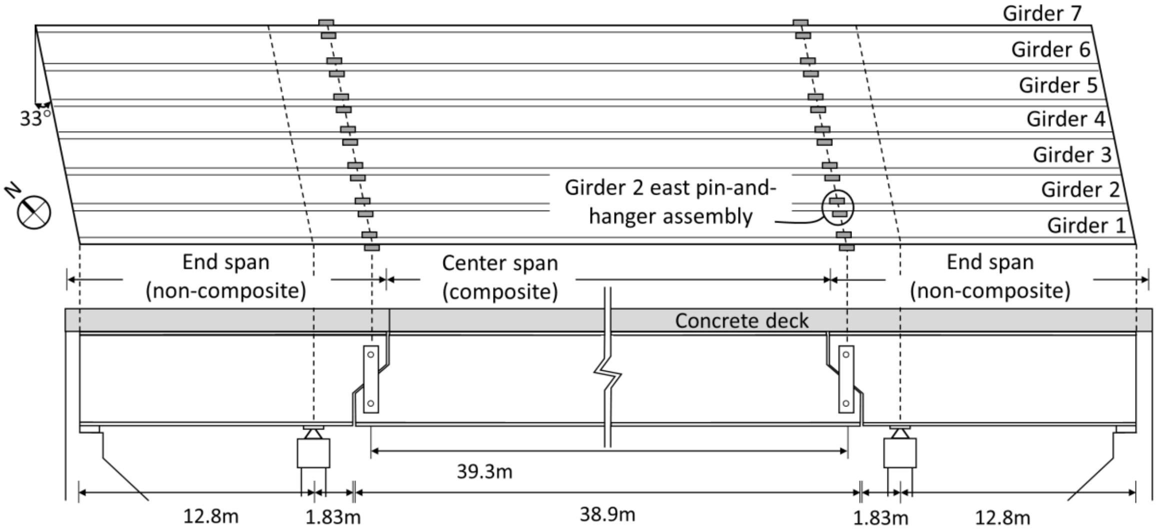
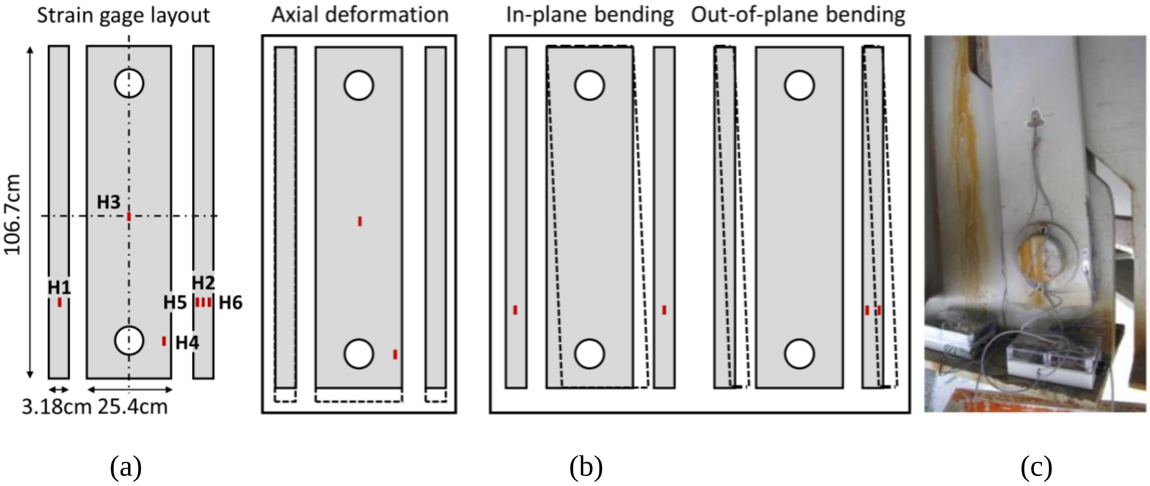


Figure 3-6. Girder elevation including center and end span dimensions.



(a) (b) (c)

Figure 3-7. (a) Strain gage layout with major dimensions; (b) expected behavior (i.e. axial deformation) and unexpected behavior (i.e. in-plane and out-of-plane bending) of the plate response; (c) strain gages installed on an instrumented TRB hanger plate.

structural response of the plate. Figure 3-8 provides a sample of the time series data collected at all six strain gage locations. The peak strain corresponds to a truck crossing the bridge and is highlighted. Some smaller responses are also observed in Figure 3-8 and likely correspond to cars that also load the bridge. Peak strain values of truck events are extracted in order to study the extreme values of long-term strain measurements using the

Block Maxima Approach (Dey and Yan 2016). Consequently, the response data collected at strain gage locations H1 through H6 are characterized by the Generalized Extreme Value (GEV) distribution, which has the following probability density function

$$f(x) = \frac{1}{\kappa} \cdot exp\left\{ -\left[1 + \xi \cdot \left(\frac{x-\eta}{\kappa}\right)\right]^{-\frac{1}{\xi}} \right\} \cdot \left[1 + \xi \cdot \left(\frac{x-\eta}{\kappa}\right)\right]^{-1-\frac{1}{\xi}} \tag{3.6}$$

for

$$1 + \xi \cdot \frac{(x-\eta)}{\kappa} > 0 \tag{3.7}$$

Here, η is denotes the location parameter, κ is the scale parameter, and $\xi \neq 0$ is the shape parameter. Strain measurements for each 100-second interval of data collection are assumed to be independently and identically distributed because it only takes a vehicle approximately 2.2 seconds to cross the bridge when driving at the posted speed limit.

3.4 Establish Reliability Index Thresholds on the TRB Pin-and-Hanger Assembly (Step 1)

This section applies Step 1 to the illustrative example of the instrumented pin-and-hanger assembly on the TRB. First, an overview of the development, calibration, and validation of an FE model of the TRB's in-service pin-and-hanger assembly is presented. Under load response histories informed by long-term monitoring data, physical deterioration specified by the condition ratings is then simulated using the FE model. Output data extracted at the sensor locations of the FE model is used in a FORM analysis (Algorithm 3-1) to calculate the reliability index value thresholds, β_i^r, associated with each condition rating. This explicitly assigns a quantifiable level of safety and performance to each decision that is currently made in practice after inspection for the TRB's pin-and-hanger assembly.

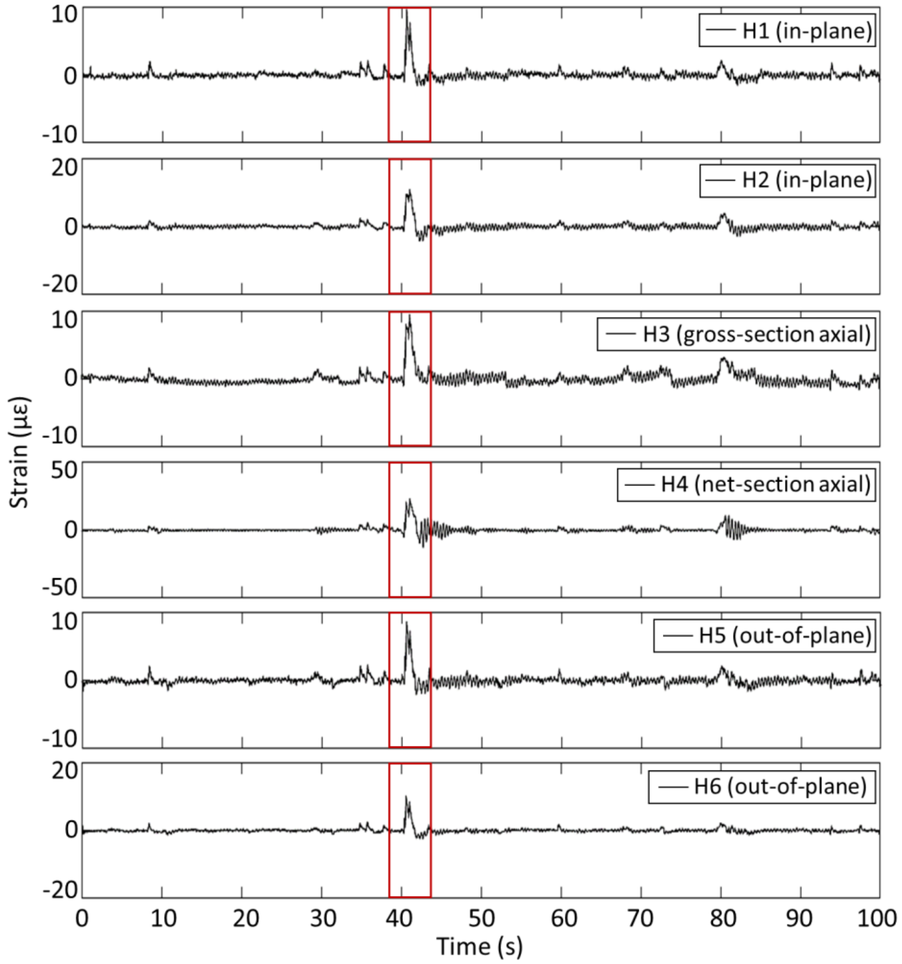
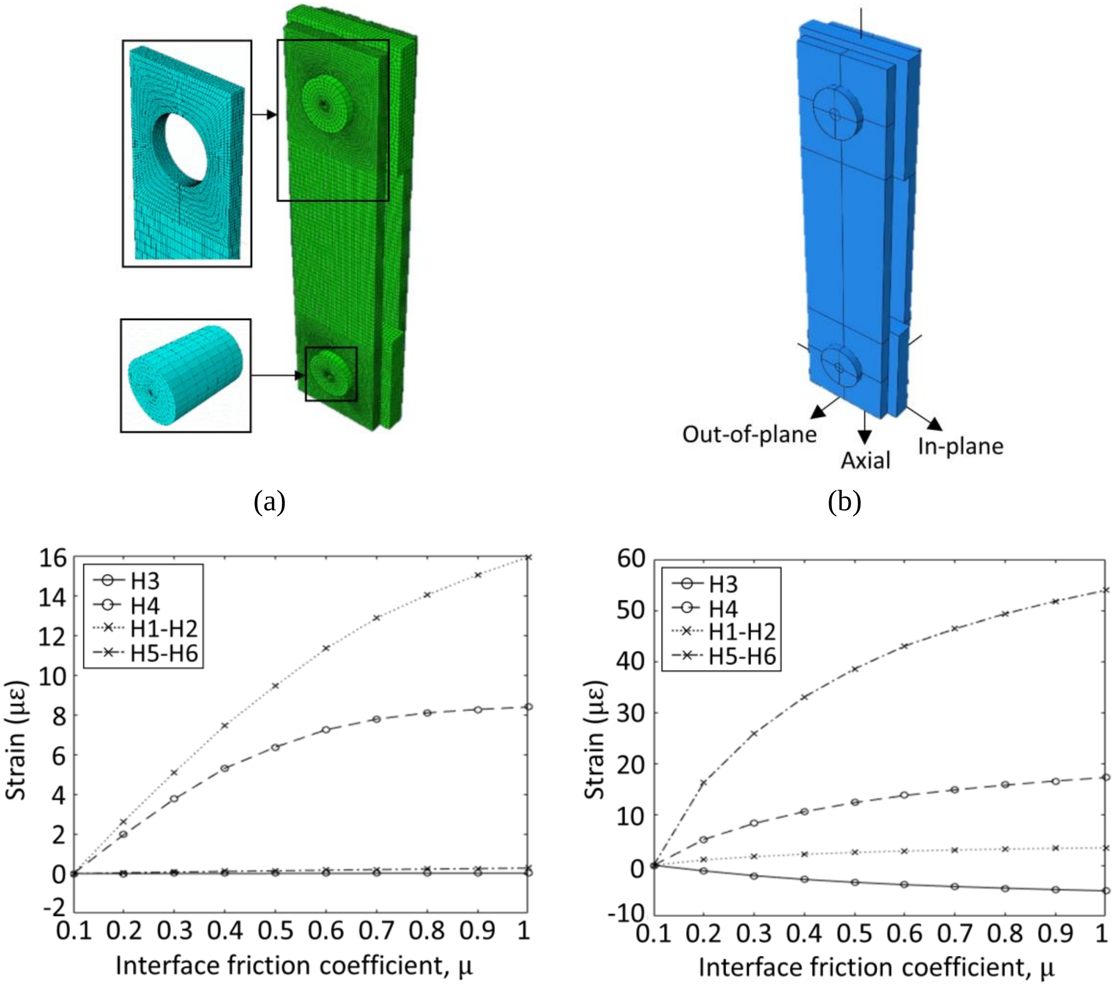


Figure 3-8. Time series data for all six strain gage locations collected during an event on May 28, 2017. The strain response produced by a passing truck is highlighted.

3.4.1 FE Model Development

A high-fidelity FE model of the in-service, undamaged instrumented pin-and-hanger assembly is developed in ABAQUS (Hibbitt 2013). The dimensions and material properties of the assembly are derived from original design drawings provided by MDOT and are used to inform the pin-and-hanger assembly's configuration, which consists of two hanger plates, two pins, and two girders. The hanger plates and girders are meshed using 8-node linear brick reduced integration elements (C3D8R) and the pins are meshed using linear wedge elements (C3D6). The entire assembly, including the plates, pins, and girders, comprises of 60,612 elements. The fully assembled and meshed model is shown in Figure 3-9(a). Since the interface at the pin-plate boundary at the upper and lower pins is integral to this study, the hanger plates and pins are partitioned to enable well-structured and symmetric meshing around this sensitive region (Figure 3-9(b)). Additionally, the partitioned regions of the hanger plate surrounding the upper and lower pins are meshed with a denser grid than the rest of the assembly because these areas are subject to contact between the plates, pins, and girders, and undergo an increased rate of change in stress under applied loads. Coulomb friction is introduced between components that undergo surface-to-surface contact. The inner surface of the plate hole and outer surface of the pin are defined as the slave and master contact surfaces, respectively, experiencing a hard-normal interaction, tangential friction interaction, and finite sliding. The FE model serves two primary purposes. First, the FE model validates the sensor instrumentation plan by confirming the influence of corrosion and unexpected bending, such as in-plane and out-of-plane bending, on the sensor locations. Second, as described in Step 1 of the methodology, the FE model plays an important role in quantifying the reliability index threshold values, β_i^r, corresponding to each condition rating's lower limit state by simulating physical deterioration.

Three criteria of the undamaged pin-and-hanger assembly are assessed as a preliminary validation of the modeled pin-plate interaction when the assembly is treated as a purely tensile element. First, the expected bridge dead load is derived from the engineering design drawings and applied to the girders in the FE model. The stress at the FE model's gross-section is confirmed in order to validate that the load transfers as expected between the modeled girders, pins, and hanger plates. Second, the stress



(a) (b)

(c) (d)

Figure 3-9. (a) Meshed FE model of pin-and-hanger assembly; (b) partitioned FE model of pin-and-hanger assembly; sensitivity of strain to interface tangential friction at the proposed sensor locations for (c) in-plane bending and (d) out-of-plane bending.

concentration at the net-section area near the pin-plate interface of the model is compared to the theoretical stress concentration. The expected stress concentration, k, is lower bounded by k = 2.16 (Young and Budynas 2002) and upper bounded by k = 2.50 (Pilkey 2005) which correspond to open-hole and tight-fitting pin-loaded stress concentration values, respectively. It is recommended to ignore stress concentrations when monitoring yielding at the net-section area of hanger plates (Jansson 2008). The FE model complies with this recommendation because sensor location H4 is located far enough away from the

Table 3-2. FE model preliminary validation results.

Validation step	Analytical result	FE result	Experimental result
Gross-section stress from dead load (MPa)	22.1	22.1	N/A
Net-section stress from dead load (MPa)	95.8 (k = 2.16) – 111 (k = 2.50)	96.5 (k = 2.18)	N/A
First Mode (Hz)	105 – 240	175.2	189

Table 3-3. Strain measurements collected at sensor locations H3 and H4 compared to FE model outputs after calibrating the pin-plate boundary conditions.

GEV distribution parameter	Axial force	In-plane prescribed displacement	H3 (data)	H3 (model)	H4 (data)	H4 (model)
Location parameter, η	11557 N	$1.39\cdot(10^{-5})$ m	3.59 με	3.59 με	6.26 με	6.27 με
Scale parameter, κ	8271 N	$7.17\cdot(10^{-5})$ m	2.56 με	2.56 με	4.80 με	4.96 με
Shape parameter, ξ	0.16	5.14	0.16	0.16	0.31	0.31

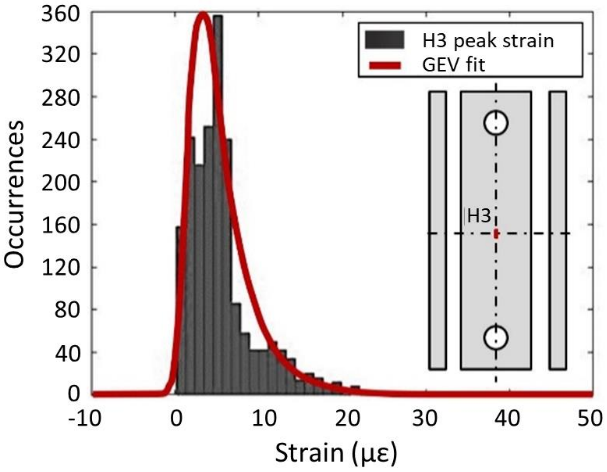
pin that the stress concentrations fully dissipate at the sensor location. Third, the first mode for out-of-plane bending of the modeled hanger plate is compared to results from experimental tests. Impulse hammer tests are conducted manually on the TRB pin-and-hanger assembly in the field during installation of the sensors. Comparing the results of the three aforementioned criteria in Table 3-2, the FE model, analytical, and experimental results are in close agreement given that the boundary condition of the hanger plate is a partially fixed support in the out-of-plane direction.

3.4.2 Sensing Instrumentation Plan Validation for Limit State Functions

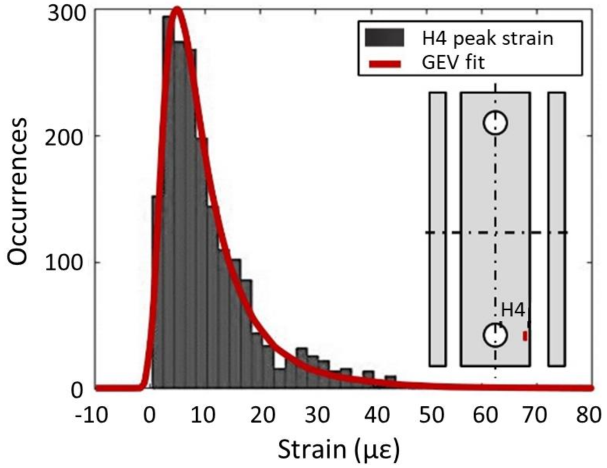
Following the preliminary FE model validation, the FE model is used to justify the sensor instrumentation plan proposed in Figure 3-7. That is, the author validates that the proposed sensing strategy can measure the sensitivity of data-driven input parameters used in this illustrative study (i.e. random variables characterized by monitoring data that inform the net-section stress and pin shear stress limit states) to corrosion-induced progressive locking modeled by tangential friction. In-plane and out-of-plane bending are simulated while the interface tangential friction between the pin and plate is varied incrementally from $\mu = 0.1$ to $\mu = 1.0$, where $\mu = 1.0$ corresponds to complete pin-plate locking. The strains extracted at the sensor locations of the FE model are centered at $\mu = 0.1$ in order to evaluate changes in relative strain with respect to free rotation. Based on the results in Figure 3-9(c)-(d), the net-section axial (H4) and in-plane (H1-H2) sensing configurations are sensitive to in-plane bending, while the net-section axial (H4) and out-of-plane (H5-H6) sensing configurations are most sensitive to out-of-plane bending under progressive corrosion-induced pin- plate locking (namely, increasing μ). As expected, the gross-section axial sensing configuration shows little sensitivity to both in-plane and out-of-plane bending. These results confirm that the sensing instrumentation plan can measure limit states that are a function of the structural response of any combination of axial deformation, in-plane bending, and out-of-plane bending on the plate, and torsional forces applied to the plate pins with responses sensitive to changes in boundary conditions that are not considered during visual inspections (e.g., interface friction).

3.4.3 FE Model Calibration to Reflect In-Service Structural Response Behavior

Long-term monitoring data is used to calibrate the corrosion-induced locking (modeled as interface tangential friction, μ) at the pin-plate interfaces to reflect the boundary conditions of the in-service, undamaged pin-and-hanger assembly. To accomplish this, monitoring data is first used to infer the live loads transferred to the in-service assembly during the one-year monitoring period. Under these loads, the interface friction at the pin-plate interfaces is adjusted until the strain response at the sensor locations in the FE model match the responses characterized by the experimental monitoring data. Since there should be little to no impact from in-plane or out-of-plane bending at sensor location H3 (i.e. the gross-section area), strain measurements from sensor location H3 are used to estimate the axial live load applied to the pins. The distribution of measured peak strain at sensor location H3 is shown in Figure 3-10. Additionally, strain measurements from sensor locations H1 and H2 indicate that there is a small amount of in-plane bending in response to each event that must be considered if the true in-service behavior of the assembly is to be modeled. Under the derived axial and in-plane loads directly modeled from the measured strain response at sensor locations H3, H1, and H2, the interface friction at the top, μ_{top}, and bottom, μ_{bot}, pins is adjusted following convex optimization such that the strain output from the FE model at sensor location H4 matches the strain measurement data collected from sensor location H4 on the in-service bridge. The distribution of measured peak strain at sensor location H4 is shown in Figure 3-10. Static friction coefficients typically range from 0.1 to 0.8, which correspond to steel-on-steel lubricated and dry contact, respectively (Engineering Toolbox 2018). Following this procedure, interface tangential friction values of $\mu_{\mathrm{top}} = 0.4$ and $\mu_{\mathrm{bot}} = 0.4$ at the top and bottom pins, respectively, result in an accurate representation of the structural response of the in-service assembly. Table 3-3 reflects these results. As expected, the response at the gross-section (i.e. sensor location H3) is not sensitive to in-plane bending. The distribution of the strain at sensor location H4 of the FE model is in close agreement with the experimental data collected from the monitoring system at sensor location H4. The small discrepancy in the distribution parameters is likely due to measurement noise and global effects on the boundary conditions that cannot be measured with the existing instrumentation plan.



(a)



(b)

Figure 3-10. Distributions of peak strain at sensor locations (a) H3 (gross-section) and (b) H4 (net-section) with superimposed GEV fit.

3.4.4 Calculate Reliability Index Threshold Values

Recall that the descriptions of physical deterioration in Table 3-1 provide quantitative information about the level of deterioration that is tolerable for each condition rating. The FHWA condition ratings for pin-and-hanger assemblies are equivalently represented in Table 3-4 in terms of the bounds on the percent net-section loss. Condition ratings 6 through 9 correspond to 0% section loss with increasing levels of superficial

Table 3-4. Updated representation of the FHWA condition ratings for pin-and-hanger assemblies to include the reliability index threshold values, β_i^r, that bound each condition rating.

Code	% Section loss	Condition	Qualitative repair action	Cost
9	$0\% +$ Superficial Damage	$\beta_{G_j}^r \geq \beta_9^r$	None	c_9
8		$\beta_9^r > \beta_{G_{j,m}}^r \geq \beta_8^r$	None	c_8
7		$\beta_8^r > \beta_{G_{j,m}}^r \geq \beta_7^r$	Minor maintenance	c_7
6		$\beta_7^r > \beta_{G_{j,m}}^r \geq \beta_6^r$	Major maintenance	c_6
5	$> 0\%, \ll 10\%$	$\beta_6^r > \beta_{G_{j,m}}^r \geq \beta_5^r$	Minor repair	c_5
4	$\gg 0\%, \leq 10\%$	$\beta_5^r > \beta_{G_{j,m}}^r \geq \beta_4^r$	Major repair	c_4
3	$> 10\%, \leq 25\%$	$\beta_4^r > \beta_{G_{j,m}}^r \geq \beta_3^r$	Rehabilitate	c_3
2	$> 25\%, \leq 85\%$	$\beta_3^r > \beta_{G_{j,m}}^r \geq \beta_2^r$	Replace	c_2
1	$\geq 85\%$	$\beta_2^r > \beta_{G_{j,m}}^r \geq \beta_1^r$	Close bridge and evacuate	c_1
0		$\beta_{G_{j,m}}^r < \beta_1^r$	Beyond corrective action	c_0

deterioration. As the section loss increases, the condition rating decreases and the reliability index values corresponding to the net-section yielding limit state decrease. These quantitative descriptions enable the quantification of the reliability index thresholds, β_i^r, by providing sufficient information to model each condition rating and simulate the effect of deterioration on the demand using the FE model. Output data extracted at the sensor locations of the FE model for each condition rating's physical deterioration is used in a FORM analysis (Algorithm 3-1) to calculate the reliability index threshold values, β_i^r, associated with each condition rating with respect to the net-section yielding limit state. The objective of this subsection is to calculate the reliability index threshold values, β_i^r, which serve as a metric to continuously assess the condition of the asset with respect to a measured limit state function $G_j(\boldsymbol{X}_j)$ and corresponding reliability index, $\beta_{G_{j,m}}^r$. This is reflected in Table 3-4, where each existing condition rating is bounded by reliability index threshold values, β_i^r, and corresponds to a repair action (Estes and Frangopol 2001) to ensure safety (at a cost, c_i).

The assumed net-section stress limit state function governing the pin-and-hanger assembly's condition ratings is denoted $G_1(\boldsymbol{X}_1)$, where

$$G_1(\boldsymbol{X}_1) = F_y - \frac{DL}{A_n} - \varepsilon_{H4} \cdot E \qquad (3.8)$$

Here, the distribution parameters for yield stress, F_y, and elastic modulus, E, are taken from literature (Hess et al. 2002). Under the assumption that the deck and parapet weights are evenly distributed across the seven girders, the dead load, DL, is estimated from engineering design drawings and the coefficient of variation is taken from literature (Nowak 1993). The statistical parameters for the net-section area, A_n, are extracted from measured data. Since corrosion-induced pitting and section loss typically initiate at the lower pin and move upward (for example, Figure 3-5(b)), it can be safely assumed that the gross-section area at the middle of the plate maintains its full section properties. Since the gross-section area, A_g = 80.6 cm^2, remains constant, a ratiometric measure of strains from sensor locations H3 and H4 is used to characterize the net-section area, A_n, as

$$A_n = A_g \cdot \frac{\varepsilon_{H3}}{\varepsilon_{H4}} \qquad (3.9)$$

The random variable distribution parameters for the yield stress, F_y, elastic modulus, E, dead load, DL, and net-section area, A_n, are provided in Table 3-5.

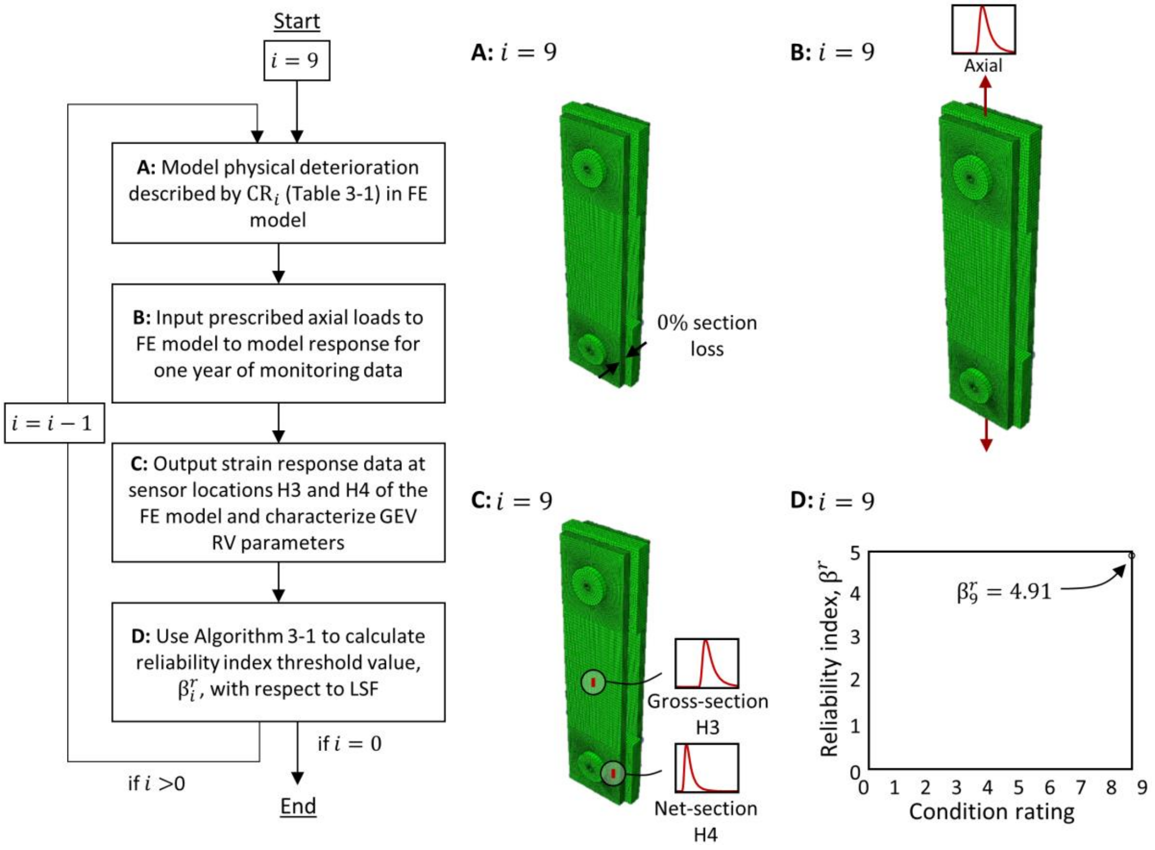


Figure 3-11. Process for modeling condition rating physical deterioration descriptions and calculating the corresponding reliability index threshold value, β_i^r.

In order to calculate the reliability index threshold values, β_i^r, using Equation (3.8), the output distributions of peak strain at sensor locations H3 and H4 of the FE model (i.e. ε_{H3} and ε_{H4}) need to be characterized for each condition rating's level of physical deterioration. Following the procedure outlined in Figure 3-11, first, model the physical deterioration described by CR_i (Table 3-1) in the FE model. Under these conditions, apply the axial loads derived in Section 3.4.3 (Table 3-3). Next, characterize the GEV distribution random variable parameters of the output strain response data at sensor locations H3 and H4 in the FE model. At this point, all random variable parameters necessary to inform $\boldsymbol{X}_1$ and evaluate the net-section stress limit state function (Equation (3.8)) are defined. Use Algorithm 3-1 to calculate the reliability index threshold value, β_i^r,

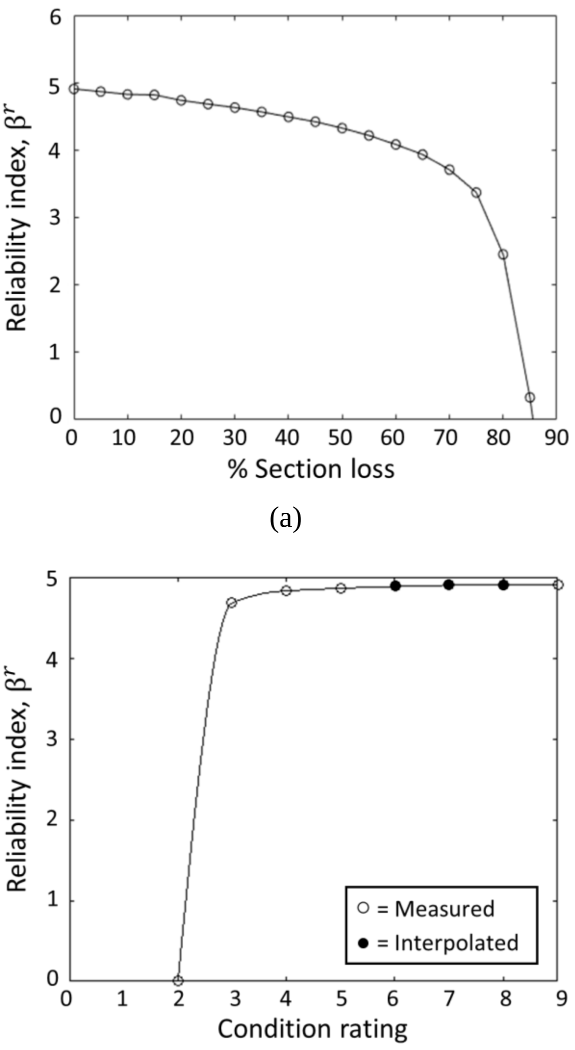
Table 3-5. Random variable distribution parameters used in the FORM analysis for Step 1.

Random variable	Mean	COV	Distribution type	Source
Yield stress, F_y (MPa)	342	0.0890	Lognormal	(Hess et al. 2002)
Shear yield stress, F_v (MPa)	205	0.0890	Lognormal	(Hess et al. 2002)
Elastic modulus, E (GPa)	205	0.0179	Lognormal	(Hess et al. 2002)
Dead load, DL (kg)	$19.5 \cdot (10^3)$	0.100	Normal	Monitoring data, (Nowak 1993)
Net-section area, A_n (cm^2)	41.2	0.0455	Normal	Monitoring data

Random variable	Location parameter, η	Scale parameter, ξ	Shape parameter, κ	Distribution type	Source
Gross-section strain, ε_{H3}	3.59 με	2.56 με	0.16	GEV	Monitoring data
Net-section strain, ε_{H4}	6.26 με	4.80 με	0.31	GEV	Monitoring data
In-plane-bending strain, $\varepsilon_{\mathrm{H1}} - \varepsilon_{\mathrm{H2}}$	1.90 με	1.51 με	0.15	GEV	Monitoring data

with respect to the net-section stress limit state function in Equation (3.8). Repeat this procedure for all condition ratings, $i = 0,1,\ldots,9$. If the physical deterioration description for a condition rating is predominantly qualitative and not sufficiently quantitative to inform the FE model (e.g., condition ratings 6, 7, and 8 for pin-and-hanger assemblies), a shape-preserving piecewise cubic interpolation of the values at neighboring known values can be used to interpolate the missing threshold values.

The results of this procedure implemented on the instrumented TRB pin-and-hanger assembly are illustrated in Figure 3-12. Figure 3-12(a) plots the reliability index threshold values, β_i^r, corresponding to twenty analyses in which section loss at the net-section area of the FE model is increased incrementally from 0% to 95%. Figure 3-12(b) provides a plot of the reliability index threshold values, β_i^r, corresponding to each of the ten condition ratings ($i = 0,1,\ldots,9$) for pin-and-hanger assemblies. These values, β_i^r, in Figure 3-12(b) are used to populate the threshold values of β_0^r through β_9^r in Table 3-4.



(a)

(b)

Figure 3-12. Reliability index values for (a) varying section loss, and (b) condition ratings.

The result in Figure 3-12(b) provides an explicit measure of the quantitative level of safety and performance that bridge decision makers (e.g., MDOT) currently use as thresholds for condition ratings to make asset management decisions for pin-and-hanger assemblies. This process also provides decision makers with a quantitative assessment metric to objectively reevaluate and improve the criteria that currently define condition ratings.

3.5 Assessing the In-Service TRB Pin-and-Hanger Assembly Reliability (Step 2)

Given the lower limit states established in Section 3.4, data collected from the in-service assembly can now be used to monitor any limit state and assign a condition rating to each limit state at any point in time. To illustrate this concept, consider Figure 3-13, where D represents the domain of all types and magnitudes of damage that are deteriorating a pin-and-hanger assembly. These might include, for example, damage due to in-plane and out-of-plane stresses that result from changes in boundary conditions, corrosion-induced locking at the pin-plate interface, and section loss at the net-section area. Different limit state functions, denoted as $G_1(\boldsymbol{X}_1)$ through $G_k(\boldsymbol{X}_k)$, can be impacted by any or all types of deterioration, and result in respective condition ratings, $\beta^r_{G_1,m}$ through $\beta^r_{G_k,m}$. The current safety and performance of a limit state is assigned to an existing condition rating based on its reliability index using Table 3-4, where the values β^r_0 through β^r_9 are shown in Figure 3-12(b). The following two subsections demonstrate the data-driven mapping of two distinct limit states of the TRB's pin-and-hanger assembly into the quantitative condition ratings identified in Section 3.4.

3.5.1 Monitoring Net-Section Stress of the In-Service Assembly

Consider the net-section stress limit state function, $G_1(\boldsymbol{X}_1)$, presented in Equation (3.8). As previously discussed in Section 3.4.2, the distribution parameters of the random variables for yield stress, F_y, elastic modulus, E, dead load, DL, and net-section area, A_n, are provided in Table 3-5. The net-section peak strain, ε_{H3}, and gross-section peak strain, ε_{H4}, GEV distribution parameters based on the monitoring data collected on the TRB during the one-year monitoring period are also provided in Table 3-5. Implementing

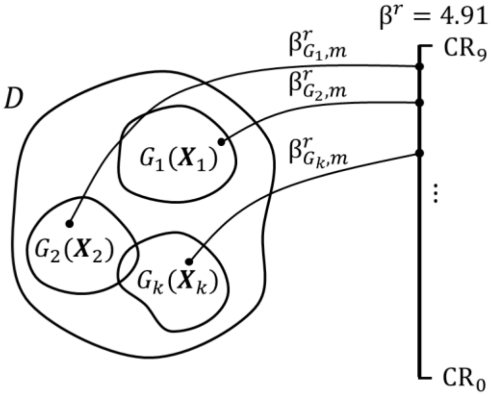


Figure 3-13. Domain of entire damage profile is quantified by a single set of quantified condition ratings.

FORM (Algorithm 3-1) with respect to the net-section stress limit state function, $G_1(\boldsymbol{X}_1)$, for the instrumented hanger, the reliability index, $\beta^r_{G_1,m}$, equals to 4.91. As summarized in Table 3-6, this result indicates that the probability of failure for the measured limit state of the in-service assembly is equivalent to the probability of failure corresponding to condition rating 9 (Figure 3-12(b)), and no repair action is necessary. In order to verify that FORM is sufficient, and that second-order reliability methods are not necessary to increase the accuracy of the approximation, the reliability index is also calculated using Monte Carlo methods. The author concludes that it is justified to use FORM to estimate the probability of failure for this limit state function because the reliability index generated from the Monte Carlo simulation with $5(10)^8$ iterations is 4.90, which is within 0.2% of the FORM results.

In order to highlight the importance of making decisions based on quantitative data rather than visual or ultrasonic inspection, once again consider net-section yielding as the failure mechanism. The long-term monitoring system described in this chapter calculates condition ratings based on the net-section stress limit state function, $G_1(\boldsymbol{X}_1)$. It is important

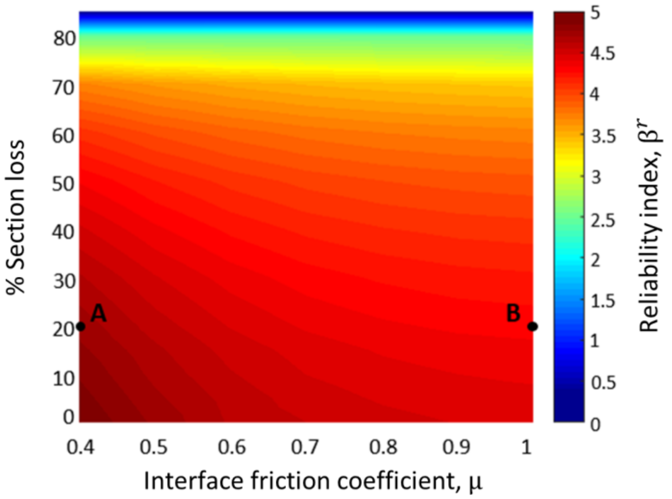


Figure 3-14. Reliability index values for varying percent section loss and interface friction coefficient for the net-section stress limit state.

to note that due to the use of a sensing strategy that captures the targeted deterioration, use of data alone is sufficient to estimate $\beta^{r}_{G_1,m}$. Unlike in Step 1, Step 2 does not require assumptions on section loss nor pin-plate locking because the impact of those deterioration states are implicit to the measurements ε_{H3} and ε_{H4}. Visual and ultrasonic inspection might capture of the deterioration (such as section loss), it may fail to quantitatively assess other deterioration types (e.g., corrosion-induced locking at the pin-plate interface) and their impact on the safety and performance of the pin-and-hanger assembly.

To illustrate this distinction, Figure 3-14 shows the results of seventy-seven FE model simulations following the process in Figure 3-11 for which both the percent section loss at the net-section *and* the interface friction coefficient of the lower pin-plate boundary are adjusted, where $\mu = 0.4$ is the interface friction coefficient of the undamaged pin-and-hanger assembly. Consider the case where the hanger plate has 20% section loss and corrosion is modeled as $\mu = 1.0$, which corresponds to complete pin-plate locking. Based on the results illustrated in Figure 3-14, the influence of both section loss and interface

Table 3-6. Structural condition monitoring of in-service assembly with respect to net-section stress and pin shear stress limit states.

Measured Limit State	FORM Result	Condition Rating	Repair Action
Net-section stress, $G_1(\boldsymbol{X}_1)$	4.91	9	No repair action necessary
Pin shear stress, $G_2(\boldsymbol{X}_2)$	6.80	9	No repair action necessary

friction are considered and the assembly has a reliability index of 4.30 ($P_f = 8.54(10)^{-6}$) with respect to the net-section yielding limit state (Location B in Figure 3-14). This reliability index corresponds to a condition rating of 2 based on Figure 3-12(b). On the other hand, using traditional visual and ultrasonic inspection methods that, as specified by condition ratings, only consider section loss and cannot quantify locking, the same hanger plate with 20% section loss has a reliability index of 4.73 ($P_f = 1.12(10)^{-6}$) and would be assigned a condition rating of 3 (Location A in Figure 3-14). In addition to resulting in different condition ratings, the results derived through visual and ultrasonic inspection lead decision makers to believe that the probability of failure is 7.6 times lower than it actually is when corrosion-induced locking is considered. Consequently, this case study illustrates that data-driven inspection using monitoring data captures the effects of damage that cannot be perceived using visual inspection when assessing structural performance and assigning condition ratings.

3.5.2 Monitoring Pin Shear Stress of the In-Service Assembly

When corrosion at the pin-plate interface initiates locking, the dead and live loads transferred from the suspended middle span to the pin result into two shear planes. Full or partial fixity between the pin and plate introduces torsional stress, τ, where $\tau = Tx/J$. Here, T is the applied torque, c is the outer radius of the pin, $J = \frac{\pi}{2} \cdot c^4$ is the polar moment of inertia, and x is the radial location. The maximum shear, τ_{max}, at the pin neutral axis due to the parabolically distributed shear stress resulting from axial load in the plate is $\tau_{max} = \frac{4}{3} \cdot \tau_{avg}$, where τ_{avg} is the average shear stress. A second limit state function, $G_2(\boldsymbol{X}_2)$, is formulated for failure due to corrosion-induced pin shear stress. This limit state cannot be assessed using traditional visual inspection-based evaluation methods and is not considered by existing condition ratings. The assumed limit state function $G_2(\boldsymbol{X}_2)$ is

$$G_2(\boldsymbol{X}_2) = 0.6 \cdot F_y - \frac{DL+LL}{A_{pin}} \cdot \frac{4}{3} - \frac{T \cdot c}{J} = 0.6 \cdot F_y - \frac{DL+\varepsilon_{H3} \cdot E \cdot A_g}{A_{pin}} \cdot \frac{4}{3} - \frac{(\varepsilon_{H1}-\varepsilon_{H2}) \cdot E \cdot b \cdot h^2 \cdot c}{12 \cdot J} \tag{3.10}$$

Here, the monitoring data collected at strain gage location H3 is used to estimate the live load, LL, carried by the pin-and-hanger assembly, and the strain gages at locations H1 and H2 are used to determine the moment in the hanger near the pin connection, M, where $T = M$. Since ε_{H1}, ε_{H2}, and ε_{H3} are dependent non-normal random variables, the Nataf Transformation (Nikolaidis et al. 2004) presented in Case 1.2 of Algorithm 3-1 is used to uncorrelated the random variables. The shear yield stress, $F_v = 0.6 \cdot F_y$, and has the same coefficient of variation as F_y (Hess et al. 2002). The distribution parameters necessary to define the random variables in $\boldsymbol{X}_2$ are provided in Table 3-5. The area of the exposed pin face, A_{pin} = 127 cm^2, gross-section area, A_g = 80.6 cm^2, thickness of the hanger plate, b = 3.18 cm, width of the hanger plate, h = 25.4 cm and pin radius, c = 6.35 cm are taken to be deterministic. Implementing FORM (Algorithm 3-1) with respect to the pin shear stress limit state function, $G_2(\boldsymbol{X}_2)$, for the instrumented hanger, the reliability index, $\beta^r_{G_2,m}$, equals 6.80. As summarized in Table 3-6, this result indicates that the probability of failure for the measured limit state of the in-service assembly is equivalent to the probability of failure corresponding to condition rating 9 (Figure 3-12(b)), and no repair action is necessary.

3.6 Conclusions

This chapter presents a method to explicitly link long-term monitoring data to condition ratings through a reliability-based framework. In order to link structural monitoring data with existing condition ratings using reliability methods, a method is offered to quantify the reliability index values, β_i^r, corresponding to the lower limit states described by condition ratings. This explicitly assigns a quantifiable level of safety and performance to each decision that is currently made in practice after inspection. Once the reliability index threshold values, β_i^r, are established for a set of condition ratings, the data-driven reliability index, β_m^r, of the in-service asset (measured from monitoring data) can be monitored continuously and explicitly mapped to a condition rating to track the structural condition of the asset at any point in time and to trigger upkeep decisions based on condition ratings. The results in Figure 3-12 demonstrate that a human-independent, SHM-based operation can successfully quantify existing condition ratings and facilitate monitoring-based decision making using quantitative condition ratings. This method creates a truly data-driven decision-making strategy that is synergistic with the existing state of practice, eliminates risks associated with periodic (rather than continuous) inspections, and expands condition ratings to encompass the entire measurable domain of damage that may exist in a system or component.

The proposed framework is successfully implemented with long-term monitoring data collected from the TRB, which is a standard steel girder highway bridge with pin-and-hanger assemblies. The application of the proposed framework to the TRB pin-and-hanger assembly demonstrates that long-term monitoring is necessary to assess the influence of diverse types of damage on the safety of a structural assembly. For hanger plates, these damage cases are historically tied to corrosion-induced locking at the pin-plate interface, and in-plane and out-of-plane bending. To illustrate the importance of this finding, the author compares the reliability index value and corresponding condition rating derived from the proposed methodology to the conclusions drawn from existing visual and ultrasonic-based methods of inspection when the net-section stress limit state is considered with both section loss and locking at the pin-plate interface. The results indicate that current inspection methods can fail to quantify and assess the impact that underlying damage and changes in boundary conditions can have on the safety and performance of the pin-and-hanger assembly. As a result, deterioration that increases the probability of failure of a pin-and-hanger assembly can go unnoticed during condition rating inspections.

Moving forward, the author will tie the proposed data-driven condition ratings more concretely to owners' decisions by quantifying the cost implications of actions that are associated with each condition rating (i.e. c_i in Table 3-4). This will extend the existing reliability-based framework to a more general risk-based decision-making framework where each element has to be assessed for risk, where risk is based on both the probability of failure tied to condition ratings, as defined in this chapter, and consequences (i.e. cost of action or inaction). Such an extension would enable stakeholders to concretely understand which asset management decisions are considered financially optimal given an asset's current condition rating. While decoupled from condition ratings, a preliminary theoretical basis for such a decision-making process has been investigated by Cappello et al. (Cappello et al. 2016) and Thöns (Thöns 2018). The proposed method investigated in this chapter also sets the foundation for implementing better decision-making strategies based on risk—and even setting new, quantitatively informed criteria for condition ratings themselves—without needing to change the already established, and nationally mandated bridge management system that is founded on condition ratings. Future work will also include exploring the potential difficulties of extending the proposed methodology to other bridge components that require less straightforward analytical modeling.

3.7 References

Agrawal A, Kawaguchi A and Qian G. Bridge element deterioration rates: phase I report. Report, New York State Department of Transportation Project# C-01-51, Albany, 2009.

American Association of State Highway and Transportation Officials (AASHTO). *LRFD bridge design specifications*. 4th ed. Washington: AASHTO, DC, 2007.

Cappello C, Zonta D and Glisic B. Expected utility theory for monitoring-based decision making. *Proc IEEE* 2016; 104(8): 1647-1661.

Catbas FN, Susoy M and Frangopol DM. Structural health monitoring and reliability estimation: long span truss bridge application with environmental monitoring data. *Eng Struct* 2008; 30(9): 2347-2359.

Cawley P and Adams RD. The location of defects in structures from measurements of natural frequencies. *J Strain Anal Eng Des* 1979; 14(2): 49-57.

Clark AV, Hehman CS, Gallagher D, et al. Ultrasonic measurement of stress in pin-and-hanger connections. *J Nondestr Eval* 1999; 18(3): 103-113.

Dey DK and Yan J. *Extreme value modeling and risk analysis: methods and applications*. Boca Raton: CRC Press, 2016.

Doebling SW, Farrar C and Prime MB. A summary review of vibration-based damage identification methods. *Shock Vib Dig* 1998; 30(2): 91-105.

Elishakoff I. *Probabilistic methods in the theory of structures*. 1st ed. Hoboken: Wiley-Interscience, 1983.

The Engineering Toolbox. Friction and friction coefficients, https://www.engineeringtoolbox.com/friction-coefficients-d_778.html (2004, accessed 11 August 2018).

Estes AC and Frangopol DM. Repair optimization of highway bridges using system reliability approach. *J Struct Eng* 1991; 125(7): 766-775.

Estes AC and Frangopol DM. Bridge lifetime system reliability under multiple limit states. *J Bridge Eng* 2001; 6(6): 523-528.

Ettouney M and Alampalli S. *Risk management in civil infrastructure*. Boca Raton: CRC Press, 2016.

Fisher JW and Yuceoglu U. A survey of localized cracking in steel bridges. Lehigh University Fritz Laboratory Report No. 448-2(81), 1981.

Frangopol DM and Estes AC. Lifetime bridge maintenance strategies based on system reliability. *Struct Eng Int* 1997; 7(3): 193-198.

Frangopol DM, Kong JS and Gharaibeh ES. Reliability-based life-cycle management of highway bridges. *J Comput Civ Eng* 2001; 15(1): 27-34.

Frangopol DM, Strauss A and Kim S. Bridge reliability assessment based on monitoring. *J Bridge Eng* 2008; 13(3): 258:270.

Graybeal BA, Walther RA and Washer GA. Ultrasonic inspection of bridge hanger pins. *Transp Res Rec J Transp Res Board* 2007; 1697(1): 19-23.

Hess PE, Bruchman D, Assakkaf IA, et al. Uncertainties in material and geometric strength and load variables. *Nav Eng J* 2002; 114: 139-166.

Hibbitt K. *ABAQUS: user's manual: version 6.13*. Providence: Dessault Systèmes, 2013.

Jansson PO. Analysis of stress distribution in link plates used for suspending bridge beams. Michigan Department of Transportation Report No. R-1517, 2008.

Juntunen DA. Study of Michigan's link plate and pin assemblies. Michigan Department of Transportation Report No. R-1358, 1998.

Kim JT, Ryu YS, Cho HM, et al. Damage identification in beam-type structures: frequency-based method vs mode-shape-based method. *Eng Struct* 2003; 25(1): 57-67.

Michigan Department of Transportation (MDOT). *Bridge safety inspection NBI rating guidelines*. MDOT, 2016.

Moore M, Phares BM and Washer GA. Guidelines for ultrasonic inspection of hanger pins. Federal Highway Administration Report No. FHWA-HRT-04-042, 2004.

National Transportation Safety Board (NTSB). Highway accident report: collapse of a suspended span of interstate Route 95 highway bridge over the Mianus River. NTSB Bureau of Accident Investigation Report No. NTSB/HAR-84-03, 1984.

Nikolaidis E, Ghiocel DM and Singhal S. *Engineering design reliability*. 1st ed. Boca Raton: CRC Press, 2004.

Nowak AS. Live load model for highway bridges. *Struct Saf* 1993; 13(1-2): 53-66.

Nowak AS and Zhou J. System reliability models for bridges. *Struct Saf* 1990; 7(2-4): 247-254.

O'Connor SM, Zhang Y., Lynch JP, et al. Long-term performance assessment of the Telegraph Road Bridge using a permanent wireless monitoring system and automated statistical process control analytics. *Struct Infrastruct Eng* 2017; 13(5): 604-624.

Peeters B and De Roeck G. One-year monitoring of the Z24-bridge: environmental influences versus damage events. *Earthquake Eng Struct Dyn* 2001; 30(2): 149-171.

Pilkey WD. *Formulas for stress, strain, and structural matrices*. 2nd ed. Hoboken: John Wiley & Sons, 2005.

Rausand M. *System reliability theory: models, statistical methods, and applications*. 2nd Ed. Hoboken: John Wiley & Sons, 2004.

Seo J, Hu JW and Lee J. Summary review of structural health monitoring applications for highway bridges. *J Perform Constr Facil* 2016; 30(4): 04015072.

South JM, Hahin C and Telford RO. Analysis, inspection, and repair methods for pin connections on Illinois bridges. United States Department of Transportation Report No. FHWA/IL/PR-01,1992.

Swartz RQ, Jung D, Lynch JP, et al. Design of a wireless sensor for scalable distributed in-network computation in a structural health monitoring system. In: *Proc of the 5th international workshop on structural health monitoring* (eds F Chang and F Kopsaftopoulos), Menlo Park, CA, 12 September – 14 September 2005, pp. 1570-1577. Lancaster: DEStech Publications, Inc.

Tabsh SW and Nowak AS. Reliability of highway girder bridges. *J Struct Eng* 1991; 117(8): 2372-2388.

Thöns S. On the value of monitoring information for the structural integrity and risk management. *Comput-Aided Civ Infrastruct Eng* 2018; 33: 79-94.

Whelan M and Janoyan K. In-service diagnostics of a highway bridge from a progressive damage case study. *J Bridge Eng* 2010; 15(5): 597-607.

Young WC and Budynas RG. *Roark's formulas for stress and strain*. 7th ed. New York: McGraw-Hill, 2002.

CHAPTER 4

Optimal Stochastic Data Collection and Transmission Policy for Remote Parameter Estimation in Wireless Sensing Architectures Under Resource and Hardware Constraints

4.1 Introduction

The proliferation of low-cost and miniaturized, yet high-performing, sensors has enabled sensing in the natural and built environments across a broad range of applications including in civil and transportation systems, health care (e.g., wearable devices), surveillance, industrial control, and environmental systems, to name a few. Within these domains, sensing and computation have been aided by advancements in wireless communication that have lowered costs, improved communication latency, permitted varying protocols, and enabled cloud computing in wireless sensor networks (WSNs). Despite these advances, the availability of energy remains a bottleneck in WSNs that do not have access to fixed power to enable continuous and reliable operation; such WSNs often rely on energy harvesting (EH) from the environment. Consideration of harvested energy as a limited, and uncertain, resource makes the optimal operation of WSNs challenging. This is especially pronounced in WSN architectures that are used to collect and transmit measured data to a remote estimator. Since only a subset of the measured information can be wirelessly transmitted to the estimator for processing, there is an inherent tradeoff between the quality of the parameter estimates and the energy-constrained communication. Such sensing architectures should transmit high-value data that increases the quality of the parameter estimates at the expense of discarding low-value data. This is in contrast to deterministic data collection methods such as schedule-based and "transmit-all" data collection methods. In the context of this chapter, a deterministic schedule-based policy refers to a WSN architecture that initiates and terminates data collection periodically on a pre-defined and regular schedule, and a transmit-all policy refers to the attempted transmission of all measurements.

The problems of 1) energy-aware data transmission in WSNs under resource constraints and 2) controlled parameter estimation over WSNs, have gained considerable traction in the wireless communication and signal processing communities, respectively. Communication schemes for WSNs with wide-ranging energy recharging models and objectives have emerged in recent years to augment advances in the efficiency of energy harvesting hardware circuits and design. These communication schemes improve upon schedule-based and transmit-all strategies by considering the availability of energy and differentiating the importance of measured data. Seyedi and Sikdar (Seyedi and Sikdar 2008) proposes a unified Markov-based model for energy harvesting nodes in WSNs that determines the state of the node based on the remaining energy supply and the state of the energy harvesting process. In Uysal-Biyikoglu et al. (Uysal-Biyikoglu et al. 2002) the transmission of data packets is scheduled to minimize energy depletion subject to a delay constraint or deadline. Other works explicitly account for the reward or value corresponding to the transmission of data where the primary objective is to maximize rewards. For example, Michelusi et al. (Michelusi et al. 2012) derives a transmission policy characterized by a threshold vector that maximizes the average reward per slot for a slotted-time system with independent and identically distributed (i.i.d.) Bernoulli energy arrivals under the condition that data arrivals are deterministic, the device can measure the exact energy level in its battery, and each data packet has a random importance value. Lei et al. (Lei and Yates 2009) also introduces a transmission policy that maximizes the expected reward rate of data transmitted from a replenishable sensing architecture, but instead considers a continuous-time system. However, the optimal policy derived in Lei et al. (Lei and Yates 2009) assumes that each collected data packet is transmitted immediately—meaning there is no storage buffer utilized—and implementation requires prefect knowledge of the state of charge (SoC). Tang and Tan (Tang and Tan 2017) expands upon the work of Lei et al. (Lei and Yates 2009) by using a multi-layer Markov fluid queue model to account for the rejection of any data that arrives when a sensing node's energy buffer is fully depleted. Ozel et al. (Ozel at al. 2015) and Ulukus et al. (Ulukus et al. 2015) provide detailed reviews of recent advances in EH WSNs. None of these representative works consider the quality of parameter estimates that are generated remotely from data collect according to their proposed scheduling policies.

Considerable work has also been carried out to develop controlled parameter estimation policies for WSNs. One class of problems that is considered widely for optimal remote estimation in WSNs is that of event-based scheduling. In event-based scheduling, a threshold or transmission criterion is derived that optimizes the tradeoff between communication rate and quality of the estimate (such as reducing the variance of parameter estimates). Event-based policies control whether data is transmitted or discarded after the event occurs and the importance or value of the data is known. Wu et al. (Wu et al. 2013) and You et al. (You et al. 2013) each propose an event-based scheduler for linear dynamical systems in which an appropriate event-triggering threshold is used to ensure quality estimation under communication rate constraints; while Wu et al. (Wu et al. 2013) only assesses the estimator's stability under the proposed schedule, You et al. (You et al. 2013) evaluates its performance. Han et al. (Han et al. 2015) presents an event-based stochastic scheduling policy that reduces the transmission rate while preserving acceptable estimation accuracy when compared to the standard maximum likelihood estimator (MLE) with full measurements. Shi et al. (Shi et al. 2011) proposes an event-based sensor scheduler that minimizes the average estimation error given a deterministic constraint governing the number of times the sensing architecture can communicate with the remote estimator over finite time horizon. In these works, objective constraints (e.g., transmission rate) are assumed *a priori*, as opposed to being derived from stochastic energy harvesting and recharging models that account for a WSN's primary resource and hardware constraints. Consequently, while the areas of optimal energy-aware data transmission in WSNs and controlled remote parameter estimation over WSNs are closely related, they remain decoupled in the existing literature.

This chapter reconciles these two problems and presents the derivation and implementation of an optimal data collection and transmission policy for remote parameter estimation in wireless sensing architectures under resource constraints. The goal of the proposed event-based policy is to derive a single optimal threshold value τ_s^* (with * denoting optimal) that maximizes the collected data's average reward rate such that, given a measured process $f(y;\boldsymbol{\theta}_0)$ with true (but unknown) parameters $\boldsymbol{\theta}_0$ (where, in general, $f(y;\boldsymbol{\theta})$ denotes the probability density function (PDF) of a random variable Y parameterized by $\boldsymbol{\theta} \in \Theta$ in the parameter space $\Theta \subset \mathbb{R}^p$), candidate data y is stored and transmitted if and only if $y \geq \tau_s^*$. The event values assigned to candidate data, $v = h_s(y)$,

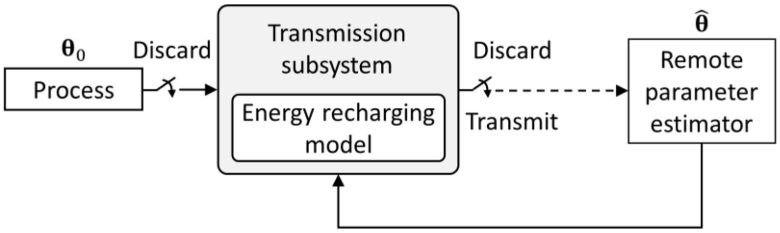


Figure 4-1. Event-based parameter estimation framework based on an energy recharging model.

follow a Bernoulli distribution such that $v = 1$ if $y \geq \tau_S^*$ and $v = 0$ if $y < \tau_S^*$. Since the value of each transmitted measurement is equal to one, maximizing the collected data's average reward rate is equivalent to maximizing the average transmission rate. Under this policy, the missingness of non-transmitted data is nonignorable and depends on the observed values: it is known that data, y, is missing because either 1) the sensing architecture's energy buffer is depleted and unable to measure the arriving message, 2) the sensing architecture measures the data and rejects it because $y < \tau^*$, or 3) the message is already stored in the data buffer but discarded before it is transmitted because the energy buffer is fully depleted. The subset of data successfully transmitted to the remote estimator according to the proposed optimal policy minimizes the variance of the measured process' parameter component estimates given the system constraints. The proposed policy accounts for system constraints imposed by the WSN architecture's energy and data buffer sizes, stochastic models of energy and event arrivals, the value of data, and temporal death. Here, temporal death refers to the rejection of any incoming data and the loss of any data that is already in the data buffer that has not yet been transmitted when a sensing node's energy buffer is fully depleted. Because this policy is governed by a single threshold value, τ_S^*, its implementation does not require knowledge of the SoC.

Following the problem formulation in Section 4.2 and the derivation of the optimal threshold, τ_s^*, in Section 4.3, Section 4.4 introduces a modified likelihood function that accounts for the missingness of data that results from the implementation of the optimal policy. We show that under regulatory conditions the modified MLE retains desirable properties that are characteristic of the standard full-information MLE (e.g., consistency, asymptotic normality) and minimizes the variance of parameter component estimates given the imposed constraints. Finally, numerical results are presented in Section 4.5 in which the proposed theoretical framework is used to control data collection and transmission in an EH WSN architecture subject to stringent energy constraints imposed by the availability of incoming energy and battery size. We illustrate that implementation of the optimal threshold produces the best possible estimate of the process parameters given the system constraints.

This work differentiates itself from, and improves upon, existing literature in two notable ways. First, the proposed policy accounts for—and places no restrictions on the size of—a WSN architecture's data storage buffer. Data transmission is often the most significant source of energy consumption in a wireless sensing node (Gastpar and Vetterli 2005); accounting for the storage of (potentially) large amounts of data in a buffer that are communicated to the remote estimator in batch transmissions—as opposed to transmitting each value individually immediately upon collection (as seen in Seyedi and Sikdar 2008; Uysal-Biyikoglu et al. 2002; Michelusi et al. 2012; Lei and Yates 2009; Tang and Tan 2017)—leads to significant gains in the transmission rate. Second, to the best of the author's knowledge this is the first chapter to derive a stochastic policy that governs data collection and transmission with the purpose of explicitly linking optimal remote parameter estimation to a rigorous stochastic energy recharging model that accounts for all significant hardware constraints.

4.2 Problem Formulation

The proposed framework guiding the energy-aware stochastic scheduling policy for remote parameter estimation discussed herein comprises of a transmission subsystem and a remote parameter estimator (Figure 4-1). Within the transmission subsystem, a sensing architecture measures a process with true parameters, $\boldsymbol{\theta}_0$. A binary value, $v \in \{0,1\}$, is assigned to each realization of the measured process, y, through the function $v = h_s(y)$ such that

$$v = h_s(y) = \begin{cases} 0, & \text{if } y < \tau_s^* \\ 1, & \text{if } y \geq \tau_s^* \end{cases} \tag{4.1}$$

If $v = 0$ the candidate data y is discarded and if $v = 1$ the candidate data y is stored in the data buffer, regardless of the SoC of the energy buffer. Once the data buffer is full, all stored data (denoted by the vector $\boldsymbol{\Lambda}_s$) is communicated to the remote parameter estimator in a single batch transmission. The subscript $s = 1,2,\ldots,s_t$ is introduced herein as an index to indicate that data collected and transmitted in batch s, $\boldsymbol{\Lambda}_s$, corresponds to the optimal threshold, τ_s^*, where $s = 1$ represents the first batch transmission after monitoring initiates and $s = s_t$ represents the most recent batch transmission. The proposed model utilizes batch transmission because the energy overhead of packet transmission is relatively high regardless of the payload size (Han et al. 2015). If the energy buffer is fully depleted at any point in time, temporal death occurs; the sensing node shuts down, all stored data is discarded, and all candidate data that arrive when the sensing node is powered down are rejected. Upon receiving the collected data, $\boldsymbol{\Lambda}_s$, the remote parameter estimator updates the maximum likelihood estimate, $\widehat{\boldsymbol{\theta}}_{s+1}$, recalculates the optimal threshold based on the updated estimate, and returns the updated threshold to the sensing node. When $s = 1$ (at monitoring initiation), the prior estimate is selected by the user, whereas the subsequent estimates, $\widehat{\boldsymbol{\theta}}_s$ for $s = 2,3, \dots, s_t$, used to calculate the optimal threshold for period s are the maximum likelihood estimates calculated from the previous period. Given this event-based parameter estimation framework, the goal is to derive the unique optimal threshold value, τ_s^*, that maximizes the collected data's average transmission rate given a WSN architecture's energy and data buffer sizes, stochastic models of energy and event arrivals, the value of data, and consideration of temporal death. Recall from Equation (4.1) that maximizing the expected transmission rate is equivalent to maximizing the expected reward rate because the reward (or value) of each transmitted measurement is one.

4.2.1 Transmission Subsystem Model

Consider a replenishable sensing architecture with finite-size energy and data storage buffers. While the system has a continuous state space (i.e., remaining energy) and a continuous parameter space (i.e., time), the state space is approximately modeled as a discrete state space. Let $X(t) \in \{0,1, \dots, N + k - 1\}$ be a finite continuous-time Markov chain with $N + k$ states as shown in Figure 4-2, where k is the number of candidate data measurements the storage buffer can hold. For $X(t) = k - 1 \leq n \leq N + k - 1$, the state at time t, $X(t) = n$, represents the remaining energy which can support the transmission of $n - (k - 1)$ data packets. Given the maximum capacity of the finite-size energy storage buffer is E_{max} and the amount of energy required to transmit the full data buffer of k values is E_T (this includes operational overhead of the embedded sensing architecture), then $N =$

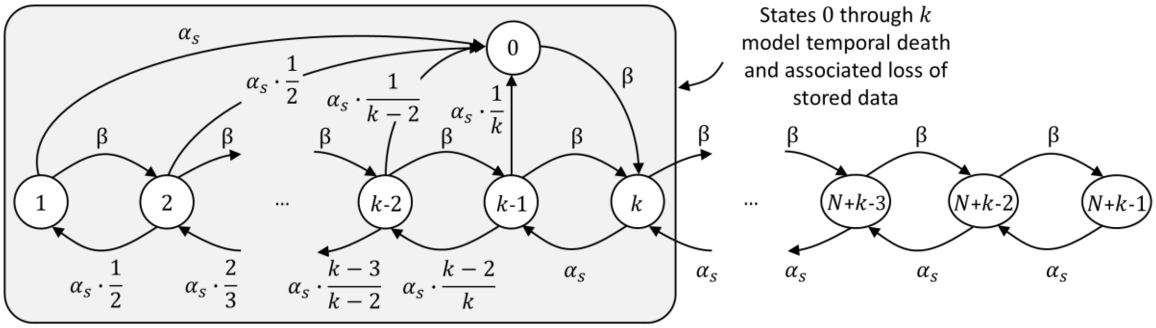


Figure 4-2. Markov chain rate transition diagram representing the energy renewal system.

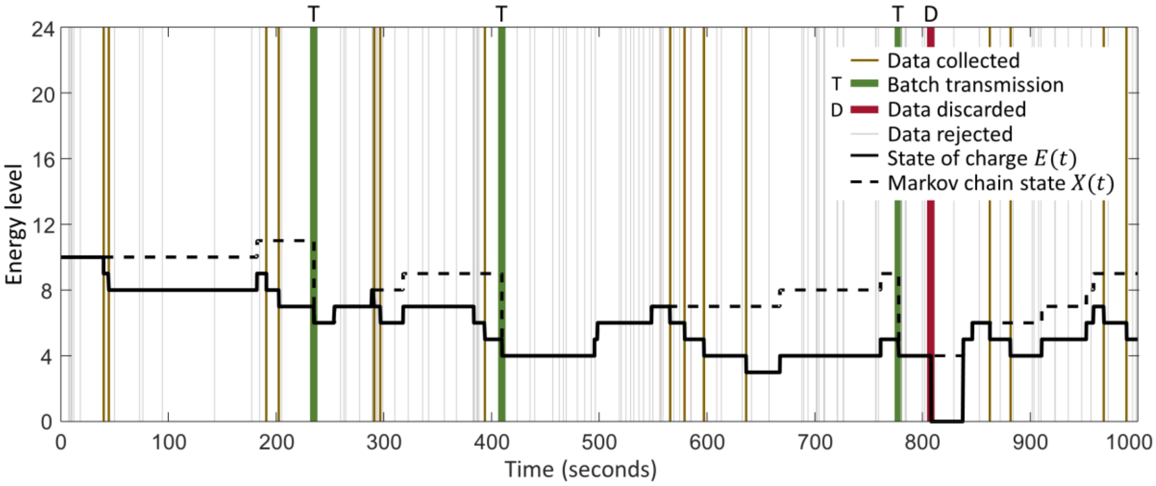


Figure 4-3. Comparison between the state transitions modeled by the proposed Markov chain and the state of charge.

$\left\lfloor \frac{E_{max} \cdot k}{E_T} \right\rfloor$ where $\lfloor x \rfloor$ is an operator rounding x to the nearest integer less than x for $x > 0$. This implies that the energy necessary for the batch transmission of k values is divided equally among the k stored values and "spent" as soon as each value is accepted and stored (i.e., the state transitions from $X(t) = n$ to $X(t) = n - 1$). In reality, the energy necessary to transmit all k values in the data buffer is spent all at once during the single batch transmission once the data buffer is full, not incrementally as each value is collected. This modelling assumption is justified because the proposed optimal policy is independent of the current SoC, meaning that implementation of the proposed policy does not require continuous measurement of the SoC to control the event-based decision making (i.e., accept or reject the candidate data). Additionally, let ℓ denote the number of values stored in the data buffer at time t. When the data buffer is empty $\ell = 0$, and when it is full $\ell = k$ and all data is transmitted. Upon each batch transmission when $\ell = k$, the modeled current state $X(t) = n$ is always equal to the actual state corresponding to the SoC. As long as $X(t) \neq 0 \leq n \leq k - 1$ upon transmission when $\ell = k$, then the transmission is successful. If $X(t) = 0 \leq n \leq k - 1$ upon transmission when $\ell = k$, the energy buffer runs out of energy and all data in the data buffer is discarded. Consequently, the Markov chain used to represent the energy renewal system and the actual physical model will have the same steady-state result. This is illustrated in Figure 4-3. Figure 4-3 serves as a representative example and compares the state transitions modeled by the Markov chain and the actual state of charge. As expected, the modeled state is always equal to the state of charge after every batch transmission. Figure 4-4 provides an overview of the aforementioned key components of the transmission subsystem and illustrates the process for either storing and transmitting data, storing and later discarding data due to temporal death, rejecting data according to the optimal policy, or rejecting data due to temporal death.

Candidate data, y, arrives as a memoryless Poisson process with rate λ (events per unit time). The energy storage buffer recharges based on arriving energy which arrives as a Poisson process with rate β (energy level per unit time). If $X(t) = k \leq n \leq N + k - 1$ $\forall \ell$ or $X(t) = 2 \leq n < k$ for $\ell < k$, the state transitions from n to $n - 1$ when candidate data is collected and stored. Here, the transition rate from state n to state $n - 1$ is $\alpha_s = \lambda \cdot P[y \geq \tau_s^*] = \lambda \cdot \bar{F}(\tau_s^*; \widehat{\boldsymbol{\theta}}_s)$ where $\bar{F}(\cdot)$ denotes the complementary cumulative distribution function (CDF). The expected reward corresponding to the optimal threshold, τ_s^*, is $R_s =$

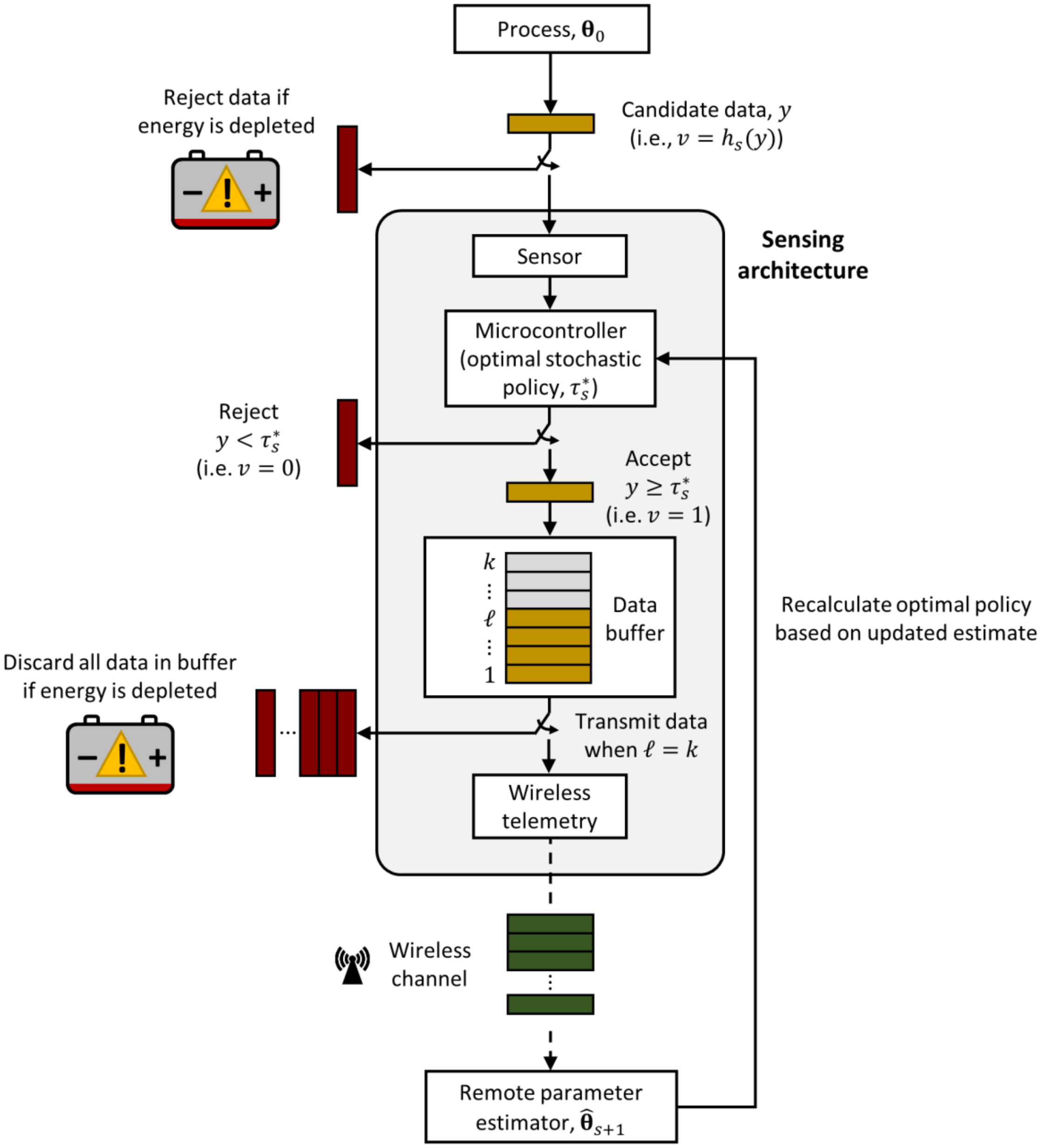


Figure 4-4. Transmission subsystem highlighting key components as well as the possible terminal conditions for candidate data.

$P[y \geq \tau_s^*]$. Formulating the reward process as a discrete-time Markov chain in which state transitions are made at infinitesimal time steps, Δ, the expected reward during each time step is $\lambda \cdot \Delta \cdot R_s$ for states $1 \leq n < k - 1$ and $k - 1 < n \leq N + k - 1$, and $\lambda \cdot \Delta \cdot \frac{k-1}{k} \cdot R_s$ for state $n = k - 1$., where $\lambda \cdot \Delta$ is the probability that an event will occur during the next time step, Δ. The time step, Δ, is assumed to be so small that at most one event can occur during the time step. A cost is incurred when the energy storage buffer is fully depleted and enters state $n = 0$ because all stored data is discarded when the sensing node shuts down. Based on the method of discretization of the finite-size battery—and consistent with the Markov chain in Figure 4-2—a state transition into state $n = 0$ can only occur when $\ell = k$, meaning that if the sensing architecture shuts down because there is not sufficient energy to complete the batch transmission, then it is known that k messages are lost. When this happens, the expected amount of time spent in state $n = 0$ is $\frac{1}{\beta}$ (or $\frac{1}{\beta\cdot\Delta}$ time steps). If we denote the expected cost incurred during a single time step, Δ, as C, then the expected total cost incurred in state $n = 0$ over time period $\frac{1}{\beta}$ is $\frac{1}{\beta\cdot\Delta}\cdot C$. The total expected cost incurred during the sojourn time in state $n = 0$ should equal the expected loss of the reward that would have been gained by the k discarded values that were discarded when the energy buffer was depleted. As a result,

$$C = -k \cdot \beta \cdot \Delta \cdot \frac{R_S}{P[y \geq \tau_S^*]} \tag{4.2}$$

where $\frac{R_S}{P[y \geq \tau_S^*]}$ is the conditional expected value of a transmitted realization of the measured process; C is not a function of τ_S^* because $\frac{R_S}{P[y \geq \tau_S^*]} = 1$ given the mapping between candidate data and value in Equation (4.1). Given the expected reward during each time step over states $1 \leq n \leq N + k - 1$ and the expected cost incurred during each time step in state $n = 0$, the reward vector is

$$\boldsymbol{r}_{S_{(N+k)\mathrm{x}1}} = \lambda \cdot \Delta \cdot \left[\frac{C}{\lambda\cdot\Delta},\ \ R_S \cdot \boldsymbol{e}_{1\mathrm{x}(k-2)},\ \ \frac{k-1}{k}\cdot R_S,\ \ R_S \cdot \boldsymbol{e}_{1\mathrm{x}N}\right]^T \tag{4.3}$$

The transition matrix corresponding to the Markov chain in Figure 4-2 is

$$\boldsymbol{P}_{s_{(N+k)\mathrm{x}(N+k)}}=\begin{bmatrix}
\gamma_0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \beta & 0 & \cdots & 0 & 0 & 0 & 0 \\
\alpha_s & \gamma_1 & \beta & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\alpha_s\cdot\frac{1}{2} & \alpha_s\cdot\frac{1}{2} & \gamma_2 & \beta & \ddots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\alpha_s\cdot\frac{1}{3} & 0 & \alpha_s\cdot\frac{2}{3} & \gamma_3 & \ddots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\alpha_s\cdot\frac{1}{k-3} & 0 & 0 & 0 & \cdots & \gamma_{k-3} & \beta & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\alpha_s\cdot\frac{1}{k-2} & 0 & 0 & 0 & \cdots & \alpha_s\cdot\frac{k-3}{k-2} & \gamma_{k-2} & \beta & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\alpha_s\cdot\frac{1}{k} & 0 & 0 & 0 & \cdots & 0 & \alpha_s\cdot\frac{k-2}{k} & \gamma_{k-1} & \beta & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & \alpha_s & \gamma_k & \beta & \ddots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \alpha_s & \gamma_{k+1} & \ddots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & \gamma_{N+k-4} & \beta & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & \alpha_s & \gamma_{N+k-3} & \beta & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \alpha_s & \gamma_{N+k-2} & \beta \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \alpha_s & \gamma_{N+k-1}
\end{bmatrix}\cdot\Delta \tag{4.4}$$

where Δ is sufficiently small such that the self-transitions (i.e., from state n to n) are $P_{n,n} \geq 0$. Here,

$$\gamma_n = \begin{cases}
\frac{1}{\Delta} - \beta, & \text{for} \quad n = 0 \\
\frac{1}{\Delta} - \beta - \alpha_s, & \text{for} \quad n = 1 \\
\frac{1}{\Delta} - \beta - \alpha_s, & \text{for} \quad 2 \leq n < k - 1 \\
\frac{1}{\Delta} - \beta - \alpha_s\cdot\frac{k-1}{k}, & \text{for} \quad n = k - 1 \\
\frac{1}{\Delta} - \beta - \alpha_s, & \text{for} \quad k \leq n \leq N + k - 2 \\
\frac{1}{\Delta} - \alpha_s, & \text{for} \quad n = N + k - 1
\end{cases}$$

The goal is to derive an optimal threshold vector, $\boldsymbol{\tau}^*_{s(N+k-1)\mathrm{x}1} = \tau^*_s \cdot \boldsymbol{e}$ —where $\boldsymbol{e}$ is a vector of ones—that is optimal in the sense that the policy maximizes the recurrent class' average reward rate (i.e., average transmission rate) given the recharge rate, β, candidate data arrival rate, λ, the size of the energy and data buffers, N and k, respectively, and a prior estimate of the estimated process parameters, $\widehat{\boldsymbol{\theta}}_1$. This entails calculating the value of a single optimal threshold value, τ_s^*, such that regardless of the current state, $X(t) = n$ for states $1 \leq n \leq N + k - 1$, it is optimal to transmit data if and only if the candidate value y is greater than or equal to τ_s^*. Note that we are able to evaluate the infinite horizon problem discussed herein because we are considering the average reward rate.

4.3 Deriving the Optimal Data Collection and Transmission Policy

4.3.1 Expected Aggregate Reward

For the Markov chain under consideration, let the aggregate expected reward be denoted

$$\boldsymbol{v}_{s(N+k)\text{x}1} = \boldsymbol{r}_s + \boldsymbol{P}_s \cdot \boldsymbol{r}_s + \cdots + \boldsymbol{P}_s^{m-1} \cdot \boldsymbol{r}_s = \sum_{h=0}^{m-1} \boldsymbol{P}_s^h \cdot \boldsymbol{r}_s \tag{4.5}$$

where $\boldsymbol{r}_s$ and $\boldsymbol{P}_s$ are defined in Equation (4.3) and Equation (4.4), respectively. Since Equation (4.5) does not have a limit as $m \to \infty$, we introduce a relative-gain vector, $\boldsymbol{w}$, which has a limit given by

$$\boldsymbol{w}_{(N+k)\text{x}1} = \lim_{m\to\infty} [\boldsymbol{v}_s - m \cdot g_s \cdot \boldsymbol{e}\,] = \lim_{m\to\infty} \left[\sum_{h=1}^{m-1} \left(\boldsymbol{P}_s^h - \boldsymbol{e} \cdot \boldsymbol{\pi}_s\right) \cdot \boldsymbol{r}_s\right] \tag{4.6}$$

Here, $g_s \cdot \boldsymbol{e}$ is the steady-state reward for the Markov chain's recurrent class (where $g_s \cdot \boldsymbol{e} = \lim_{m\to\infty} \boldsymbol{P}_s^m \cdot \boldsymbol{r}_s = \boldsymbol{e} \cdot \boldsymbol{\pi}_s \cdot \boldsymbol{r}_s$) and $\boldsymbol{\pi}_s$ is the steady-state vector (i.e., $\boldsymbol{\pi}_s = \boldsymbol{\pi}_s \cdot \boldsymbol{P}_s$). A proof for the existence of the limits in Equation (4.6) can be found in Gallager (Gallager 2013). There is certainly a single unique steady-state vector, $\boldsymbol{\pi}_s$, for the Markov chain shown in Figure 4-2 because the Markov chain has finite states and is ergodic.

Due to the complexity of calculating $\boldsymbol{w}$ from Equation (4.6) when there are a large number of states, we multiply both sides of Equation (4.6) by the transition matrix, $\boldsymbol{P}_s$, to get the following set of value-determination equations

$$\boldsymbol{P}_s \cdot \boldsymbol{w} = \lim_{m\to\infty} \sum_{h=0}^{m} \left(\boldsymbol{P}_s^h - \boldsymbol{e} \cdot \boldsymbol{\pi}_s\right) \cdot \boldsymbol{r}_s - (\boldsymbol{P}_s^0 - \boldsymbol{e} \cdot \boldsymbol{\pi}_s) \cdot \boldsymbol{r}_s = \boldsymbol{w} - \boldsymbol{r}_s + g_s \cdot \boldsymbol{e} \tag{4.7}$$

Since there are $N + k$ equations and $N + k + 1$ unknowns in Equation (4.7) and only the relative values of $\boldsymbol{w}$ are important, one component of $\boldsymbol{w}$ can be set to zero. The relative-gain vector, $\boldsymbol{w}$, then satisfies the equation $\boldsymbol{w} + g_s \cdot \boldsymbol{e} = \boldsymbol{r}_s + \boldsymbol{P}_s \cdot \boldsymbol{w}$ with $w_0 = 0$ and has a single unique solution. Analytically, the objective is to determine the unique optimal threshold value, τ_s^*, that maximizes the steady-state reward, g_s, in Equation (4.7).

4.3.2 Policy Improvement

Howard's Policy Improvement algorithm is used to determine the necessary and sufficient conditions that must be imposed on the value-determination equations in Equation (4.7) to derive the optimal policy that maximizes the average reward rate (Howard 1960). The use of Howard's Policy Improvement algorithm has been used widely in literature, such as in Lei et al. (Lei and Yates 2009). Howard's Policy Improvement algorithm consists of two primary stages which are applied sequentially and iteratively: the value-determination stage (Step 1) and the policy-improvement stage (Step 2 and Step 3). This procedure is outlined in Algorithm 4-1. Section 4.3.2.1 and Section 4.3.2.2 carry out Algorithm 4-1 for the energy renewal system under consideration in this chapter.

Algorithm 4-1. Howard's Policy Improvement algorithm.

Step 1:

Start with an arbitrary policy, $\boldsymbol{\tau}^A = [\tau^A, \tau^A, \dots, \tau^A]^T$, and calculate $\boldsymbol{w}^A$ and g^A from Equation (7).

Step 2:

Evaluate an alternate policy, $\boldsymbol{\tau} = [\tau, \tau, \dots, \tau]^T$, by defining a contraction mapping $T\colon \mathbb{R}^n \to \mathbb{R}^n$ as (Gedergruen et al. 1978)

$$T(\boldsymbol{\tau})\boldsymbol{w}^A = \boldsymbol{r}_s(\boldsymbol{\tau}) + \boldsymbol{P}_s(\boldsymbol{\tau}) \cdot \boldsymbol{w}^A \tag{4.8}$$

Step 3:

An optimal policy is achieved when, for all policies $\boldsymbol{\tau}$,

$$T(\boldsymbol{\tau})\boldsymbol{w}^* = \boldsymbol{r}_s(\boldsymbol{\tau}) + \boldsymbol{P}_s(\boldsymbol{\tau}) \cdot \boldsymbol{w}^* \leq T(\boldsymbol{\tau}_s^*)\boldsymbol{w}^* = \boldsymbol{r}_s(\boldsymbol{\tau}_s^*) + \boldsymbol{P}_s(\boldsymbol{\tau}_s^*) \cdot \boldsymbol{w}^* \tag{4.9}$$

4.3.2.1 Value Determination for the Energy Renewal System

Starting with an arbitrary policy, $\boldsymbol{\tau}^{\boldsymbol{A}} = \tau^A \cdot \boldsymbol{e}$, and implementing Equation (4.7) we get the following set of value determination equations

$$\boldsymbol{w}^A + g^A \cdot \boldsymbol{e} = \boldsymbol{r}_s + \boldsymbol{P}_s \cdot \boldsymbol{w}^A \tag{4.10}$$

For $n = 0$:

$$g^A = C + \beta \cdot \Delta \cdot w_k^A \tag{4.10a}$$

For $n = 1$:

$$w_2^A + \frac{\lambda}{\beta} \cdot \left(R_s - \frac{C}{\lambda \cdot \Delta} - \frac{\alpha_s}{\lambda} \cdot w_1^A \right) = w_1^A + w_k^A \tag{4.10b}$$

For $2 \leq n < k - 1$:

$$w_{n+1}^{A} + \frac{\lambda}{\beta} \cdot \left\{R_s - \frac{C}{\lambda \cdot \Delta} - \frac{\alpha_s}{\lambda} \cdot \left[w_n^{A} - \left(1 - \frac{1}{n}\right) \cdot w_{n-1}^{A}\right]\right\} = w_n^{A} + w_k^{A} \tag{4.10c}$$

For $n = k - 1$:

$$\frac{\lambda}{\beta} \cdot \left[R_s - \frac{C}{\lambda \cdot \Delta} - \frac{\alpha_s}{\lambda} \cdot \left(\frac{k-1}{k} \cdot w_{k-1}^{A} - \frac{k-2}{k} \cdot w_{k-2}^{A}\right)\right] = w_{k-1}^{A} \tag{4.10d}$$

For $k \leq n < N + k - 1$:

$$w_{n+1}^{A} + \frac{\lambda}{\beta} \cdot \left[R_s - \frac{C}{\lambda \cdot \Delta} - \frac{\alpha_s}{\lambda} \cdot (w_n^{A} - w_{n-1}^{A})\right] = w_n^{A} + w_k^{A} \tag{4.10e}$$

For $n = N + k - 1$:

$$\frac{\lambda}{\beta} \cdot \left[R_s - \frac{C}{\lambda \cdot \Delta} - \frac{\alpha_s}{\lambda} \cdot \left(w_{N+k-1}^{A} - w_{N+k-2}^{A}\right)\right] = w_k^{A} \tag{4.10f}$$

4.3.2.2 Policy Iteration for the Energy Renewal System

We define a contraction mapping $T: \mathbb{R}^n \to \mathbb{R}^n$ (Equation (4.8)) to evaluate an alternate policy, $\boldsymbol{\tau}$. Since an optimal policy is achieved when, for all policies $\boldsymbol{\tau}$, Equation (4.9) holds, then

$$[T(\boldsymbol{\tau})\boldsymbol{w}^*]' = 0 \tag{4.11}$$

at $\boldsymbol{\tau} = \boldsymbol{\tau}_s^*$, where $(\cdot)'$ denotes the derivative with respect to τ. The necessary and sufficient conditions imposed by Howard (Howard 1960) are characterized by Equation (4.11) and imposed on the value-determination equations in Equation (4.10). Because the optimal policy must be a function of a single threshold value, τ_s^*, carrying out the policy iteration stage is complicated by the fact that all $N + k$ equations described by Equation (4.10) have a common dependency on τ^A. To simplify the value-determination equations, we consolidate Equation (4.10e) and Equation (4.10f) into a single equation using backward recursion from the terminal condition in Equation (4.10f) such that w_k^A is only a function of w_{k-1}^A, τ^A, and known inputs, where

$$w_k^A = \left[\frac{\lambda}{\alpha_s} \cdot \left(R_s - \frac{C}{\lambda \cdot \Delta}\right) + \frac{w_{k-1}^A}{\sum_{j=0}^{N-1}\left(\frac{\beta}{\alpha_s}\right)^j}\right] \cdot \left[\frac{1}{\sum_{j=0}^{N-1}\left(\frac{\beta}{\alpha_s}\right)^j} + \frac{\beta}{\alpha_s}\right]^{-1} \tag{4.12}$$

This reduces Equation (4.10) from $N + k$ equations to the k equations shown in Equation (4.13), where w_k^A can be solved for explicitly as a function of only τ_s^A and known input parameters.

$$\begin{bmatrix} \alpha_s + \beta & -\beta & 0 & \cdots & 0 & 0 & \beta \\ -\alpha_s \cdot \frac{1}{2} & \alpha_s + \beta & -\beta & \ddots & 0 & 0 & \beta \\ 0 & -\alpha_s \cdot \frac{1}{3} & \alpha_s + \beta & \ddots & 0 & 0 & \beta \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & \alpha_s + \beta & -\beta & \beta \\ 0 & 0 & 0 & \ddots & -\alpha_s \cdot \frac{k\text{-}2}{k} & \alpha_s \cdot \frac{k\text{-}1}{k} + \beta & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\frac{1}{\sum_{j=0}^{N-1}\left(\frac{\beta}{\alpha_s}\right)^j} & \frac{1}{\sum_{j=0}^{N-1}\left(\frac{\beta}{\alpha_s}\right)^j} + \frac{\beta}{\alpha_s} \end{bmatrix} \cdot \begin{bmatrix} w_1^A \\ w_2^A \\ w_3^A \\ \vdots \\ w_{k-2}^A \\ w_{k-1}^A \\ w_k^A \end{bmatrix} = \begin{bmatrix} \lambda \cdot R_s - \frac{C}{\Delta} \\ \lambda \cdot R_s - \frac{C}{\Delta} \\ \lambda \cdot R_s - \frac{C}{\Delta} \\ \vdots \\ \lambda \cdot R_s - \frac{C}{\Delta} \\ \lambda \cdot \frac{k\text{-}1}{k} \cdot R_s - \frac{C}{\Delta} \\ \frac{\lambda}{\alpha_s} \cdot \left(R_s - \frac{C}{\lambda \cdot \Delta}\right) \end{bmatrix} \tag{4.13}$$

Returning to Equation (4.11), we note that since the optimal threshold is the same across all states, we need only calculate the optimal threshold such that one component of Equation (4.11) equals zero. Consequently,

$$[T(\boldsymbol{\tau})w_0^*]' = C + \beta \cdot \Delta \cdot \left(w_k^A\right)' = 0 \tag{4.14}$$

where w_k^A comes directly from Equation (4.13). An explicit solution for the optimal threshold, τ_s^*, is calculated from Equation (4.14); further simplification of the expression for the optimal threshold, τ_s^*, requires knowledge of the characterization of $f(y; \boldsymbol{\theta}_0)$ (i.e., the process distribution type). It follows that the optimal policy calculated from Equation (4.14) has a maximum average reward rate of $\frac{C}{\Delta} + \beta \cdot w_k^*$.

4.4 Remote Parameter Estimation (MLE) Using a Modified Likelihood Function

The optimal policy derived in Section 4.3 maximizes the expected transmission rate of candidate data measured from a process, $f(y; \boldsymbol{\theta}_0)$. If there is unlimited energy supplied to the WSN—meaning all candidate data are transmitted—then the standard likelihood function given full information is

$$L(\boldsymbol{\theta}) = \prod_{s=1}^{s_t} \left[\prod_{i=1}^{k} f(\Lambda_{s,i}; \boldsymbol{\theta}) \right] \tag{4.15}$$

For the case of WSNs under energy constraints, the information that is collected and sent to the remote parameter estimator during batch transmission s reflects only a subset of candidate. By implementing the optimal policy, τ_s^*, the missingness of non-transmitted data is classified as not missing at random (NMAR) (Arnab 2017) because it depends on the observed values: it is known that data, y, is missing because either 1) the sensing architecture's energy buffer is depleted and unable to measure the arriving message, 2) the sensing architecture measures the data and rejects it because $y < \tau_s^*$, or 3) the message is already stored in the data buffer but discarded before it is transmitted because the energy buffer is fully depleted. Because the rejected and discarded data that is never transmitted to the remote estimator is categorized as NMAR, the missing data is nonignorable and the likelihood function characterizing transmitted data must account for consideration of left- and right-censored data. Given the optimal threshold, τ_s^*, corresponding to batch transmission s, we introduce the following modified likelihood function that allows us to have a maximizer even when data is censored

$$L(\boldsymbol{\theta}) = \prod_{s=1}^{S_t} \left[\prod_{i=1}^{k} f(\Lambda_{s,i}; \boldsymbol{\theta}) \cdot \prod_{i=1}^{l_{c,s}} F(\tau_s^*; \boldsymbol{\theta}) \cdot \prod_{i=1}^{r_{c,s}} \bar{F}(\tau_s^*; \boldsymbol{\theta}) \right] \tag{4.16}$$

Here, $l_{c,s}$ is the number of candidate data measurements that are left censored and $r_{c,s}$ is the number of candidate data measurements that are right censored, where

$$r_{c,s} = k \cdot N_{D,s} + T_s \cdot \lambda \cdot \bar{F}(\tau_s^*, \boldsymbol{\theta}) \cdot \left(\pi_0 + \frac{\pi_{k-1}}{k-1} \right) \tag{4.17a}$$

$$l_{c,s} = k \cdot \left(1 + N_{D,s}\right) \cdot \frac{F(\tau_s^*, \boldsymbol{\theta})}{\bar{F}(\tau_s^*, \boldsymbol{\theta})} + T_s \cdot \lambda \cdot F(\tau_s^*, \boldsymbol{\theta}) \cdot \left(\pi_0 + \frac{\pi_{k-1}}{k-1} \right) \tag{4.17b}$$

In Equation (4.17), $N_{D,s}$ is the number of times the sensing unit runs out of energy during data collection period s, T_s is the amount of time between the start and finish of data collection for batch transmission s, and π_0 and π_{k-1} are the first and k^{th} components of the steady-state vector, $\boldsymbol{\pi}_s$. The first term in Equation (4.17a) accounts for the number of accepted messages that are stored but ultimately discarded before batch transmission occurs, and the first term in Equation (4.17b) accounts for the expected number of rejected messages. The second terms in Equation (4.17a) and Equation (4.17b) account for the loss of candidate data that could not be measured because messages arrived while the sensing architecture's energy buffer was fully depleted.

Under regulatory conditions, the standard full-information MLE described by the likelihood function in Equation (4.15) is a consistent, asymptotically normal, and efficient estimator. In this section we show that for the modified likelihood function characterized by Equation (4.16), the MLE is the maximizer, there exists a Cramer-Rao bound (CRB) on the covariance matrix of the estimator, the MLE is consistent, asymptotically unbiased, and asymptotically normal, and minimum variance of the MLE components is achieved under the optimal data collection and transmission policy given the system constraints. For ease of notation we will consider $s_t = 1$ for all proofs since the following discussion extends naturally for $s_t > 1$.

Assumptions:

We consider the following regularity conditions herein. Let $\boldsymbol{\theta} \in \Theta$ be a p x 1 vector and assume:

A1. $\boldsymbol{\Theta}$ is an open subset of $\mathbb{R}^p$

A2. $f(y; \boldsymbol{\theta})$ is smooth (Ibragimov and Has'minskii 1981) and differentiable in $\boldsymbol{\theta}$

A3. The covariance matrix, $\mathrm{cov}_{\boldsymbol{\theta}}(\widehat{\boldsymbol{\theta}})$, and the Fisher information matrix, $\mathbf{F}(\boldsymbol{\theta})$ (defined in *Theorem 1*), are non-singular matrices

A4. The support of y, $\{y : f(y; \boldsymbol{\theta}) > 0\}$, does not depend on $\boldsymbol{\theta}$

A5. The model is identifiable, meaning that for every $\boldsymbol{\theta} \in \Theta$, there does not exist another $\widetilde{\boldsymbol{\theta}} \in \Theta$ such that $f(y; \boldsymbol{\theta}) = f(y; \widetilde{\boldsymbol{\theta}})$ for all $\boldsymbol{\Lambda}_s$ in the sample space

Theorem 1:

If the modified log-likelihood of y, $\ell(\boldsymbol{\theta})$, satisfies the regulatory conditions, then the modified log-likelihood is a concave function.

Proof: The modified log-likelihood function, $\ell(\boldsymbol{\theta})$, is a concave function if $-\nabla^2_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta})$ is a positive semi-definite matrix. Let $\nabla_{\boldsymbol{\theta}}$ and $\nabla^2_{\boldsymbol{\theta}}$ denote the gradient and Hessian operators with respect to $\boldsymbol{\theta}$, respectively. The modified log-likelihood function for a single observation is

$$\ell(\boldsymbol{\theta}) = (1 - P_D) \cdot \bar{F}(\tau_s^*; \boldsymbol{\theta}_0) \cdot \ln f(y; \boldsymbol{\theta}) + F(\tau_s^*; \boldsymbol{\theta}_0) \cdot \ln F(\tau_s^*; \boldsymbol{\theta}) + P_D \cdot \bar{F}(\tau_s^*; \boldsymbol{\theta}) \cdot \ln \bar{F}(\tau_s^*; \boldsymbol{\theta})$$

(4.18)

where P_D is the probability that candidate data is discarded due to temporal death (which is not a function of $\mathbf{\Lambda}_{s,i}$).

Step 1: First, we show that the gradient of the expectation of the modified log-likelihood function is maximum at the true parameter and is the unique maximum.

$$E_{\boldsymbol{\theta}}[\nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta})]$$

$$= E_{\boldsymbol{\theta}}[(1-P_D)\cdot\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)\cdot\nabla_{\boldsymbol{\theta}}\ln f(y;\boldsymbol{\theta}) + F(\tau_S^*;\boldsymbol{\theta}_0)\cdot\nabla_{\boldsymbol{\theta}}\ln F(\tau_S^*;\boldsymbol{\theta}) + P_D\cdot\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)\cdot$$
$$\nabla_{\boldsymbol{\theta}}\ln\bar{F}(\tau_S^*;\boldsymbol{\theta})]$$

$$= E_{\boldsymbol{\theta}}\left[(1-P_D)\cdot\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)\cdot\frac{\nabla_{\boldsymbol{\theta}}f(y;\boldsymbol{\theta})}{f(y;\boldsymbol{\theta}_0)} + F(\tau_S^*;\boldsymbol{\theta}_0)\cdot\frac{\nabla_{\boldsymbol{\theta}}F(\tau_S^*;\boldsymbol{\theta})}{F(\tau_S^*;\boldsymbol{\theta}_0)} + P_D\cdot\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)\cdot\frac{\nabla_{\boldsymbol{\theta}}\bar{F}(\tau_S^*;\boldsymbol{\theta})}{\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)}\right]$$

$$= \nabla_{\boldsymbol{\theta}}\mathbf{1}$$

$$= \mathbf{0} \tag{4.19}$$

Step 2: Second, we let $\widehat{\boldsymbol{\theta}}$ denote an unbiased estimator of $\boldsymbol{\theta}$ and demonstrate that the correlation between the estimator and the gradient of the log-likelihood is constant.

$$E_{\boldsymbol{\theta}}\left[\nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta})\cdot\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\right]$$

$$= E_{\boldsymbol{\theta}}\left[\nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta})\cdot\widehat{\boldsymbol{\theta}}^T\right] - E[\nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta})]\cdot\boldsymbol{\theta}^T \quad \text{(where } E_{\boldsymbol{\theta}}[\nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta})] = \mathbf{0} \text{ from Step 1)}$$

$$= E_{\boldsymbol{\theta}}\left[(1-P_D)\cdot\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)\cdot\nabla_{\boldsymbol{\theta}}\ln f(y;\boldsymbol{\theta})\cdot\widehat{\boldsymbol{\theta}}^T + F(\tau_S^*;\boldsymbol{\theta}_0)\cdot\nabla_{\boldsymbol{\theta}}\ln F(\tau_S^*;\boldsymbol{\theta})\cdot\widehat{\boldsymbol{\theta}}^T + P_D\cdot\right.$$
$$\left.\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)\cdot\nabla_{\boldsymbol{\theta}}\ln\bar{F}(\tau_S^*;\boldsymbol{\theta})\cdot\widehat{\boldsymbol{\theta}}^T\right]$$

$$= E_{\boldsymbol{\theta}}\left[(1-P_D)\cdot\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)\cdot\frac{\nabla_{\boldsymbol{\theta}}f(y;\boldsymbol{\theta})}{f(y;\boldsymbol{\theta}_0)}\cdot\widehat{\boldsymbol{\theta}}^T + F(\tau_S^*;\boldsymbol{\theta}_0)\cdot\frac{\nabla_{\boldsymbol{\theta}}F(\tau_S^*;\boldsymbol{\theta})}{F(\tau_S^*;\boldsymbol{\theta}_0)}\cdot\widehat{\boldsymbol{\theta}}^T + P_D\cdot\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)\cdot\right.$$
$$\left.\frac{\nabla_{\boldsymbol{\theta}}\bar{F}(\tau_S^*;\boldsymbol{\theta})}{\bar{F}(\tau_S^*;\boldsymbol{\theta}_0)}\cdot\widehat{\boldsymbol{\theta}}^T\right]$$

$$= \nabla_{\boldsymbol{\theta}} \cdot E_{\boldsymbol{\theta}}\left(\hat{\boldsymbol{\theta}}^T\right)$$

$$= \mathbf{I} \tag{4.20}$$

Step 3: The covariance matrix of the concatenated estimator error and gradient gives the relation between the estimator covariance and the Fisher information matrix. Define a random vector $\boldsymbol{U}$ as

$$\boldsymbol{U} = \begin{bmatrix} \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \\ \nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}) \end{bmatrix} \tag{4.21}$$

Since any matrix expressed as an outer product of two vectors is non-negative definite,

$$E_{\boldsymbol{\theta}}[\boldsymbol{U}\boldsymbol{U}^T] \geq 0 \tag{4.22}$$

Using the results of Steps 1 and 2, we have

$$E_{\boldsymbol{\theta}}[\boldsymbol{U}\boldsymbol{U}^T] = \begin{bmatrix} \mathrm{cov}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}) & \mathbf{I} \\ \mathbf{I} & \mathbf{F}(\boldsymbol{\theta}) \end{bmatrix} \geq 0 \tag{4.23}$$

Since $E_{\boldsymbol{\theta}}[\boldsymbol{U}\boldsymbol{U}^T]$ is a partitioned symmetric matrix that is positive semi-definite and it is assumed that $\mathbf{F}(\boldsymbol{\theta})$ is a non-singular matrix, then $\mathrm{cov}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}})$ is positive semi-definite, $\mathbf{F}(\boldsymbol{\theta})$ is positive semi-definite, and

$$\mathrm{cov}_{\boldsymbol{\theta}}\left(\hat{\boldsymbol{\theta}}\right) - \mathbf{F}^{-1}(\boldsymbol{\theta}) \geq 0 \tag{4.24}$$

Since, by definition,

$$\mathbf{F}(\boldsymbol{\theta}) = -E_{\boldsymbol{\theta}}\left[\nabla_{\boldsymbol{\theta}}^2 \ell(\boldsymbol{\theta})\right] \tag{4.25}$$

and $\mathbf{F}(\boldsymbol{\theta}) \geq 0$, the modified log-likelihood function, $\ell(\boldsymbol{\theta})$, is a concave function and there exists a unique maximum likelihood estimator $\hat{\boldsymbol{\theta}}$.

Theorem 2: Suppose that the modified log-likelihood of y, $\ell(\boldsymbol{\theta})$, satisfies the regulatory conditions. There exists a Cramer-Rao bound (CRB) on the covariance matrix of the estimator.

Proof: The result comes directly from Equation (4.24) in *Theorem 1*.

Theorem 3: If the modified log-likelihood of y, $\ell(\boldsymbol{\theta})$, satisfies the regulatory conditions, then the MLE is consistent, where consistency is defined as

$$\widehat{\boldsymbol{\theta}}_n \xrightarrow{P} \boldsymbol{\theta}_0 \text{ as } n \to \infty \tag{4.26}$$

where n is the number of samples.

Proof: In *Theorem 1* we prove that $\widehat{\boldsymbol{\theta}}$ is the value of $\boldsymbol{\theta}$ which maximizes the modified log-likelihood in Equation (4.18). We now note that for any $\boldsymbol{\theta} \in \Theta$ the Law of Large Numbers implies the convergence in probability of

$$\frac{1}{n} \cdot \ell(\boldsymbol{\theta}) \xrightarrow{P} E_{\boldsymbol{\theta}_0}[\ell(\boldsymbol{\theta})] \tag{4.27}$$

Since, by *Theorem 1*, $E_{\boldsymbol{\theta}}[\ell(\boldsymbol{\theta})]$ is uniquely maximum at $\boldsymbol{\theta}_0$, then $\widehat{\boldsymbol{\theta}}_n \xrightarrow{p} \boldsymbol{\theta}_0$.

Theorem 4: If the modified log-likelihood of y, $\ell(\boldsymbol{\theta})$, satisfies the regulatory conditions, then the MLE is asymptotically unbiased and normal, where asymptotic normality is defined as

$$\sqrt{n} \cdot \left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) \xrightarrow{d} \boldsymbol{\mathcal{N}}_p(0, \boldsymbol{F}(\boldsymbol{\theta})^{-1}) \text{ as } n \to \infty \tag{4.28}$$

Here, $\boldsymbol{F}(\boldsymbol{\theta})$ is the single sample Fisher information matrix.

Proof: In *Theorem 1* we prove that $\widehat{\boldsymbol{\theta}}$ maximizes the modified log-likelihood function, meaning that $\nabla_{\boldsymbol{\theta}}\ell(\widehat{\boldsymbol{\theta}}) = \mathbf{0}$. From *Theorem 3*, consistency of the MLE ensures convergence in probability of $\widehat{\boldsymbol{\theta}}_n$ to $\boldsymbol{\theta}_0$ as $n \to \infty$. This justifies the application of a first-order Taylor expansion to $\nabla_{\boldsymbol{\theta}}\ell(\widehat{\boldsymbol{\theta}}) = \mathbf{0}$ around $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0$ such that $\mathbf{0} \approx \nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta}_0) + (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \cdot \nabla_{\boldsymbol{\theta}}^2\ell(\boldsymbol{\theta}_0)$ and

$$\sqrt{n} \cdot (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \approx -\left[\frac{1}{\sqrt{n}} \cdot \nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta}_0)\right] \cdot \left[\frac{1}{n} \cdot \nabla_{\boldsymbol{\theta}}^2\ell(\boldsymbol{\theta}_0)\right]^{-1} \tag{4.29}$$

By the Law of Large Numbers, the second term in Equation (4.29) is

$$\frac{1}{n} \cdot \nabla_{\boldsymbol{\theta}}^2\ell(\boldsymbol{\theta}_0) \xrightarrow{P} E_{\boldsymbol{\theta}_0}\left[\nabla_{\boldsymbol{\theta}}^2\ell(\boldsymbol{\theta})\right] = -\boldsymbol{F}(\boldsymbol{\theta}_0) \tag{4.30}$$

in probability. Further, by the Central Limit Theorem, since $\nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta}_0)$ has mean 0 and covariance $\boldsymbol{F}(\boldsymbol{\theta}_0)$, the first term in Equation (4.29) can be expressed as

$$\frac{1}{\sqrt{n}} \cdot \nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta}_0) \xrightarrow{d} \boldsymbol{\mathcal{N}}_p(0, \boldsymbol{F}(\boldsymbol{\theta}_0)) \tag{4.31}$$

in distribution. We can conclude by substituting Equations (4.30)-(4.31) into Equation (4.29) that by the Continuous Mapping Theorem and Slutsky's Lemma, Equation (4.28) holds and the MLE is asymptotically unbiased and asymptotically normal.

Lemma 1: Suppose that the modified log-likelihood of y, $\ell(\boldsymbol{\theta})$, satisfies the regulatory conditions. Let $X = H(y)$ be some statistic with arbitrary dimension of the original data. Let $\ell_{H(y)}(X; \boldsymbol{\theta})$ denote the log-likelihood of $X = H(y)$. Then

$$\boldsymbol{F}(\boldsymbol{\theta}) - \boldsymbol{F}_{H(y)}(\boldsymbol{\theta}) \geq 0 \tag{4.32}$$

where

$$\boldsymbol{F}_{H(y)}(\boldsymbol{\theta}) = -E_{\boldsymbol{\theta}}\left[\nabla^2_{\boldsymbol{\theta}}\ell_{H(y)}(\boldsymbol{\theta})\right] \tag{4.33}$$

Proof: Refer to Rao (Rao 2017) and Ly et al. (Ly et al. 2017) for the complete proof.

Theorem 5: Minimum variance of the MLE parameter components is achieved under the optimal data collection and transmission policy, given the system constraints.

Proof: Let P_T denote the probability that candidate data is accepted and transmitted. From Equation (4.18),

$$P_T = (1 - P_D) \cdot \bar{F}(\tau_s^*; \boldsymbol{\theta}_0) \tag{4.34}$$

Since the optimal policy maximizes the average transmission rate, $\frac{g_s}{\Delta}$, then

$$\underset{\tau \in \mathbb{R}_{>0}}{\operatorname{argmax}}\left(\frac{g_s}{\Delta}\right) = \tau_s^* \Leftrightarrow \underset{\tau \in \mathbb{R}_{>0}}{\operatorname{argmax}}(P_T) = \tau_s^* \tag{4.35}$$

where the steady-state reward, g_s, is defined in Equation (4.7). By *Lemma 1*, a transmitted measurement y contains more information about $\boldsymbol{\theta}$ than a censored measurement. It follows directly from Equation (4.35) and *Lemma 1* that

$$\boldsymbol{F}(\boldsymbol{\theta}) - \boldsymbol{F}^A(\boldsymbol{\theta}) \geq 0 \tag{4.36}$$

where $\boldsymbol{F}^A(\boldsymbol{\theta})$ is the Fisher Information Matrix corresponding to any alternate policy that is not the optimal policy. Since $\boldsymbol{F}(\boldsymbol{\theta})$ and $\boldsymbol{F}^A(\boldsymbol{\theta})$ are symmetric matrices and $\boldsymbol{F}(\boldsymbol{\theta}) \geq \boldsymbol{F}^A(\boldsymbol{\theta}) \Leftrightarrow \boldsymbol{F}(\boldsymbol{\theta})^{-1} \leq \boldsymbol{F}^A(\boldsymbol{\theta})^{-1}$, then $\boldsymbol{F}(\boldsymbol{\theta})_{ii}^{-1} \leq \boldsymbol{F}^A(\boldsymbol{\theta})_{ii}^{-1}\ \forall i$. Consequently, we conclude that the scalar lower bounds on the variance of each component of $\widehat{\boldsymbol{\theta}}$ of the CRB are minimized under the optimal policy.

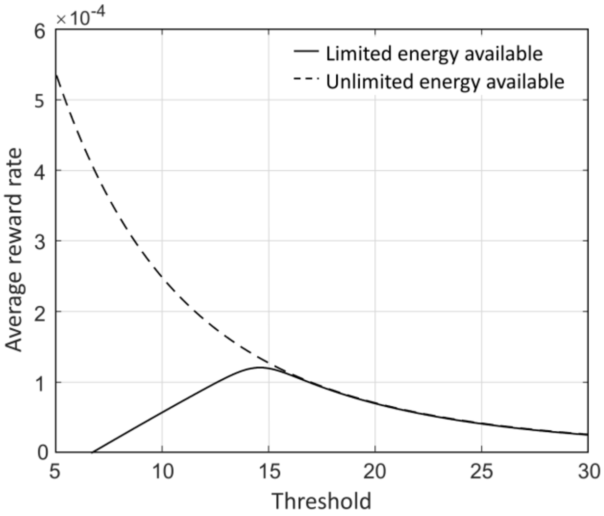


Figure 4-5. Comparison of the average reward rate corresponding to varying threshold values for WSNs with limited ($\beta = 0.8$) and unlimited ($\beta = \infty$) available energy assuming $\xi_1 = 0.2$ and $\sigma_1 = 5$.

4.5 Numerical Results

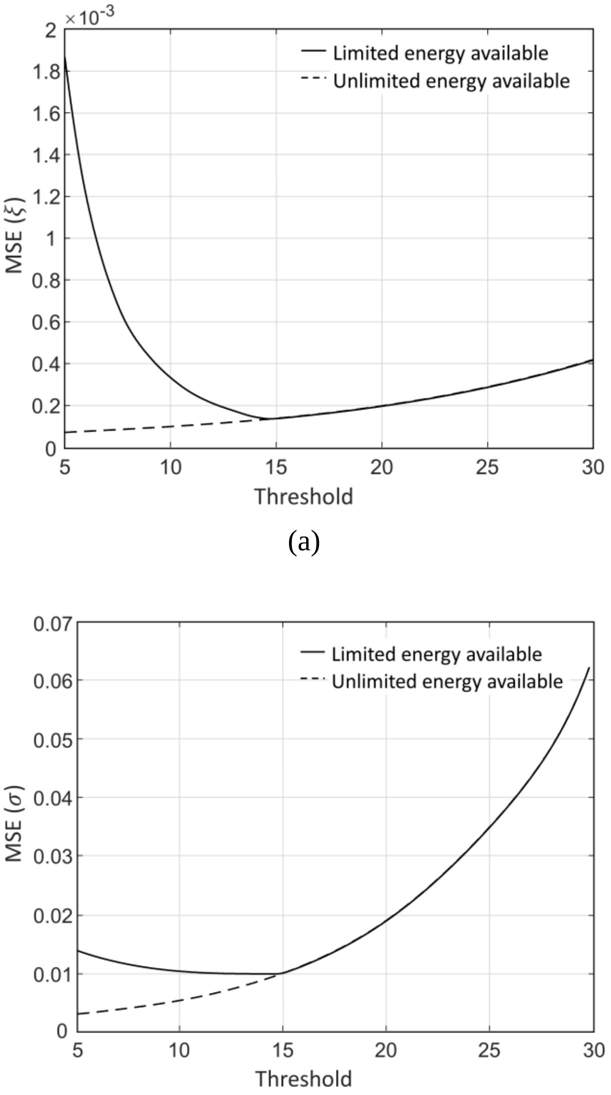
Consider an energy renewal system modeled by the Markov chain in Figure 4-2 with the following parameters:

- The total monitoring period is 50 hours
- Energy arrives with rate $\beta = 0.8$ (energy-level/minute)
- Candidate data arrive with rate $\lambda = 8$ (events/minute)
- The measured process follows a Generalized Pareto distribution (GPD) characterized by

$$f(y; \boldsymbol{\theta}_0) = \frac{1}{\sigma_0} \cdot \left[1 + \xi_0 \cdot \frac{y}{\sigma_0}\right]^{-\frac{1}{\xi_0}-1} \tag{4.37}$$

  with true parameters $\boldsymbol{\theta}_0 = [\xi_0, \sigma_0]^T$, where $\xi_0 = 0.2$ and $\sigma_0 = 5$

- The energy storage buffer can support four transmissions with $k = 5$ messages each (i.e., $N = 20$)



(a)

(b)

Figure 4-6. Comparison of the mean square error (MSE) of the (a) shape parameter and (b) scale parameter corresponding to varying threshold values for 1000 iterations of the first four hours of the monitoring period.

- The original belief about the measured process before the monitoring period begins follows a GPD with parameters $\widehat{\boldsymbol{\theta}}_1 = [\xi_1, \sigma_1]^T$ (various initial parameters are considered and specified in the subsequent plots)

Given the input parameters characterizing the energy renewal system, Figure 4-5 shows the average reward rate corresponding to varying threshold values for the WSN assuming

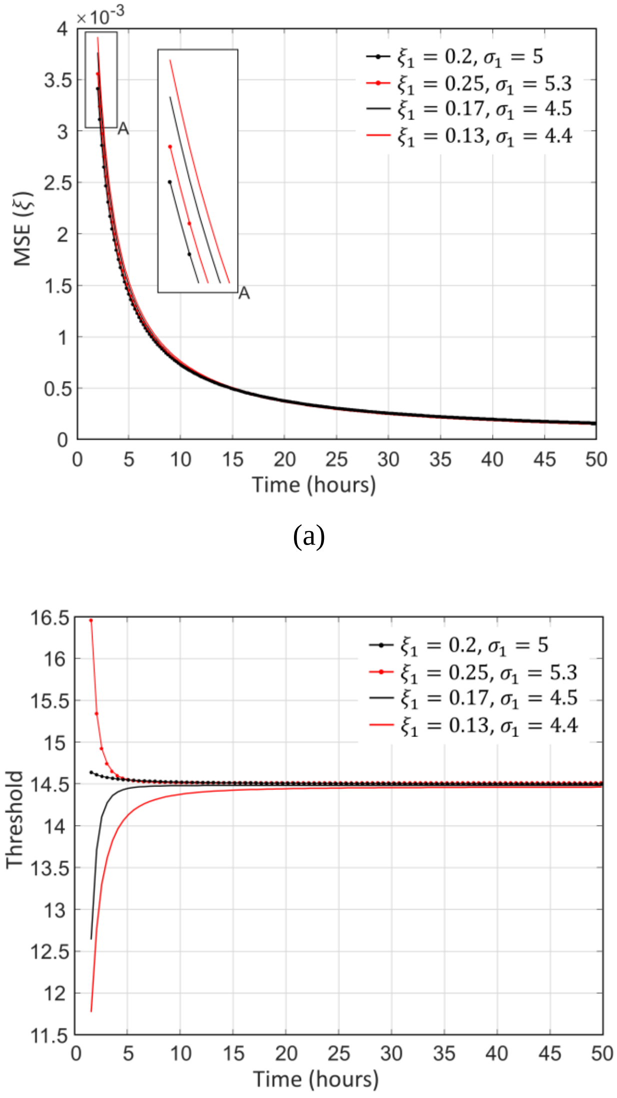
$\xi_1 = 0.2$ and $\sigma_1 = 5$ (i.e., the measured process' parameters are known). Figure 4-5 compares the average reward rate corresponding to varying threshold values for the case of limited energy to that of unlimited energy to illustrate that as $\tau_s \gg \tau_s^*$, the probability of storing candidate data decreases and the limited energy and unlimited energy average reward rates converge. Conversely, as $\tau_s \ll \tau_s^*$, the availability of energy becomes a bottleneck when energy is limited because the probability of storing candidate data increases, leading to repeated depletion of the energy buffer.

In Figure 4-6 we compare the mean square error (MSE) of the GPD's estimated shape and scale parameters corresponding to varying threshold values for 1000 iterations over the first four hours of the monitoring period. As expected based on the results of *Theorem 5* in Section 4.4, the MSE of the MLE components is minimized at the optimal threshold, τ_s^*, which is illustrated in Figure 4-5 and computed analytically from Equation (4.14) in Section 4.3. In order to assess the robustness of the proposed method to (potentially large) uncertainty in the estimate of the measured process parameters before the monitoring period begins, Figure 4-7(a) illustrates the convergence of the shape parameter MSE over the entire monitoring period for 1000 iterations given different initial estimates of the measured process parameters. Despite having incorrect estimates of the process parameters that govern data collection during the first batch transmission period, the estimates rapidly converge. This is confirmed in Figure 4-7(b) which shows the convergence of the implemented threshold value towards the optimal threshold value corresponding to $\boldsymbol{\theta}_0$.

4.6 Conclusion

This chapter presents the derivation and implementation of an optimal data collection and transmission policy for remote parameter estimation in wireless sensing architectures under resource constraints. Given a WSN architecture's energy and data buffer sizes, stochastic models of energy and event arrivals, the value of data, and temporal death, the proposed policy controls the storage and transmission of candidate data such that the variance of the estimated measured process' parameter components is minimized. Because this work models a replenishable WSN node as a continuous-time Markov chain for which a single unique threshold value is optimal across the entire state space for $X(t) = 1 \leq n \leq N + k - 1$, the implemented policy is independent of the energy buffer's SoC and does not place assumptions on exact knowledge or the ability to measure the energy level in real time. The proposed policy also accounts for—and places no restrictions on the size of—a WSN architecture's data storage buffer, which broadens the scope of its possible applications.

Based on the optimal policy, only a subset of the measured information is wirelessly transmitted to the estimator for processing; candidate data that are not transmitted are missing because the sensing architecture's energy buffer is depleted and unable to measure the arriving message, the sensing architecture measures the data and rejects it because $y < \tau_s^*$, or the message is already accepted and stored in the data buffer but discarded before it is transmitted because the energy buffer is fully depleted. A modified likelihood function is proposed that accounts for this missingness of data. We show that under regulatory conditions the modified MLE retains desirable properties that are characteristic of the full-information MLE assuming unlimited energy (i.e., all data is transmitted).



(a)

(b)

Figure 4-7. Convergence of the (a) estimated shape parameter MSE and (b) implemented threshold value over the entire monitoring period for 1000 iterations given different initial estimates of the measured process parameters.

Numerical results are presented and reflect the objectives of the theoretical results: given the proposed EH WSN architecture is subject to stringent energy constraints imposed by the availability of incoming energy and battery size, implementation of the optimal threshold, τ_S^*, produces the best possible estimate of the process parameters given the system constraints. The numerical results also indicate that the proposed framework for optimal remote parameter estimation under resource constraints is robust against even significant amounts of uncertainty in the parameter estimates at the outset of the monitoring period.

Since energy arrival and event arrival rates are rarely constant over long periods of time when WSNs rely on EH from the environment, this research would benefit from future work that incorporates predictive estimation of the recharge rate, β, and event arrival rate, λ. Short-term forecasting will also enable the integration of online supervisory control schemes that dynamically control whether the sensing architecture is accepting measurements or in a low-power sleep state in which all incoming data is rejected, even when the energy buffer is not fully depleted. Next steps should also include expanding this theoretical framework from a single-hop to network-level scale.

4.7 References

Arnab R. Nonsampling errors. In: *Survey sampling theory and applications*. London: Academic Press, 2017, pp. 469-504.

Gallager RG. *Stochastic processes: theory for applications*. Cambridge: Cambridge University Press, 2013.

Gastpar M and Vetterli M. Power, spatio-temporal bandwidth, and distortion in large sensor networks. *IEEE J Sel Areas Commun* 2005; 23(4): 745-754.

Han D, You K, Xie L et al. Optimal parameter estimation under controlled communication over sensor networks. *IEEE Trans Signal Process* 2015; 63(24): 6473-6485.

Howard RA. *Dynamic programming and Markov processes*. Cambridge: MIT Press, 1960.

Ibragimov IA and Has'minskii RZ. *Statistical estimation: asymptotic theory.* New York: Springer-Verlag, 1981.

Lei J and Yates R. A generic model for optimizing single-hop transmission policy of replenishable sensors. *IEEE Trans Wireless Commun* 2009; 8(2): 547-551.

Ly A, Marsman M, Verhagen J, et al. A tutorial on Fisher information, https://arxiv.org/abs/ 1705.01064 (2017, accessed 2 June 2020).

Michelusi N, Stamatiou K and Zorzi M. On optimal transmission policies for energy harvesting devices. In: *2012 Information Theory and Applications Workshop*, San Diego, 5-10 February 2012.

Ozel O, Tutuncuoglu K, Ulukus S, et al. Fundamental limits of energy harvesting communications. *IEEE Commun Mag* 2015; 53(4): 126-132.

Rao SS. *Advanced statistical inference*. College Station: Texas A&M University, 2017.

Seyedi A and Sikdar B. Modeling and analysis of energy harvesting nodes in wireless sensor networks. In: *Proceedings of the 46th Annual Allerton Conference*, Urbana-Champaign, 23-26 September 2008.

Shi L, Johansson KH and Qiu L. Time and event-based sensor scheduling for networks with limited communication resources. *IFAC World Congress* 2011; 44(1): 13263-13268.

Tang S and Tan L. Reward rate maximization and optimal transmission policy of EH device with temporal death in EH-WSNs. *IEEE Trans Wireless Commun* 2017; 16(2): 1157-1167.

Ulukus S, Yener A, Erkip E, et al. Energy harvesting wireless communications: a review of recent advances. *IEEE J Sel Areas Commun* 2015; 33(3): 360-381.

Uysal-Biyikoglu E, Prabhakar B and Gamal AE. Energy-efficient packet transmission over a wireless link. *IEEE/ACM Trans Networking* 2002; 10(4): 487-499.

Wu J, Jia QS, Johansson KH, et al. Event-based sensor data scheduling: trade-off between communication rate and estimation quality. *IEEE Trans Autom Control* 2013; 58(4): 2013.

You K, Xie L and Song S. Asymptotically optimal parameter estimation with scheduled measurements. *IEEE Trans Signal Process* 2013; 61(14): 3521-3531.

CHAPTER 5

Optimal Stochastic Data Collection and Transmission Policy for Self-Sustaining SHM Systems Guiding Asset Management

5.1 Introduction

The proliferation of low-cost and high-performing sensors has led to sensing in the built environment across a broad range of civil engineering applications. The applicability of such sensing technologies has been profoundly impacted by advancements in wireless communication and cloud computing, which have enabled the implementation of wireless sensing networks (WSNs) in structural, transportation, geotechnical, and environmental systems, to name a few. Within the civil engineering domain, the research community has widely embraced the use of long-term monitoring data for use in structural health monitoring (SHM). SHM systems using WSNs emerged in the mid-1990’s and have been growing in popularity as a lower cost and easily deployable alternative to traditional wired sensing systems (Kane et al. 2014). Despite these advances, the availability of energy remains a bottleneck in SHM WSNs that rely on energy harvesting (EH) from the environment and can hinder their continuous and reliable operation.

Due to the limited and uncertain nature of EH, WSNs used for SHM have historically employed deterministic schedule-based or “transmit-all” sensing strategies to inform data collection protocol (see, for example, Hou et al. 2019, Kim et al. 2007, O’Connor et al. 2017, Flanigan et al. 2020a). In the context of this chapter, a deterministic schedule-based policy refers to a sensing architecture that transmits all data collected periodically on a pre-defined and regular schedule, where T_{ON} denotes the fixed monitoring period length and T_{OFF} denotes the fixed time between monitoring periods. A transmit-all policy refers to the attempted transmission of all measurements (this is equivalent to a schedule-based policy where $T_{OFF} = 0$). These two data collection strategies have been used almost exclusively within the field of SHM because their implementation is straightforward and because the statistical properties of a stationary ergodic process can be deduced from realizations of the process over a sufficiently long period of time—meaning all data that is not transmitted is missing at random. Despite their widespread implementation, schedule-based and transmit-all sensing strategies are suboptimal because they do not account for resource constraints, hardware constraints, the stochastic nature of energy and event arrivals, or the value of data. For a schedule-based sampling scheme, if $\frac{T_{ON}}{T_{OFF}}$ is too large, then the sensing architecture is at risk of rapidly and repeatedly running out of energy, resulting in temporal death. Here, temporal death refers to the rejection of all incoming data and the loss of any data in the data buffer that has not yet been transmitted when a sensing architecture's energy buffer is fully depleted. If $\frac{T_{ON}}{T_{OFF}}$ is too small, then the sensing architecture will require an excessively long period of time to collect enough data to sufficiently characterize the measured process' statistical parameters—especially the parameters governing the extreme value behavior.

This is problematic because the quality of the estimate of a measured process' parameters is critical to accurately assessing the probability of failure of a monitored asset using reliability methods; parameter estimates should be updated as frequently as possible to track structural condition. Reliability methods account for uncertainty in an asset's capacity and demand and quantify the reliability index, β^r, which is a scalar measure of the safety of a system or component and directly related to the probability of failure (Nikolaidis et al. 2004). An asset with a higher reliability index value has a higher level of safety (i.e., lower probability of failure) than an asset with a lower reliability index value. Reliability methods have matured in civil engineering applications over the past several decades and are playing an increasingly critical role in bridging the gap between SHM and decision-making practices. For example, Estes and Frangopol (Estes and Frangopol 1991) and Frangopol and Estes (Frangopol and Estes 1997) use reliability methods to develop lifetime bridge maintenance strategies that inform optimal repair actions on deteriorating bridges. Additionally, in Chapter 3, reliability methods are used to explicitly link bridge response data collected from a long-term monitoring system to condition ratings, which are the starting point for infrastructure asset management decisions made in practice today. Due to conservatism in structural design—for which the designed probability of failure is very low—the reliability index is particularly sensitive to the quality of the parameter estimates governing extreme value behavior; the demand will only exceed the capacity when realizations of the structural response are near the distribution's extremum. Consequently, there is a need for an automated data collection method that produces a high-quality estimate of a component or system's reliability index by minimizing the variance of the measured process' parameter component estimates under energy and hardware constraints imposed by EH WSNs.

Chapter 4 presents the derivation of an optimal data collection and transmission policy for remote parameter estimation in wireless sensing architectures under resource constraints. In this work, the author proposes an event-based policy and derive a single optimal threshold value, τ_s^* (with * denoting optimal), that maximizes the collected data's average reward rate such that, given a measured process $f(y; \boldsymbol{\theta}_0)$ with true (but unknown) parameters $\boldsymbol{\theta}_0$ (where, in general, $f(y; \boldsymbol{\theta})$ denotes the probability density function (PDF) of a random variable Y parameterized by $\boldsymbol{\theta} \in \Theta$ in the parameter space $\Theta \subset \mathbb{R}^p$), candidate data y is stored and transmitted if and only if $y \geq \tau_s^*$. This work considers constraints imposed by the WSN architecture's energy and data buffer sizes, stochastic models of energy and event arrivals, the value of data, and temporal death. Because the proposed policy is governed by a single optimal threshold value, its implementation does not require knowledge of the state of charge (SoC). Additionally, the policy places no restrictions on the size of the WSN architecture's data storage buffer. Since only a subset of the candidate data can be transmitted when a WSN relies on EH, Chapter 4 introduces a modified likelihood function that accounts for the missingness of data that result from the implementation of the optimal policy. The author shows that under regulatory conditions, the modified MLE retains desirable properties that are characteristic of the standard full-information MLE assuming unlimited energy (i.e., all data is transmitted). Implementation of the optimal threshold produces the best possible estimate of the process parameters given the system constraints.

This chapter draws on the theoretical basis presented in Chapter 4 and extends this framework to an SHM application in which the WSN relies on solar energy harvesting. The Telegraph Road Bridge (TRB) is a steel girder highway bridge located in Monroe, MI that has been continuously monitored since 2011. Since monitoring initiated nearly a decade ago, the SHM system on the TRB has been the subject of numerous research efforts (Flanigan et al. 2020b, O'Connor et al. 2017, Jeong et al. 2019). This chapter focuses its attention on one of the bridge's monitoring subsystems that measures the strain response of a pin-and-hanger assembly. The pin-and-hanger assembly is an integral structural element within the bridge's design, and also a primary concern for the bridge owner due to the well-documented history of bridge failures associated with corroding pin-and-hanger assemblies in the past (Fisher and Yuceoglu 1981, Juntunen 1998, NTSB 1984, South et al. 1992). This chapter extends the optimal data collection and transmission policy proposed in Chapter 4 to estimate the parameters characterizing the net-section strain response of one of the TRB hanger plates. The estimated parameters are then used as inputs to a first-order reliability method (FORM) analysis to calculate the reliability index (i.e., probability of failure) associated with the net-section yielding limit state function; the American Association of State Highway and Transportation Officials load-and-resistance factor design (AASHTO-LRFD) of pin-connected assemblies assumes net-section yielding to be the dominant failure mechanism (AASHTO 2007).

In order to illustrate the gains achieved using the optimal policy on the TRB to characterize the net-section strain response, the author simulated the data collection and transmission process over the monitoring period assuming 1) the optimal policy is implemented, 2) the best schedule-based policy is implemented, and 3) the transmit-all policy is implemented over a one-year period. Strain response data measured at the TRB hanger plate's net-section area is used in each simulation to reflect the actual in-service bridge response. As expected, implementation of the derived optimal stochastic data collection and transmission policy for the self-sustaining SHM system on the TRB minimizes the variance of the estimated component parameters. This enables the bridge owner to increase the frequency with which structural condition is tracked and asset management decisions are made based on the reliability index (using, for example the reliability-based decision-making practice outlined in Chapter 3) without compromising accuracy, as compared to existing schedule-based and transmit-all data collection methods. While the proposed framework is applied to a specific type of element on the TRB, the proposed methodology extends naturally to any SHM application.

5.1.1 Chapter Outline

The remainder of this chapter is outlined as follows. Section 5.2 introduces the TRB and provides an overview of pin-and-hanger assemblies, including the well-documented failure mechanisms that ultimately led to their discontinued use. These deficiencies motivate the TRB's instrumentation plan and data collection program, which are also detailed in Section 5.2 and focus on tracking the safety of the pin-and-hanger assembly with respect to the net-section yielding failure limit state. The goal of the monitoring system is to collect maximum strain response measurements at the net-section area due to passing trucks and characterize the random variable describing the long-term response data, ε_{H4}. These parameter estimates should be updated as frequently as possible to track structural condition using reliability methods. To this end, Section 5.3 models the SHM system's data collection and transmission processes as an energy renewal system and uses Weigh-in-Motion (WIM) data, weather information, hardware specifications, and monitoring response data to define the system parameters. Using this modeled system, Section 5.3 derives an optimal data collection and transmission policy that maximizes the transmission rate given constraints imposed by the WSN node's energy and data buffer sizes, stochastic models of the incoming energy and event arrivals, the value of data, and temporal death. Section 5.3 then introduces a modified likelihood function used for remote

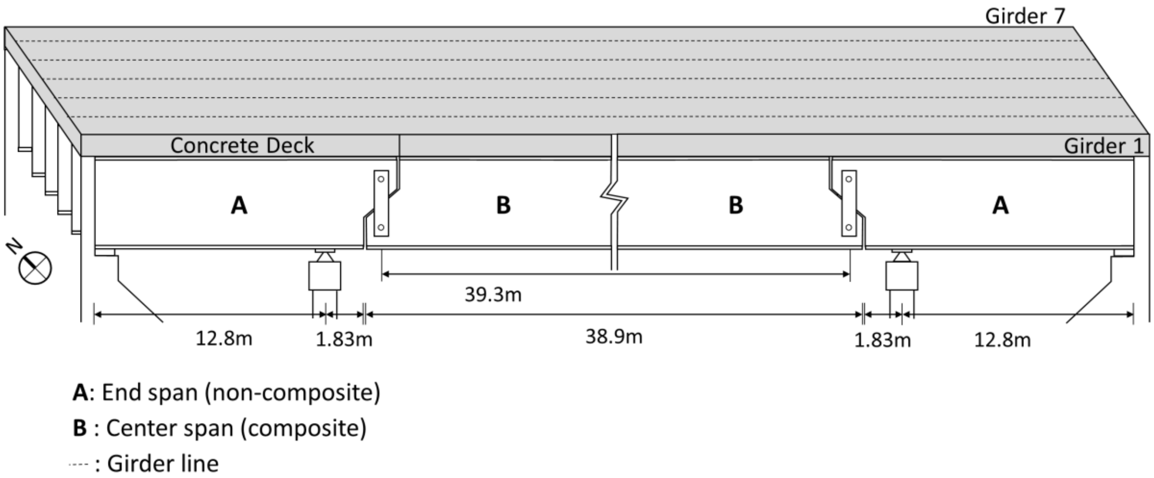
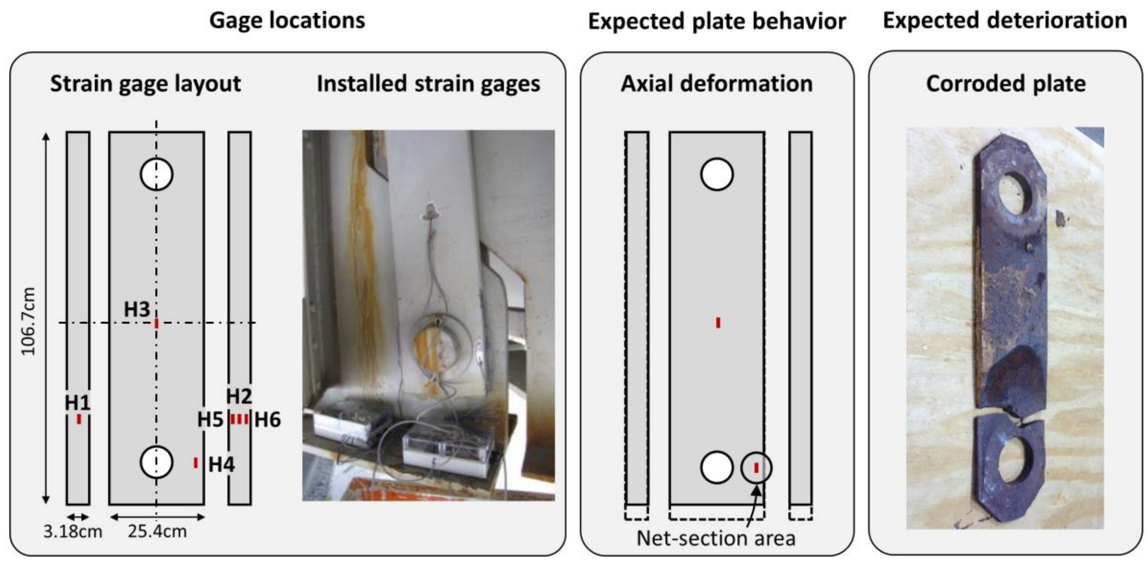


Figure 5-1. Girder elevation including primary span dimensions.

parameter estimation. In this section, we prove that the stochastic data collection and transmission policy minimizes the variance of estimated component parameters of the measured process (e.g., net-section strain, ε_{H4}), which serve as measured inputs to reliability-based decision-making practices. Section 5 provides an overview of reliability methods and illustrates how reliability methods can be used to monitor the net-section stress of the in-service TRB pin-and-hanger assembly based on the measured random variable describing the long-term response data, ε_{H4}. The reliability index is tracked based on transmitted data collected according to simulation of the optimal, schedule-based, and transmit-all policies on the TRB for the one-year monitoring period. These results illustrate the gains achieved using the optimal policy on the TRB and confirm that implementation of the optimal policy enables bridge owners to track structural condition as frequently as possible without compromising accuracy. Finally, Section 5.6 presents conclusions and introduces key directions for future work.



(a) (b) (c)

Figure 5-2. (a) Strain gage layout including major dimensions and identification of corresponding strain gages installed on an instrumented TRB hanger plate; (b) expected behavior of the plate response (i.e. axial deformation); (c) the inside face of a hanger plate with severe corrosion-induced section loss.

5.2 Telegraph Road Bridge

The proposed optimal stochastic scheduling policy for SHM systems under resource constraints is applied to the TRB, which carries two lanes of highway traffic north along Interstate 275. The TRB is a standard steel girder highway bridge (skewed by 33 degrees) that was constructed in 1973. The bridge has three spans with a 20 cm concrete deck, seven girders, and 28 pin-and-hanger assemblies, which are integral structural elements in its design. As shown in Figure 5-1, each of the TRB's seven girders comprises two sets of hanger plates (there is a hanger plate on the north and south faces of each girder) to suspend the center span between the two end spans. The two end spans are each supported by abutment structures and interior piers. The TRB's girder elevation and major span dimensions are included in Figure 5-1. The TRB is representative of a large portion of the Michigan Department of Transportation's (MDOT) highway bridge inventory. Of the

2,914 steel beam bridges owned by MDOT, around 25 percent have pin-and-hanger assemblies in their design (Jansson 2008). While the design and construction of pin-and-hanger bridges was discontinued in Michigan around 1983, many in-service bridges—such as the TRB—remain operational.

5.2.1 Pin-and-Hanger Detail

A number of design advantages of pin-and-hanger assemblies led to an increase in the frequency of their use in bridge design in the United States from 1940 to 1980. When integrated into bridge designs, pin-and-hanger connections (Figure 5-2(a)) reduce moments at supports, allow for thermal expansion between spans, and decrease corrosion-induced damage in bearing elements by enabling expansion joints to be located away from the piers and abutments. Despite these gains, over time it became apparent that pin-and-hanger assemblies exhibit a number of deficiencies as they age, which can jeopardize the structural safety of the entire structure.

Deterioration of pin-and-hanger assemblies can initiate when salt and water leak through expansion joints and corrode the pin-hanger connection. Dirt and sand behind the plate can absorb moisture, thereby accelerating corrosion, especially at the bottom pin. Corrosion-induced changes in the assembly boundary conditions and net-section loss can introduce stresses around the pins at the net-section area, leading to failure (Juntunen 1998; Nowak and Zhou 1990). Since pin-connected assemblies are designed based on the assumption that they are purely tensile elements whose primary failure mechanism is governed by net-section yielding (AASHTO 2007), this chapter focuses its attention on deterioration that manifests as corrosion-induced section loss at the net-section area, which is the section of the hanger plate with the least area (i.e., the cross-section of the hanger plate along the centroid of each pin as identified in Figure 5-2(b)). Because deterioration typically initiates between the plate and girder, it can be difficult to identify the onset of corrosion-induced pitting through visual inspection. For example, the back side of a hanger plate taken out of service is shown in Figure 5-2(c). It is evident that a significant loss of section caused by corrosion has occurred near the lower pin at the net-section area and propagates toward the gross-section area. Because nationally mandated visual inspections are infrequent (e.g., biennial) and may not uncover underlying deterioration, this assembly stands to benefit greatly from an automated monitoring system that can use measured response data to track structural performance using reliability methods.

5.2.2 Telegraph Road Bridge Instrumentation Plan and Data Collection Program

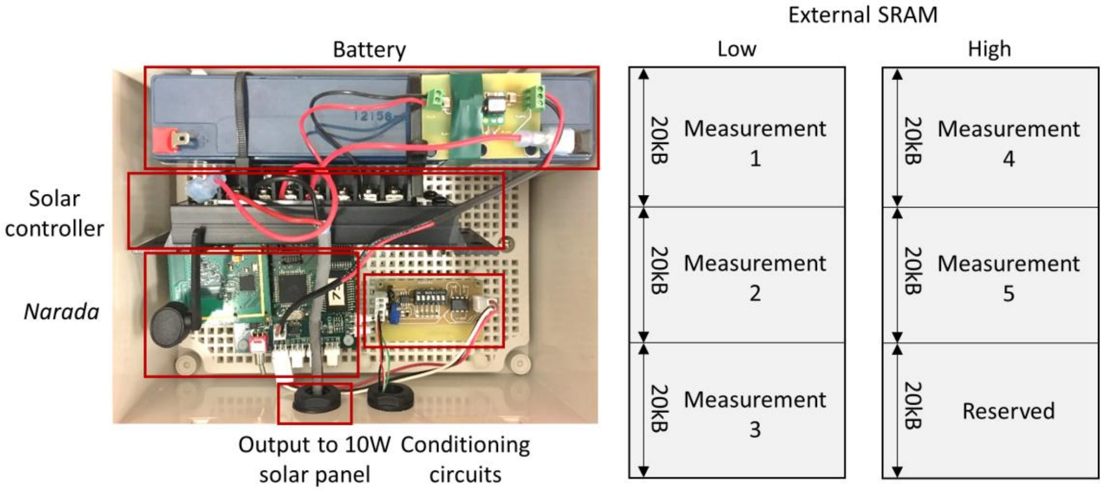
Strain gages are instrumented on a hanger plate located at the east end of the center span's Girder 2 (Figure 5-1) to monitor stresses that could be induced by damage due to long-term deterioration and corresponding changes in plate boundary conditions. Six Tokyo Sokki 120Ω uni-axial strain gages are installed on the hanger plate in order to monitor the plate's response corresponding to limit state functions associated with axial, in-plane bending, and out-of-plane bending stresses, and torsional forces applied to the hanger plate pins (Figure 5-2(a)) (O'Connor 2017). Because this chapter is investigating structural condition associated with net-section stress, only the strain measured at location H4 is considered herein. However, interested readers can refer to Chapter 3 for an extended discussion of how all six strain gages comprising the hanger plate's monitoring subsystem can be used to track structural performance with respect to other deterioration mechanisms that result from unexpected in-plane and out-of-plane bending. The strain gages are sampled using an analog-to-digital converter (ADC) with 16-bit resolution at 100 Hz for 100 seconds every two hours using the *Narada* wireless sensor (Swartz et al. 2005) from June 2016 to June 2017. Each *Narada* node—pictured in Figure 5-3(a)—contains a solar controller connected to a 10-W solar panel and a rechargeable battery for solar harvesting and energy storage, respectively. Monitoring data provides quantitative information about the loading history and structural response of the plate. For example, Figure 5-4 provides a 100-second sample of the time series data collected at strain gage location H4. The maximum peak strain corresponding to a truck crossing the bridge is highlighted. A histogram comprising all truck-induced maximum strain values at location H4 measured over the entire monitoring period from June 2016 to June 2017 is shown in Figure 5-5. The corresponding Generalized Pareto distribution (GPD) fit is superimposed on the histogram, where the GPD has a probability density function (PDF) of the following form

$$f(y;\boldsymbol{\theta}) = \frac{1}{\sigma}\cdot\left(1+\xi\cdot\frac{y}{\sigma}\right)^{-\frac{1}{\xi}-1} \tag{5.1}$$

Here, ξ is the shape parameter and σ is the scale parameter. We let $\boldsymbol{\theta}_0 = [\xi_0, \sigma_0]^T$ denote the “true” process parameters that are informed by data collected over the one-year monitoring period, where $\xi_0 = 0.311$ and $\sigma_0 = 2.80$. In the remaining sections, we show that over the fixed monitoring period, implementation of the proposed optimal policy leads to minimum variance parameter estimates given the system constraints.

5.3 Problem Formulation

The proposed energy-aware stochastic scheduling policy for remote parameter estimation discussed herein comprises a transmission subsystem and a remote parameter estimator. Within the transmission subsystem, a wireless sensing node measures the peak strain response of each passing truck. The peak strain responses induced by passing trucks



(a) (b)

Figure 5-3. (a) *Narada* wireless sensing node; (b) allocation of external memory for storing measured structural response data.

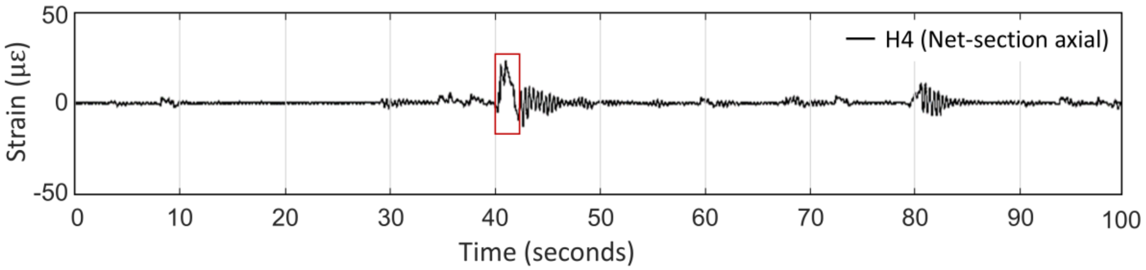


Figure 5-4. Time series data at strain gage location H4 collected during an event on May 28, 2017.

constitute the measured process, which follows a GPD and has true parameters denoted by the 2-dimensional vector $\boldsymbol{\theta}_0$. A binary value $v \in \{0,1\}$ is assigned to each realization of the measured process, y, through the function $v = h_s(y)$ such that

$$v = h_s(y) = \begin{cases} 0, & \text{if } y < \tau_s^* \\ 1, & \text{if } y \geq \tau_s^* \end{cases} \tag{5.2}$$

If $v = 0$ the candidate measurement y is discarded and if $v = 1$ the candidate measurement y is stored in external SRAM, regardless of the current SoC of the battery. Once the data buffer is full, all stored data (denoted by the vector $\boldsymbol{\Lambda}_s$) is communicated to the remote

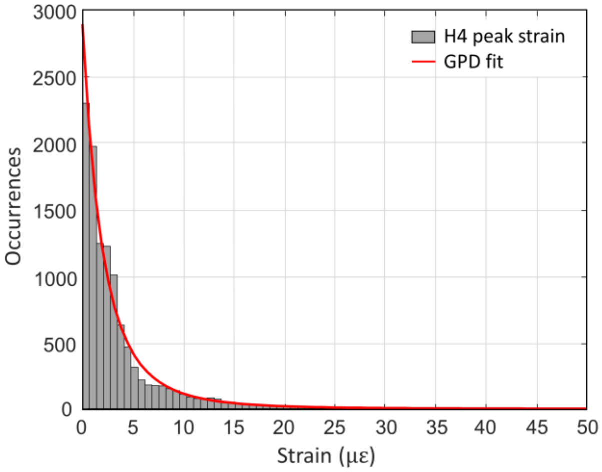


Figure 5-5. GPD fit overlaying the histogram of maximum truck-induced strain responses measured at strain gage location H4 over the one-year monitoring period.

parameter estimator in a single batch transmission. The subscript $s = 1,2,3, \ldots, s_t$ is introduced herein as an index to indicate that data collected and transmitted in batch s, $\boldsymbol{\Lambda}_s$, corresponds to the optimal threshold, τ_s^*, where $s = 1$ represents the first batch transmission after monitoring initiates and $s = s_t$ represents the most recent batch transmission. The proposed model utilizes batch transmission because the energy overhead of packet transmission is relatively high regardless of the payload size (Han et al. 2015). If the sensing node's battery is fully depleted at any point in time, temporal death occurs; the sensing node shuts down, all stored data is discarded, and all candidate data that arrive when the sensing node is powered down are rejected. Upon receiving the collected data, $\boldsymbol{\Lambda}_s$, the remote parameter estimator located on an external server updates the maximum likelihood estimate, $\widehat{\boldsymbol{\theta}}_{s+1}$, recalculates the optimal threshold based on the updated estimate, and returns the updated optimal threshold to the sensing node. Each time the parameter estimates are updated, the reliability index corresponding to the limit state governed by net-section yielding is updated based on the measured process at strain gage location H4

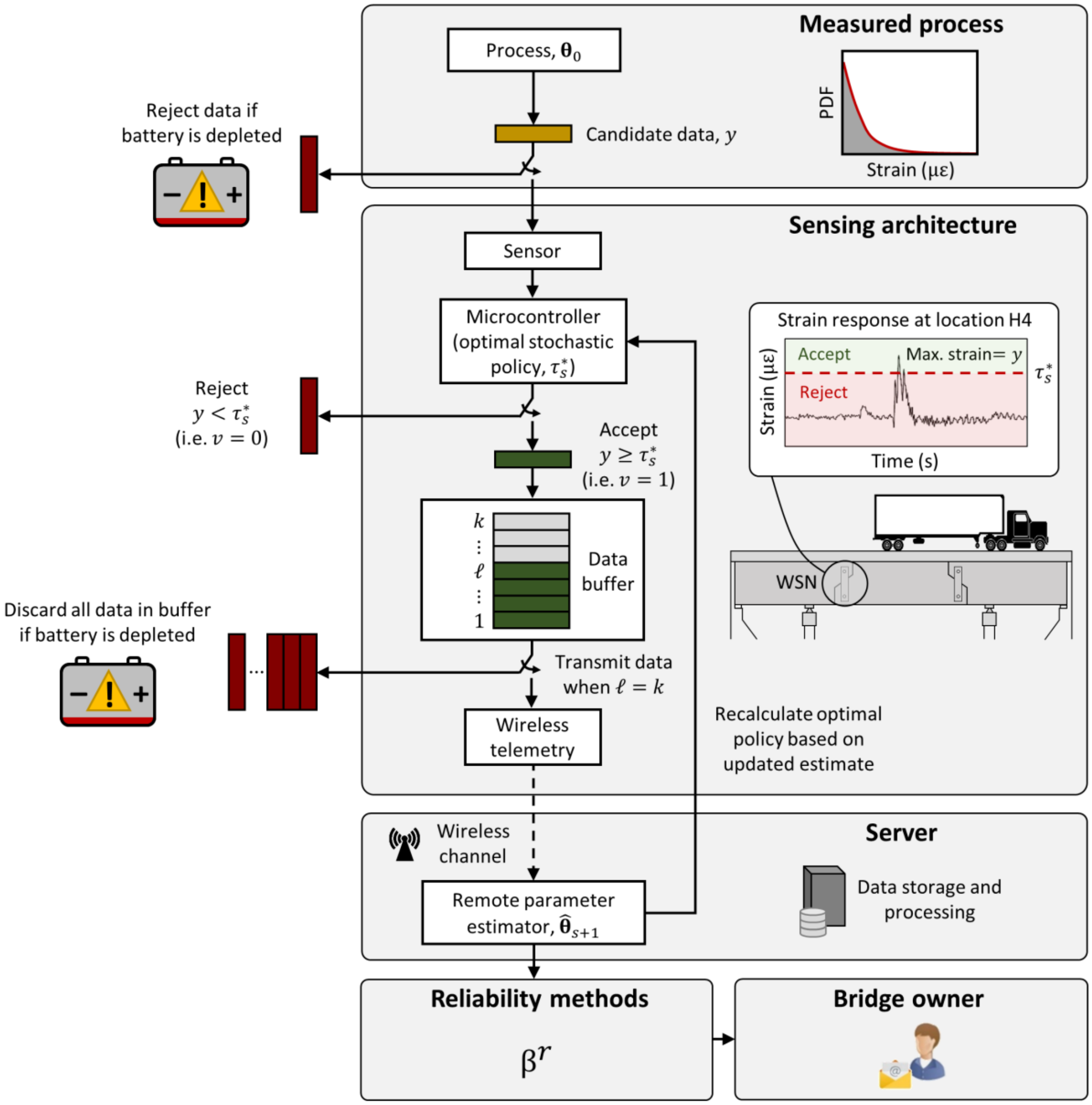


Figure 5-6. Event-based parameter estimation framework for applications in SHM.

(this will be described in greater detail in Section 5.4.1). Figure 5-6 illustrates this event-based parameter estimation framework for SHM applications and highlights the process for either storing and transmitting data, storing and discarding data, or rejecting data according to the optimal policy. Also shown in Figure 5-6, the bridge owner and field inspectors have access to the continuously monitored reliability index, which is a scalar metric for the safety of the asset. Given this framework, the goal is to derive the unique

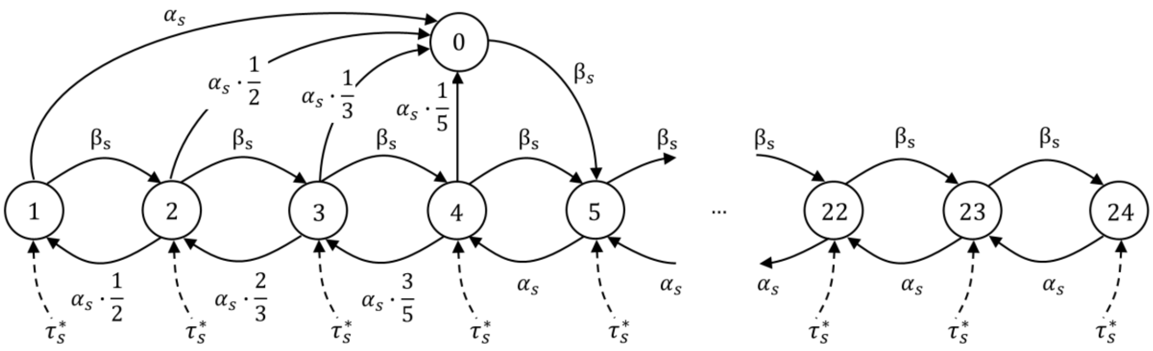


Figure 5-7. Markov chain representing the energy renewal system.

optimal threshold value, τ_s^*, that maximizes the collected data's average transmission rate given a WSN architecture's energy and data buffer sizes, stochastic models of energy and event arrivals, the value of data, and consideration of temporal death. Here, maximizing the expected transmission rate is equivalent to maximizing the expected reward rate because the reward (or value) of each transmitted measurement is one (recall Equation (5-2)).

5.3.1 Transmission Subsystem Model

Recall from Section 5.2.2 that the wireless sensing node has a replenishable, finite-size battery, as well as a data storage buffer. While the system has a continuous state space (i.e., remaining energy) and a continuous parameter space (i.e., time), the state space is approximately modeled as a discrete state space. Let $X(t) \in \{0,1,\ldots,N+k-1\}$ be a finite continuous-time Markov chain with $N+k$ states as shown in Figure 5-7, where k is the number of candidate measurements the storage buffer can hold and ℓ denotes the number of values stored in the data buffer at time t. Since there is 120kB of available external SRAM (ExSRAM) on the *Narada's* 8-bit embedded processor, $k=5$ because five 100-second measurement periods can be stored before transmission is triggered

(Figure 5-3(b)). Consistent with the system requirements on the TRB, the 100-second measurement period is sampled at 100 Hz and each measurement requires two bytes of storage in the ExSRAM. Rather than only storing and transmitting the maximum strain response to each truck at the sensor location (when $y \geq \tau_s^*$), the 50 seconds leading up to and following the peak strain response are also stored and transmitted. This allows for the time series data to be used in other post-processing applications. The remaining 20kB of storage in the ExSRAM is reserved for storing header information and for aiding in the execution of on-board processing functions. For $X(t) = 4 \leq n \leq 24$, the state at time t, $X(t) = n$, represents the remaining energy which can support the transmission of $n - 4$ data packets. Given the maximum capacity of the finite-size battery is E_{max} and the amount of energy required to transmit the full data buffer of 5 values is E_T, then $N = \left\lfloor \frac{E_{max} \cdot 5}{E_T} \right\rfloor$ where $\lfloor x \rfloor$ is an operator rounding x to the nearest integer less than x for $x > 0$. Here, E_{max} and E_T are dependent on a number of factors that are unique to the implemented SHM system, such as current consumption of attached transducers, operational overhead of the embedded sensing architecture, and battery capacity. For the sensing system considered in this chapter, $N = 20$. Interested readers can refer to Chapter 4 for a detailed description of the Markov chain representing the energy renewal system.

Trucks arrive as a memoryless Poisson process with rate λ_s (trucks per minute). The energy storage buffer recharges based on harvested energy which arrives as a memoryless Poisson process with rate β_s (energy level per minute). If $X(t) = 5 \leq n \leq 24 \; \forall \ell$ or $X(t) = 2 \leq n < 5$ for $\ell < k$, the state transitions from n to $n - 1$ when candidate data is collected and stored. Here, the transition rate from state n to state $n - 1$ is $\alpha_s = \lambda_s \cdot P[y \geq \tau_s^*] = \lambda_s \cdot \bar{F}(\tau_s^*; \widehat{\boldsymbol{\theta}}_s)$ where $\bar{F}(\cdot)$ denotes the complementary cumulative distribution function (CDF). The expected reward corresponding to the optimal threshold, τ_s^*, is $R_s = P[y \geq \tau_s^*]$. Formulating the reward process as a discrete-time Markov chain in which state transitions are made at infinitesimal time steps, Δ, the expected reward during each time step is $\lambda_s \cdot \Delta \cdot R_s$ for states $1 \leq n < 4$ and $4 < n \leq 24$, and $\lambda_s \cdot \Delta \cdot \frac{4}{5} \cdot R_s$ for state $n = 4$, where $\lambda_s \cdot \Delta$ is the probability that an event will occur during the next time step, Δ. The time step, Δ, is assumed to be so small that at most one event can occur during the time step. A cost is incurred when the energy storage buffer is fully depleted and enters state $n = 0$ because all stored data is discarded when the sensing node shuts down. Based on the method of discretization of the finite-size battery—and consistent with the Markov chain in Figure 5-7—a state transition into state $n = 0$ can only occur when $\ell = 5$, meaning that if the sensing architecture shuts down because there is not sufficient energy to complete the batch transmission then it is known that 5 messages are lost. When this happens, the expected amount of time spent in state $n = 0$ is $\frac{1}{\beta_s}$ (or $\frac{1}{\beta_s \cdot \Delta}$ time steps). If we denote the expected cost incurred during a single time step, Δ, as C_s, then the expected total cost incurred in state $n = 0$ over time period $\frac{1}{\beta_s}$ is $\frac{1}{\beta_s \cdot \Delta} \cdot C_s$. The total expected cost incurred during the sojourn time in state $n = 0$ should equal the expected loss of the reward that would have been gained by the k discarded values that were discarded when the energy buffer was depleted. As a result,

$$C_s = -5 \cdot \beta_s \cdot \Delta \cdot \frac{R_s}{P[y \geq \tau_s^*]} \tag{5.3}$$

where $\frac{R_s}{P[y \geq \tau_s^*]}$ is the conditional expected value of a transmitted realization of the measured process; C_s is not a function of τ_s^* because $\frac{R_s}{P[y \geq \tau_s^*]} = 1$ given the mapping between candidate data and value in Equation (5-2). Given the expected reward during each time step over states $1 \leq n \leq 24$ and the expected cost incurred during each time step in state $n = 0$, the reward vector is

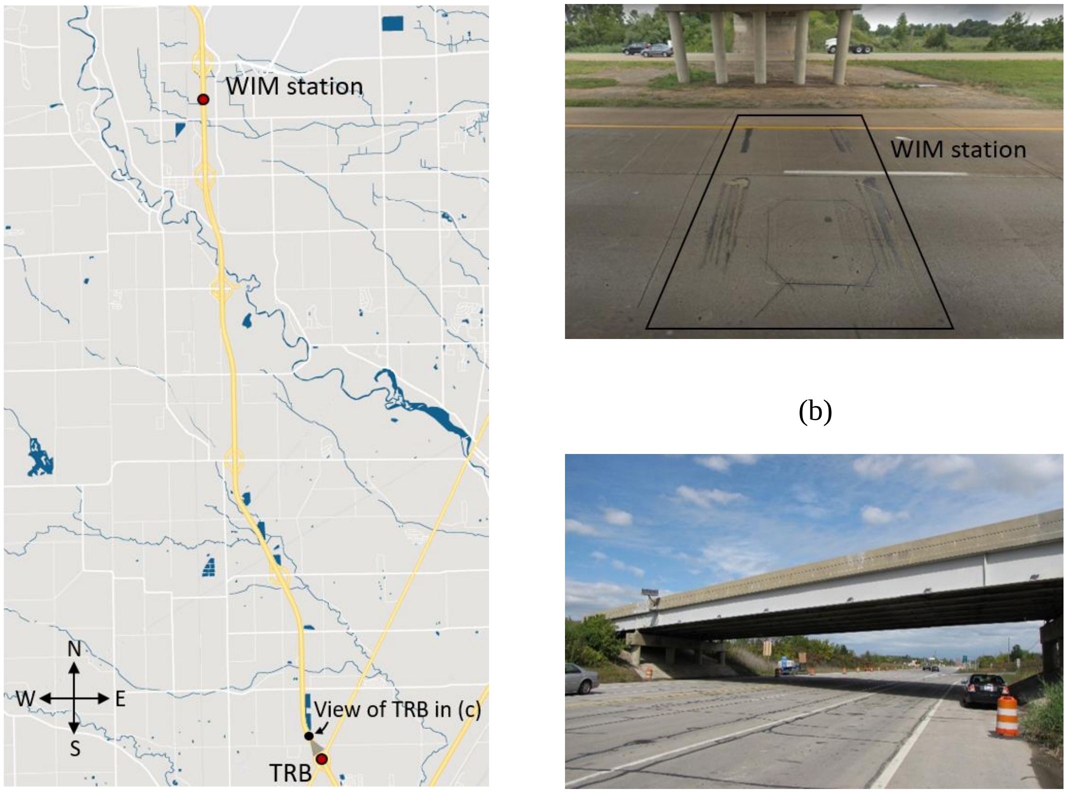
$$\boldsymbol{r}_{s(N+k)\text{x}1} = \lambda_s \cdot \Delta \cdot \left[\frac{C_s}{\lambda_s \cdot \Delta},\ \ R_s \cdot \boldsymbol{e}_{1\text{x}(k-2)},\ \ \frac{4}{5} \cdot R_s,\ \ R_s \cdot \boldsymbol{e}_{1\text{x}N}\right]^T \tag{5.4}$$

where $\boldsymbol{e}$ is a vector of ones. The transition matrix corresponding to the Markov chain in Figure 5-7 is

$$\boldsymbol{P}_{s_{25\text{x}25}} = \begin{bmatrix} \gamma_0 & 0 & 0 & 0 & 0 & \beta_s & 0 & \cdots & 0 & 0 & 0 & 0 \\ \alpha_s & \gamma_1 & \beta_s & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \alpha_s \cdot \frac{1}{2} & \alpha_s \cdot \frac{1}{2} & \gamma_2 & \beta_s & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \alpha_s \cdot \frac{1}{3} & 0 & \alpha_s \cdot \frac{2}{3} & \gamma_3 & \beta_s & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \alpha_s \cdot \frac{1}{5} & 0 & 0 & \alpha_s \cdot \frac{3}{5} & \gamma_4 & \beta_s & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha_s & \gamma_5 & \beta_s & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_s & \gamma_6 & \ddots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \gamma_{21} & \beta_s & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \alpha_s & \gamma_{22} & \beta_s & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \alpha_s & \gamma_{23} & \beta_s \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \alpha_s & \gamma_{24} \end{bmatrix} \cdot \Delta \tag{5.5}$$

where Δ is sufficiently small such that the self-transitions (i.e., from state n to n) are $P_{n,n} \geq 0$. Here,

$$\gamma_n = \begin{cases} \frac{1}{\Delta} - \beta_s, & \text{for} \quad n = 0 \\ \frac{1}{\Delta} - \beta_s - \alpha_s, & \text{for} \quad n = 1 \\ \frac{1}{\Delta} - \beta_s - \alpha_s, & \text{for} \quad 2 \leq n < 4 \\ \frac{1}{\Delta} - \beta_s - \alpha_s \cdot \frac{4}{5}, & \text{for} \quad n = 4 \\ \frac{1}{\Delta} - \beta_s - \alpha_s, & \text{for} \quad 5 \leq n \leq 23 \\ \frac{1}{\Delta} - \alpha_s, & \text{for} \quad n = 24 \end{cases}$$



(b)

(a)

(c)

Figure 5-8. (a) Locations of the TRB and WIM station along Interstate 275; (b) WIM station embedded in roadway; (c) view of the instrumented TRB.

When the transmission subsystem is implemented on a sensing architecture on an in-service structure, the recharge and truck arrival rates can be obtained directly from the measured bridge response (i.e., truck arrival) and environmental conditions (i.e., recharge rate). However, since we are using historical data collected from June 2016 to June 2017 to illustrate the gains that are achieved by implementing the proposed optimal policy, we do not have measured recharge and event arrival rates during the monitoring period. In the following three subsections, we model the recharge and event arrival rates based on continuous WIM station data collected along Interstate 275 to the north of the TRB as well as surface weather observations recorded by the National Oceanic and Atmospheric Administration (NOAA) at a nearby station.

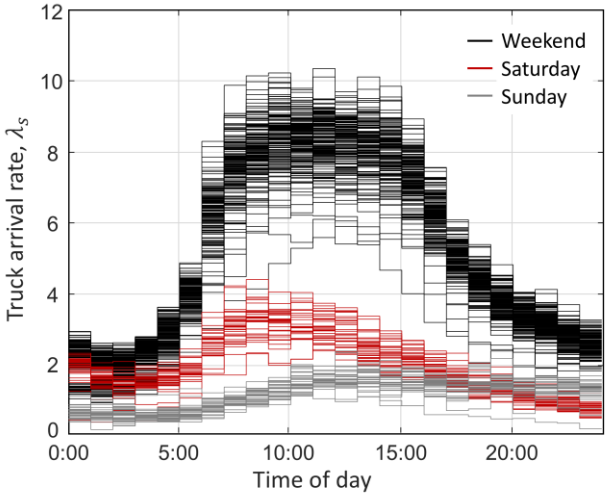


Figure 5-9. Hourly truck arrival rate measured at the WIM station during the one-year monitoring period.

5.3.1.1 Characterizing the Event Arrival Rate

The truck arrival rate (trucks per minute) is calculated based on data collected from an MDOT WIM station located at Pennsylvania Road along Interstate 275 to the north of the TRB in Romulus, MI (Figure 5-8). WIM stations are used by transportation officials to monitor the movement of freight and to track adherence to freight weight limits (McCall and Vodrazka 1997). WIM stations are an attractive alternative to traditional weigh stations because they do not require vehicles to exit the highway; WIM stations are embedded into the roadway (Figure 5-8(b)) and measure key attributes such as each truck's load profile (e.g., gross weight, axle spacing, weight carried by each axle), vehicle speed, and time. There are no major on or off ramps between the WIM station and TRB; according to Hou et al. (Hou et al. 2020) around 70% of trucks that pass over the TRB also pass over the WIM station. Consequently, we assume that the truck arrival rate recorded at the WIM station is representative of the truck arrival rate at the TRB. Figure 5-9 shows the hourly truck arrival rate measured at the WIM station during the monitoring period from June

Table 5-1. Look-up table to calculate the recharge rate based on hourly weather and sky condition information.

		Sky conditions					
		Clear	Sky clear	Few	Scattered	Broken	Overcast
Weather type	None	0.6	0.6	$0.6 \cdot \mathcal{U}(0.85,0.95)$	$0.6 \cdot \mathcal{U}(0.85,0.95)$	$0.6 \cdot \mathcal{U}(0.80,0.90)$	$0.6 \cdot \mathcal{U}(0.70,0.80)$
		0.02	0.02	$0.02 \cdot \mathcal{U}(0.85,0.95)$	$0.02 \cdot \mathcal{U}(0.85,0.95)$	$0.02 \cdot \mathcal{U}(0.80,0.90)$	$0.02 \cdot \mathcal{U}(0.70,0.80)$
	Rain	X	X	$0.6 \cdot \mathcal{U}(0.085,0.875)$	$0.6 \cdot \mathcal{U}(0.085,0.875)$	$0.6 \cdot \mathcal{U}(0.085,0.875)$	$0.6 \cdot \mathcal{U}(0.085,0.875)$
		X	X	$0.02 \cdot \mathcal{U}(0.085,0.875)$	$0.02 \cdot \mathcal{U}(0.085,0.875)$	$0.02 \cdot \mathcal{U}(0.085,0.875)$	$0.02 \cdot \mathcal{U}(0.085,0.875)$
	Snow	X	X	$0.6 \cdot \mathcal{U}(0.085,0.875)$	$0.6 \cdot \mathcal{U}(0.085,0.875)$	$0.6 \cdot \mathcal{U}(0.085,0.875)$	$0.6 \cdot \mathcal{U}(0.085,0.875)$
		X	X	$0.02 \cdot \mathcal{U}(0.085,0.875)$	$0.02 \cdot \mathcal{U}(0.085,0.875)$	$0.02 \cdot \mathcal{U}(0.085,0.875)$	$0.02 \cdot \mathcal{U}(0.085,0.875)$

☐ : Day
▨ : Night

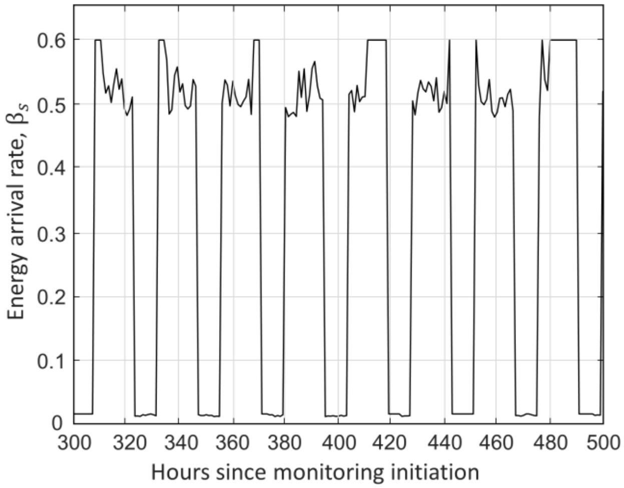


Figure 5-10. Hourly recharge rate based on surface weather observations during the one-year monitoring period.

2016 to June 2017. Figure 5-9 illustrates that there is a strong temporal correlation between truck arrival rates and the time of day, with three distinct profiles occurring on weekdays, Saturdays, and Sundays. While this point process is generally inhomogeneous, it behaves as a homogeneous Poisson process over sufficiently short intervals. Consequently, the truck arrival rate is considered constant over one-hour intervals, which is reflected in Figure 5-9.

5.3.1.2 Characterizing the Recharge Rate

Like truck arrivals, the energy recharge rate is highly correlated to the time of day and weather conditions. Consequently, the energy arrival rate (energy level per minute) is calculated based on data collected from a NOAA METAR surface weather observation station in Monroe, MI (NOAA 2020a). A number of types of weather conditions (e.g., none, rain snow), sky conditions (e.g., clear, few, scattered, broken, overcast), and the time of day (e.g., day, night) are defined by NOAA (NOAA 2020b). Given the weather conditions, sky conditions, and time of day for each hour during the monitoring period, the recharge rate is scaled based on Table 5-1 (Ho 2010). In Table 5-1, $\mathcal{U}(a,b)$ denotes a random realization drawn from the uniform distribution bounded by a and b. The solar controller and battery configuration, as well as the wattage of the solar panel used will contribute greatly to the recharge rate. For the energy harvesting system installed on the TRB, the assumed hourly recharge rates based on Table 5-1 are shown in Figure 5-10 for a sample 200-hour period.

5.3.1.3 Characterizing the Prior Maximum Strain Response to Trucks

Implementation of the optimal policy requires a prior estimate, $\widehat{\boldsymbol{\theta}}_s$, of the process parameters characterizing the maximum strain response to passing trucks. When $s = 1$ at monitoring initiation, the prior estimate is selected by the user, whereas the subsequent estimates, $\widehat{\boldsymbol{\theta}}_s$ for $s = 2,3,\dots,s_t$, used to calculate the optimal threshold for period s is equal to the maximum likelihood estimate calculated from the previous period. Chapter 4 shows that the optimal policy is robust against even large amounts of uncertainty in the prior estimate of the process parameters for $s = 1$, since the accuracy of the threshold, τ_s^*, improves with each transmission. The original belief about the measured process parameters at $s = 1$ is selected by recording the maximum strain response measured for each passing truck using a schedule-based policy over a five-hour period, where $T_{ON} = 50$ seconds and $T_{OFF} = 500$ seconds. Given the hourly truck arrival rates and hourly recharge rates over the one-year monitoring period (from Section 5.3.1.1 and Section 5.3.1.2), $T_{ON} = 50$ seconds and $T_{OFF} = 500$ seconds represent the best schedule-based policy. Here, the "best" schedule-based policy refers to the schedule-based sensing policy that results in the highest expected transmission rate over the entire monitoring period. Implementing the schedule-based policy for five hours at the beginning of the monitoring period, the initial GPD parameters corresponding to $\widehat{\boldsymbol{\theta}}_1$ are $\xi_1 = 0.301$ and $\sigma_1 = 2.84$.

5.3.2 Optimal Data Collection and Transmission Policy

Given the recharge rate, β_s, truck arrival rate, λ_s, size of the energy and data buffers, N and k, respectively, a prior estimate of the estimated process parameters at the beginning of period s, $\widehat{\boldsymbol{\theta}}_s$, and the energy renewal system, the goal is to derive an optimal threshold vector, $\boldsymbol{\tau}^*_{s_{24\text{x}1}} = \tau^*_s \cdot \boldsymbol{e}$ that is optimal in the sense that the policy maximizes the recurrent class' average reward rate (i.e., average transmission rate) This entails calculating the value of a single optimal threshold value, τ^*_s, such that regardless of the current state, $X(t) = n$ for states $1 \leq n \leq 24$, it is optimal to transmit data if and only if the candidate value y is greater than or equal to τ^*_s. Note that we are able to evaluate the infinite horizon problem discussed herein because we are considering the average reward rate.

Chapter 4 shows how to use Howard's Policy Improvement algorithm (Howard 1960) to derive the optimal policy that maximizes the average reward rate given the transmission subsystem model in Section 5.3.1. In summary, Howard's Policy Improvement algorithm consists of two primary stages which are applied sequentially and iteratively: the value-determination stage and the policy-improvement stage. We start with an arbitrary policy, $\boldsymbol{\tau}^A = [\tau^A, \tau^A, \ldots, \tau^A]^T$, and calculate the relative gain vector, $\boldsymbol{w}^A$, and steady-state reward, g^A, associated with this arbitrary policy from the following value determination equations,

$$\boldsymbol{w}^A + g^A \cdot \boldsymbol{e} = \boldsymbol{r}_s + \boldsymbol{P}_s \cdot \boldsymbol{w}^A \tag{5.6}$$

We then evaluate an alternate policy, $\boldsymbol{\tau} = [\tau, \tau, \ldots, \tau]^T$, by defining a contraction mapping $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ as (Gedergruen et al. 1978)

$$T(\boldsymbol{\tau})\boldsymbol{w}^A = \boldsymbol{r}_s(\boldsymbol{\tau}) + \boldsymbol{P}_s(\boldsymbol{\tau}) \cdot \boldsymbol{w}^A \tag{5.7}$$

An optimal policy is achieved when, for all policies $\boldsymbol{\tau}$,

$$T(\boldsymbol{\tau})\boldsymbol{w}^* = \boldsymbol{r}_s(\boldsymbol{\tau}) + \boldsymbol{P}_s(\boldsymbol{\tau}) \cdot \boldsymbol{w}^* \leq T(\boldsymbol{\tau}_s^*)\boldsymbol{w}^* = \boldsymbol{r}_s(\boldsymbol{\tau}_s^*) + \boldsymbol{P}_s(\boldsymbol{\tau}_s^*) \cdot \boldsymbol{w}^* \tag{5.8}$$

Following the results of Chapter 4 for the input parameters derived for the SHM system installed on the TRB, the optimal threshold, τ_s^*, is calculated explicitly by solving for the k^{th} component of the relative-gain vector, w_5^A, from the following reduced value determination equations,

$$\begin{bmatrix} \alpha_s + \beta_s & -\beta_s & 0 & 0 & \beta_s \\ -\alpha_s \cdot \frac{1}{2} & \alpha_s + \beta_s & -\beta_s & 0 & \beta_s \\ 0 & -\alpha_s \cdot \frac{1}{3} & \alpha_s + \beta_s & 0 & \beta_s \\ 0 & 0 & 0 & \alpha_s \cdot \frac{4}{5} + \beta_s & 0 \\ 0 & 0 & 0 & -\frac{1}{\sum_{j=0}^{N-1}\left(\frac{\beta_s}{\alpha_s}\right)^j} & \frac{1}{\sum_{j=0}^{N-1}\left(\frac{\beta_s}{\alpha_s}\right)^j} + \frac{\beta_s}{\alpha_s} \end{bmatrix} \cdot \begin{bmatrix} w_1^A \\ w_2^A \\ w_3^A \\ w_4^A \\ w_5^A \end{bmatrix} = \begin{bmatrix} \lambda_s \cdot R_s - \frac{C_s}{\Delta} \\ \lambda_s \cdot R_s - \frac{C_s}{\Delta} \\ \lambda_s \cdot R_s - \frac{C_s}{\Delta} \\ \lambda_s \cdot \frac{4}{5} \cdot R_s - \frac{C_s}{\Delta} \\ \frac{\lambda_s}{\alpha_s} \cdot \left(R_s - \frac{C_s}{\lambda_s \cdot \Delta}\right) \end{bmatrix} \tag{5.9}$$

and then substituting w_5^A into Equation (5-10) to solve for the optimal threshold, where

$$C_s + \beta_s \cdot \Delta \cdot \left(w_5^A\right)' = 0 \tag{5.10}$$

Here, $(\cdot)'$ denotes the derivative with respect to the threshold. As discussed in Chapter 4, Equation (5-10) reflects the implementation of Howard's Policy Improvement algorithm, which is used to determine the necessary and sufficient conditions that must be imposed on the value-determination equations in Equation (5-9) to derive the optimal threshold value. An infinite-horizon policy is used herein over one-hour time intervals, despite the fact that the policy will be implemented and updated over finite-time periods. The infinite and finite-horizon problems are comparable in this problem because there is no cost imposed for the number of events that have occurred—or even a requirement that data be collected at all—meaning that minimal time is not a requirement for this system (as is the case for most finite-horizon problems).

5.3.3 Remote Parameter Estimation Using a Modified Likelihood Function

The optimal policy presented in Section 5.3.2 maximizes the expected transmission rate of candidate strain data measured from the process characterized by the PDF $f(y; \boldsymbol{\theta}_0)$. If there is unlimited energy supplied to the WSN—meaning all candidate data are transmitted—then the standard likelihood function given full information is

$$L(\boldsymbol{\theta}) = \prod_{s=1}^{S_t} \left[\prod_{i=1}^{5} f(\Lambda_{s,i}; \boldsymbol{\theta})\right] \tag{5.11}$$

For the case of WSNs under energy constraints, the information that is collected and sent to the remote parameter estimator during batch transmission s reflects only a subset of candidate data of the measured process. By implementing the optimal policy, τ_s^*, the missingness of non-transmitted data is classified as not missing at random (NMAR) (Arnab 2017) because it depends on the observed values: it is known that data, y, is missing because either 1) the sensing architecture's energy buffer is depleted and unable to measure the arriving message, 2) the sensing architecture measures the data and rejects it because $y < \tau_s^*$, or 3) the message is already stored in the data buffer but discarded before it is transmitted because the energy buffer is fully depleted. Because the rejected and discarded data that is never transmitted to the remote estimator is categorized as NMAR, the missing data is nonignorable and the likelihood function characterizing transmitted data must account for consideration of left- and right-censored data.

Given the optimal threshold, τ_s^*, corresponding to batch transmission s, we introduce the following modified likelihood function that allows us to have a maximizer even when data is censored

$$L(\boldsymbol{\theta}) = \prod_{s=1}^{s_t} \left[\prod_{i=1}^{5} f(\Lambda_{s,i}; \boldsymbol{\theta}) \cdot \prod_{i=1}^{l_{c,s}} F(\tau_s^*; \boldsymbol{\theta}) \cdot \prod_{i=1}^{r_{c,s}} \bar{F}(\tau_s^*; \boldsymbol{\theta}) \right] \tag{5.12}$$

Here, $l_{c,s}$ is the number of candidate data measurements that are left censored and $r_{c,s}$ is the number of candidate data measurements that are right censored, where

$$r_{c,s} = 5 \cdot N_{D,s} + T_s \cdot \lambda_s \cdot \bar{F}(\tau_s^*, \boldsymbol{\theta}) \cdot \left(\pi_0 + \frac{\pi_4}{4}\right) \tag{5.13a}$$

$$l_{c,s} = 5 \cdot \left(1 + N_{D,s}\right) \cdot \frac{F(\tau_s^*, \boldsymbol{\theta})}{\bar{F}(\tau_s^*, \boldsymbol{\theta})} + T_s \cdot \lambda_s \cdot F(\tau_s^*, \boldsymbol{\theta}) \cdot \left(\pi_0 + \frac{\pi_4}{4}\right) \tag{5.13b}$$

In Equation (5-13), $N_{D,s}$ is the number of times the sensing unit runs out of energy during data collection period s, T_s is the amount of time between the start and finish of data collection for batch transmission s, and π_0 and π_4 are the first and k^{th} components of the steady-state vector, $\boldsymbol{\pi}_s$. Chapter 4 proves that for the modified likelihood function characterized by Equation (5-12), the MLE is the maximizer, there exists a Cramer-Rao bound (CRB) on the covariance matrix of the estimator, the MLE is consistent, asymptotically unbiased, and asymptotically normal, and minimum variance of the MLE components is achieved under the optimal data collection and transmission policy given the system constraints.

5.4 Overview of Reliability Methods

The goal of structural reliability is to calculate the reliability index, β^r, which is a scalar measure of safety of a system or component with respect to a failure limit state (Nikolaidis et al. 2004). A limit state function, $G(\boldsymbol{X}) = C - D$, is described by an n-dimensional vector of random variables, $\boldsymbol{X} = [X_1, X_2, \ldots, X_n]^T$, the load effect, D, and the resistance, C. For the single component reliability problem, the probability of failure, $P_f = P(\boldsymbol{X} \in \Omega)$, is defined as

$$P_f = \int_{\Omega} f_{\boldsymbol{X}}(\boldsymbol{X}) d\boldsymbol{X} \tag{5.14}$$

where $f_{\boldsymbol{X}}(\boldsymbol{X})$ is the joint PDF of $\boldsymbol{X}$, and $\Omega \equiv G(\boldsymbol{X}) \leq 0$ is the failure domain. The reliability index, β^r, is defined as the minimum distance from the origin to the failure domain of the limit state function, Ω, in the independent standard normal space (ISNS)—specifically, when the limit state function is linear in the ISNS. When the limit state function is not linear in the ISNS, first-order approximations can be used to estimate β^r, in which case the reliability index is defined as the approximate, rather than absolute, minimum distance. The reliability index, β^r, is a direct indicator of the probability of failure, where

$$P_f \cong \Phi(-\beta^r) \tag{5.15}$$

and $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF) (Nikolaidis et al. 2004). The equality in Equation (5-15) holds if the limit state function is linear in the ISNS.

When no closed-form solution exists for the integral in Equation (5-14), FORM can be used to approximate the probability of failure by linearizing the limit state function at an optimal point in the ISNS (Nikolaidis et al. 2004). Interested readers can refer to Chapter 3 for a complete discussion about FORM as well as the step-by-step procedure for calculating the reliability index.

5.4.1 Monitoring Net-Section Stress of the In-Service TRB Pin-and-Hanger Assembly

The assumed net-section stress limit state function governing the pin-and-hanger assembly's condition ratings is denoted $G(\boldsymbol{X})$, where

$$G(\boldsymbol{X}) = F_y - \frac{DL}{A_n} - \varepsilon_{H4} \cdot E \tag{5.16}$$

Here, the distribution parameters for yield stress, F_y, and elastic modulus, E, are taken from literature (Hess et al. 2002). Under the assumption that the deck and parapet weights are evenly distributed across the seven girders, the dead load, DL, is estimated from engineering design drawings and the coefficient of variation is taken from literature (Nowak 1993). These random variables and their associated properties are detailed in Table 5-2. Because the net-section area, A_n, can be measured directly, its value is assumed to be deterministic in this case study with a value of 40.3 cm^2 for this study. The location parameter, ξ_s, and scale parameter, σ_s, characterizing the net-section strain response at strain gage location H4, ε_{H4}, are calculated and updated after each batch transmission. Implementation of the optimal policy minimizes the variance of these estimated parameters, which leads to a more accurate assessment of the safety of the pin-and-hanger assembly (via the reliability index) over a fixed period of time. Figure 5-11 illustrates the

Table 5-2. Random variable distribution parameters characterizing the limit state function.

Random Variable	Mean	COV	Distribution Type	Source
Yield stress, F_y (MPa)	342	0.0890	Lognormal	[27]
Shear yield stress, F_v (MPa)	205	0.0890	Lognormal	[27]
Elastic modulus, E (GPa)	205	0.0179	Lognormal	[27]
Dead load, DL (kg)	19.5(10^3)	0.100	Normal	Monitoring Data, [8] [27]

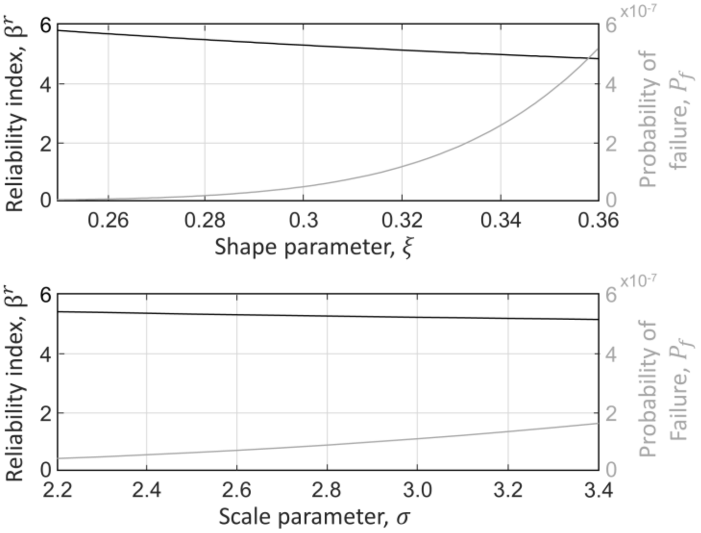


Figure 5-11. Sensitivity of the reliability index and associated probability of failure to variations in the GDP shape and scale parameters.

sensitivity of the reliability index and associated probability of failure to variations in the location and scale parameter estimates. As expected, the reliability index is particularly sensitive to the quality of the parameter estimate governing its extreme value behavior as characterized by the shape parameter.

5.5 Numerical Results

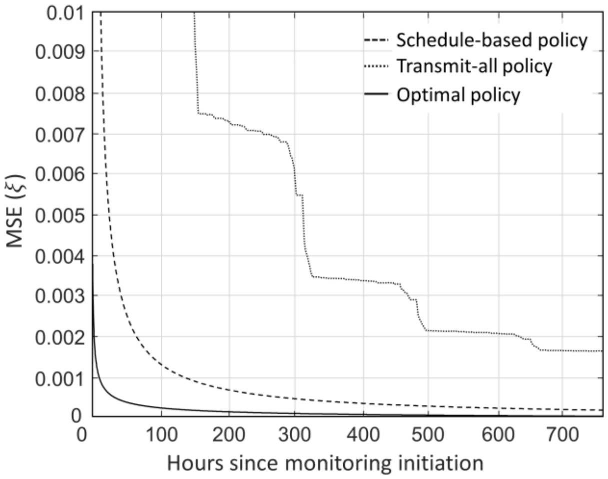
Consider the TRB's energy renewal system modeled by the Markov chain in Figure 5-7. In Section 5.3 we derive the input parameters necessary to obtain and implement the proposed optimal policy on the TRB's long-term SHM system. We provide a summary of these input and model parameters below:

- The total monitoring period runs from June 2016 to June 2017.
- Trucks arrive as a memoryless Poisson process with rate λ_s (trucks per minute). The truck arrival rate is considered constant over one-hour intervals, which are shown in Figure 5-9.
- Energy arrives as a memoryless Poisson process with recharge rate β_s (energy level per minute). The recharge rate is considered constant over one-hour intervals, which are shown in Figure 5-10.
- The measured process reflects the maximum strain response to each truck at sensor location H4, and follows a GPD characterized by

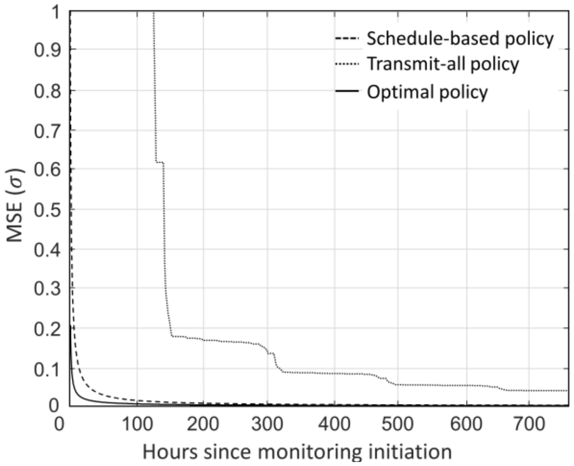
  $$f(y; \boldsymbol{\theta}_0) = \frac{1}{\sigma_0} \cdot \left[1 + \xi_0 \cdot \frac{y}{\sigma_0}\right]^{-\frac{1}{\xi_0}-1} \tag{5.17}$$

  with true parameters $\boldsymbol{\theta}_0 = [\xi_0, \sigma_0]^T$, where $\xi_0 = 0.311$ and $\sigma_0 = 2.80$.
- The battery can support four transmissions with $k = 5$ messages each (i.e., $N = 20$).
- The original belief about the measured process before the monitoring period begins follows a GPD with parameters $\xi_1 = 0.301$ and $\sigma_1 = 2.84$.
- After each batch transmission the remote parameter estimator updates the parameter estimates and calculates the measured reliability index, β^r, using the FORM algorithm presented in Chapter 3, where the assumed limit state function is defined in Equation (5-16).

Given these input parameters, we simulate the data collection and transmission process over the monitoring period assuming 1) the optimal policy is implemented, 2) the



(a)



(b)

Figure 5-12. MSE of the (a) location parameter and (b) scale parameter calculated after each batch transmission. The optimal, best schedule-based, and transmit-all policies are considered.

best schedule-based policy is implemented, and 3) the transmit-all policy is implemented. For these simulations, the strain response measurement corresponding to each truck is drawn from the GPD distribution representing the true process described in Figure 5-5. The results of these three simulations are presented in Figure 5-12. The first 750 hours are shown so that that the different convergence rates corresponding to the three different policies can be more easily distinguished. As expected, these numerical results illustrate the gains achieved using the optimal policy as compared to the best schedule-based and

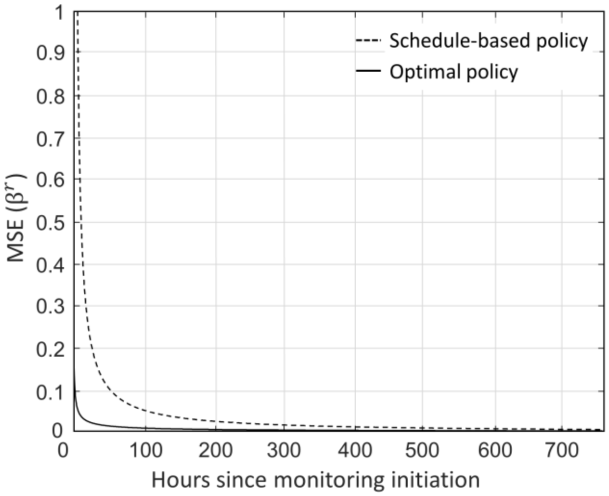


Figure 5-13. MSE of the reliability index calculated after each batch transmission. The optimal and best schedule-based policies are considered.

transmit-all policies. Based on the results of Chapter 4, the implemented optimal policy minimizes the variance of the estimated component parameters subject to the system constraints. Consequently, bridge owners can update the reliability index more frequently without compromising accuracy; as shown in Figure 5-13, the MSE of the reliability index achieved using the optimal policy after a 200-hour monitoring period would not be reached using the best schedule-based policy until after a nearly 750-hour monitoring period.

5.6 Conclusion

This chapter draws on an optimal data collection and transmission policy for remote parameter estimation in WSNs under resource constraints that was derived in Chapter 4 and extends the proposed framework to an SHM system installed on the TRB. The extension of such an optimal policy to control data collection in SHM systems is motivated by the observation that structural monitoring data is being increasingly incorporated into decision-making processes for asset management using reliability methods. Consequently, there is a need for an automated data collection and transmission strategy that facilitates the characterization of the statistical parameters of structural response data as quickly and accurately as possible so that bridge owners can track structural condition regularly. Given the TRB WSN architecture's battery and ExSRAM sizes, stochastic models of energy and truck arrivals based on NOAA weather station data and MDOT WIM station data, respectively, the measured strain response to passing trucks, and temporal death, the proposed policy controls the storage and transmission of measured strain data such that the variance of the measured process' parameter components is minimized.

Given the optimal policy, only a subset of the measured information is wirelessly transmitted to the remote estimator for processing; candidate data that are not transmitted are missing because either data is rejected when $y < \tau_S^*$ (i.e., left censored), or stored data is discarded when $y \geq \tau_S^*$ and the energy buffer is fully depleted (i.e., right censored). A modified likelihood function is implemented that accounts for the missingness of data. Under regulatory conditions, the modified MLE retains desirable properties that are characteristic of standard MLE based on a likelihood function given full-information.

When applied to the EH WSN architecture installed on the TRB, the optimal policy produces the best possible estimate of the process parameters given the system constraints. The presented numerical results illustrate the gains achieved using the optimal policy as compared to the best deterministic schedule-based and transmit-all policies. The numerical results show how implementation of the optimal policy results in faster convergence towards the true reliability index value quantifying the safety of the asset. This means that reliability-based decision-making practices can be used to track structural condition with higher frequency and accuracy in post-processing applications.

This research would benefit from future work that incorporates predictive estimation of the recharge rate, β_s, and event arrival rate, λ_s, at the beginning of each epoch, s. This would also help to inform varying lengths of epochs s that adapt to dynamically changing conditions. Next steps should also include expanding this theoretical framework from a single-hop to network-level scale. This will be facilitated by moving the computation of the optimal threshold from a remote server to a WSN node's local embedded processor. To facilitate this transition, a surrogate model should be implemented to simplify the solution for the optimal threshold value to facilitate less complex on-board embedded processing.

5.7 References

American Association of State Highway and Transportation Officials (AASHTO). *LRFD bridge design specifications*. 4th ed. Washington: AASHTO, DC, 2007.

Arnab R. Nonsampling errors. In: *Survey sampling theory and applications*. London: Academic Press, 2017, pp. 469-504.

Estes AC and Frangopol DM. Repair optimization of highway bridges using system reliability approach. *J Struct Eng* 1991; 125(7): 766-775.

Fisher JW and Yuceoglu U. A survey of localized cracking in steel bridges. Lehigh University Fritz Laboratory Report No. 448-2(81), 1981. Report No. R-1358, 1998.

Flanigan KA, Lynch JP and Ettouney M. Probabilistic fatigue assessment of monitored railroad bridge components using long-term response data in a reliability framework. *J Struct Health Monit* 2020a; DOI: 10.1177/1475921720915712.

Flanigan KA, Lynch JP and Ettouney M. Quantitatively linking long-term monitoring data to condition ratings through a reliability-based framework. *J Struct Health Monit* 2020b; DOI: 10.1177/1475921720949965.

Frangopol DM and Estes AC. Lifetime bridge maintenance strategies based on system reliability. *Struct Eng Int* 1997; 7(3): 193-198.

Han D, You K, Xie L et al. Optimal parameter estimation under controlled communication over sensor networks. *IEEE Trans Signal Process* 2015; 63(24): 6473-6485.

Hess PE, Bruchman D, Assakkaf IA, et al. Uncertainties in material and geometric strength and load variables. *Nav Eng J* 2002; 114: 139-166.

Ho CK, Khoa PD and Ming PC. Markovian models for harvested energy in wireless communications. In: *IEEE international conference on communication systems*, Singapore, 2010.

Hou R, Jeong S, Law K, et al. Reidentification of trucks in highway corridors using convolutional neural networks to link truck weights to bridge responses. In: *SPIE sensors and smart structures technologies for civil, mechanical, and aerospace systems* (eds JP Lynch, H Huang, H Sohn, et al.), Denver, CO, 4 March – 7 March 2019, paper no. 109700P, pp. 1-14. Bellingham: SPIE.

Hou R, Jeong S, Lynch JP et al. Cyber-physical system architecture for automating the mapping of truck loads to bridge behavior using computer vision in connected highway corridors. *Transp Res Part C: Emerging Technol* 2020; 111: 547-571.

Howard RA. *Dynamic programming and Markov processes*. Cambridge: MIT Press, 1960.

Jansson PO. Analysis of stress distribution in link plates used for suspending bridge beams. Michigan Department of Transportation Report No. R-1517, 2008.

Jeong S, Hou R, Lynch JP et al. A scalable cloud-based cyberinfrastructure platform for bridge monitoring. *Struct Infrastruct Eng* 2019; 15(1): 82-102.

Juntunen DA. Study of Michigan's link plate and pin assemblies. Michigan Department of Transportation Report No. R-1358, 1998.

Kane MB, Peckens C and Lynch JP. Introduction to wireless structural monitoring systems: design and selection. In: *Sensor technologies for civil infrastructures: performance assessment & health monitoring, Vol. 1*. London: Woodhead Publishing, 2014.

Kim S, Pakzad S, Culler D, Demmel J, et al. Health monitoring of civil infrastructures using wireless sensor networks. In: *6th international symposium on information processing in sensor networks,* Cambridge, MA, 25-27 April 2007, pp. 254-263. IEEE.

McCall B and Vodrazka WC Jr. *States' successful practices weigh*-in-motion *handbook*. Washington, DC: Federal Highway Administration, 1997.

National Oceanic and Atmospheric Administration (NOAA). Climate data online, https://www.ncdc.noaa.gov/cdo-web/ (2020a, accessed 12 June 2020).

National Oceanic and Atmospheric Administration (NOAA). Key to METAR surface weather observations, https://www.ncdc.noaa.gov/orders/qclcd/ExplanationofCodes.htm (2020b, accessed 12 June 2020).

National Transportation Safety Board (NTSB). Highway accident report: collapse of a suspended span of interstate Route 95 highway bridge over the Mianus River. NTSB Bureau of Accident Investigation Report No. NTSB/HAR-84-03, 1984.

Nikolaidis E, Ghiocel DM and Singhal S. *Engineering design reliability*. 1st ed. Boca Raton: CRC Press, 2004.

Nowak AS. Live load model for highway bridges. *Struct Saf* 1993; 13(1-2): 53-66.

Nowak AS and Zhou J. System reliability models for bridges. *Struct Saf* 1990; 7(2-4): 247-254.

O'Connor SM, Zhang Y, Lynch JP, et al. Long-term performance assessment of the Telegraph Road bridge using a permanent wireless monitoring system and automated statistical process control analytics. *Struct and Infrast Eng* 2017; 13: 604-624.

South JM, Hahin C and Telford RO. Analysis, inspection, and repair methods for pin connections on Illinois bridges. United States Department of Transportation Report No. FHWA/IL/PR-01,1992.

Swartz RQ, Jung D, Lynch JP, et al. Design of a wireless sensor for scalable distributed in-network computation in a structural health monitoring system. In: *Proc of the 5th international workshop on structural health monitoring* (eds F Chang and F Kopsaftopoulos), Menlo Park, CA, 12 September – 14 September 2005, pp. 1570-1577. Lancaster: DEStech Publications, Inc.

CHAPTER 6

Community Engagement Using Urban Sensing – Technology Development, Deployment Studies, and CPSS Architecture

6.1 Introduction

6.1.1 Smart Cities and Current Limitations

There has been considerable interest globally in smart cities due to the emergence of game-changing technologies including IoT platforms, cloud computing, and powerful automation architectures. The application of heterogeneous IoT technologies and network services to sensing in urban environments enables the development of smart cities, which are those cities in which the use of IoT sensor networks, massive sets of urban data, and ubiquitous access to cloud computing enhance the performance of urban systems and experiences of citizens. Already, a host of smart city applications have been deployed including connected and autonomous vehicles, controlled urban watersheds, environmental sensing, pedestrian and vehicle tracking using cameras, among many others (AOT 2015, Mehmood et al. 2017, Placemeter 2015). In addition to the realization of these applications in cities, various conceptual criteria have also been proposed regarding the integration of IoT technologies into smart cities (Ganchev et al. 2013, Jin et al. 2014, Mitton et al. 2012, Zanella et al. 2014).

However, a general-purpose urban sensing architecture has not yet emerged that is diverse enough to enable the management of wide arrays of heterogeneous IoT technologies and empower all stakeholders in a city to partake in data collection and data-driven decision making (Mehmood et al. 2017). This may be in large part due to the commercial sector and its marketing approach tailored to government stakeholders. Consequently, the proliferation of sensors and government-centric data aggregation in urban cities has disproportionately focused on improving city cores, with less attention paid to residential neighborhoods and areas. This is worsened by technological obstacles associated with current IoT platforms, such as the high power demand of existing hardware which requires access to power sources (e.g. light poles, electrical trash receptacles, and other powered street furniture), thus limiting the potential of mobile sensors and deployments in cities like Detroit where community access to power sources is severely limited in residential areas. This lack of connection between citizens and smart city initiatives is particularly pronounced in depopulated American cities (e.g., Detroit, Flint, St. Louis, Baltimore) where there has been no notable success in using smart city technologies to connect populations to their larger communities.

While most urban cities in the United States have undergone sustained periods of economic growth since the 1940s, a small number of cities have seen dramatic drops in population and economic activity (Ryan 2012). For example, Detroit experienced population reductions from 1.8 million people in 1950 to less than 700,000 in 2015 (The Detroit New 2016). However, major revitalization efforts in Detroit's business core are starting to rapidly transform the city. However, the residential areas and neighborhoods of Detroit remain underpopulated, with scarce access to important city services. Shrinking cities have resulted in extreme levels of poverty and inequity that result in stressors that disproportionately impact urban youth, who are at risk of losing connectivity to their cities and communities (Luthar 1991). As a result, there is a need to expand the use of sensing, especially by the general public, through the development of a more democratized approach to urban data collection and post-collection data uses to fully and more broadly reap the promise of smart cities. This, in part, entails engaging urban youth and citizens with their communities by architecturally embedding them within a smart city's urban CPSS. As seen in Figure 1-3, this expanded CPS architecture directly integrates humans into the CPS framework by taking into account a citizen's ability to observe and take action in response to physical and CPS elements (Wang 2010). By empowering communities to collect their own data in their neighborhoods and cities, city governments, local organizations and citizens can work in a more meaningful partnership with each other, leading to more resilient modes of smart city governance.

In addition to the lack of current urban sensing strategies that can empower all stakeholders in a city to partake in data collection and data-driven decision making, there are several technological obstacles that impede the emergence of a successful general-purpose urban sensing architecture for widespread use. These include the development of a sensing platform that supports interoperability among diverse arrays of heterogeneous IoT devices, preserves privacy and trust among citizens, supports cloud-based analytics, and supports low-power and low-cost sensing and communication, which is particularly difficult to achieve with platforms that require a continuous source of energy (Mehmood et al. 2017). For example, the Array of Things (AoT) deployment in Chicago provides environmental and air quality sensors tied to a Linux-based sensing node (AOT 2015). The high power demand of the hardware requires access to power sources supplied by street furniture (e.g. light poles) which limits its deployment potential in less populated areas and

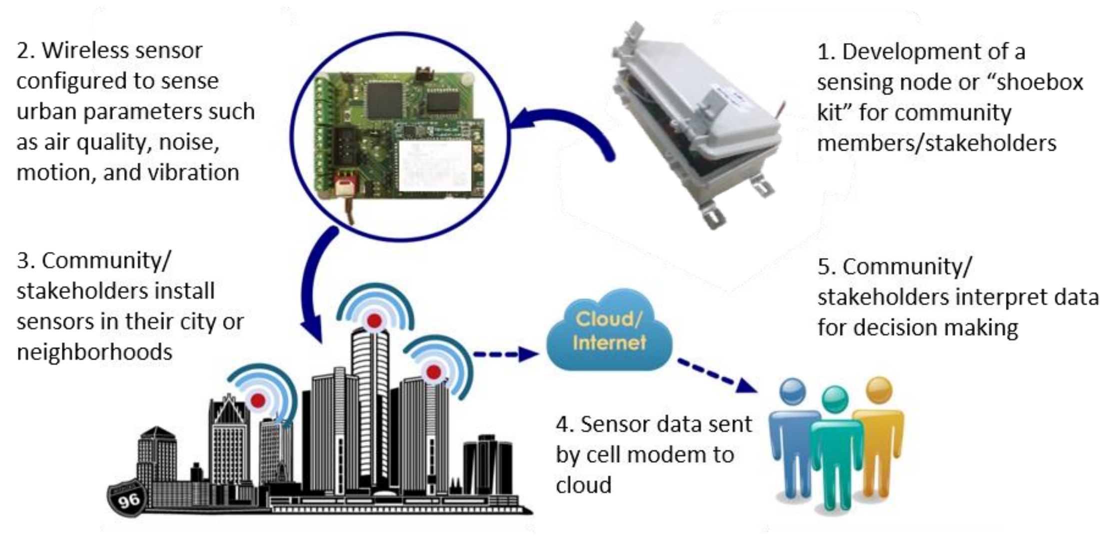


Figure 6-1. Overview of the implementation of the *Urbano* IoT platform.

residential neighborhoods where power sources are severely limited. In addition, Placemeter is a novel camera-based sensing solution designed to track vehicles and pedestrians in city spaces (Placemeter 2015). While Placemeter promises to anonymize data by processing video footage using automated data processing tools, many communities might be resistant to camera-based monitoring due to infringements on privacy and overarching notions of "Big Brother" within the community.

6.1.2 Introduction to the *Urbano* IoT Platform

In response to these needs, this chapter describes the development of the *Urbano* sensing node which is designed for dense and rapid deployment in cities for a wide variety of smart city applications (Figure 6-1). In particular, the design of the architecture is based on the belief that urban sensing can play a major role in empowering communities to collect data on urban processes of interest, and can transform how communities engage with other city stakeholders to make decisions. Hence, the design of *Urbano* emphasizes ease of use and minimizes dependence on required infrastructure for which a stakeholder may have limited or no access. *Urbano* is designed as an ultra-low power, low-cost, wireless sensor node that is capable of collecting diverse and heterogeneous sensor measurements, supporting embedded computing, and communicating using cellular or wireless communication, such as long-range radio (LoRa). A major differentiator of *Urbano* from other smart city IoT platforms is that it does not require a persistent power source (e.g., grid power available from street furniture) nor a wired communication medium (e.g., fiber network), yet can still meet standards necessary to sample, store data, and transmit information for use in SHM. Rather, it is designed to operate using solar energy and leverages a cellular radio to push data to the cloud. The case studies presented in this chapter avoid the use of cameras in order to respect the anonymity of citizens. For example, passive infrared (PIR) sensors are used for pedestrian counting instead of using cameras or Bluetooth, and GPS modules are only integrated into *Urbano* nodes with the consent of the relevant parties (e.g. food truck vendors).

Due to *Urbano's* low-power design, low cost, and independence from continuous power sources, nodes can be densely and rapidly deployed as stationary or mobile sensing units anywhere in a city. *Urbano* has analog and digital sensing interfaces, and a wide variety of sensing transducers have already been integrated with *Urbano* and deployed in urban cities. These include digital sensors such as air quality sensors (NO_2, SO_2, O_3, and particulate matter (PM)) and GPS receivers, in addition to analog sensors such as geophones for vibration measurements, strain gages, accelerometers, PIR sensors for pedestrian tracking, and temperature sensors. Sensors are connected to *Urbano's* analog and digital sensing interfaces (Figure 6-2, Figure 6-3) and are either housed inside of the node enclosure or connected to the *Urbano* externally through a water tight connection through the enclosure.

To emphasize its ability to support community uses, *Urbano* has been assembled in a user-friendly packaging with all components integrated, and consists of a variety of libraries of data processing blocks that support the different sensing applications (including those that require onboard data analytics); community members can simply deploy, turn on the device, and see the data stream to a data portal of their choice. *Urbano* nodes are designed to push their data to a database server hosted in a commercial cloud environment. A variety of data portals are exposed. The cellular modem integrated with *Urbano* can be used to issue data and alerts in the form of SMS messages and Twitter posts to allow nodes to essentially tweet alerts and updates to users subscribed to their feed. This specific approach to data dissemination is well suited to presenting urban data and information using a user-friendly interface. The second approach adopts a more robust cloud-based data management platform well suited for storage and management of time history series. Here, Exosite's One Platform (Exosite 2017) is adopted as a time series database that is ideally suited to collect and manage *Urbano* data streams. Graphical representations of the data are provided using Exosite's internal standard and customizable visualization portals. This chapter details the hardware design, software architecture, and data processing approaches implemented in an analytics layer that queries data from the database.

The *Urbano* sensor node has been deployed in a number of smart city engagements in Michigan. These applications incorporate a wide variety of stakeholders including city governments, local organizations, urban youth, and communities. All of the applications described in this chapter are carried out using the *Urbano* cloud-based sensing IoT platform. The first test case highlights the use of GPS-enabled *Urbano* nodes to track food trucks in Grand Rapids, MI to assess compliance with permit rules and curbside management by city planning officials. The second application deploys *Urbano* nodes to monitor pedestrian traffic along the Detroit waterfront; this data is desired by the Detroit Riverfront Conservancy to understand utilization of public spaces to guide future investments. The third test case packages *Urbano* nodes for community deployment in southwest Detroit as part of the *Sensors in a Shoebox* program. One of five iterations of this program is discussed in this chapter, in which youth explore why a local park is under-utilized by the community as compared to surrounding areas, as well as how the space might be improved to address the community's needs. Additionally, the youth measure air quality at the park because residents in southwest Detroit reside in one of the most polluted regions of Michigan. As a result of poor air quality there have been high rates of youth asthma and long-term cardiovascular disease in the community (EPA 2015).

6.2 Hardware Architecture

The hardware of the *Urbano* wireless sensing node is separated into three primary subsystems: analog and digital sensing interface, computational core, and wireless communication. The flexible sensing interface is compatible with a diverse array of heterogeneous analog and digital sensing transducers. The computational core is programmed to operate the hardware and carry out on-chip data processing. The wireless communication system incorporates a cellular modem to push and pull data to and from the cloud. In addition to an overview of wireless sensing, the following sub-sections provide a detailed overview of the hardware design, cellular communication design, and

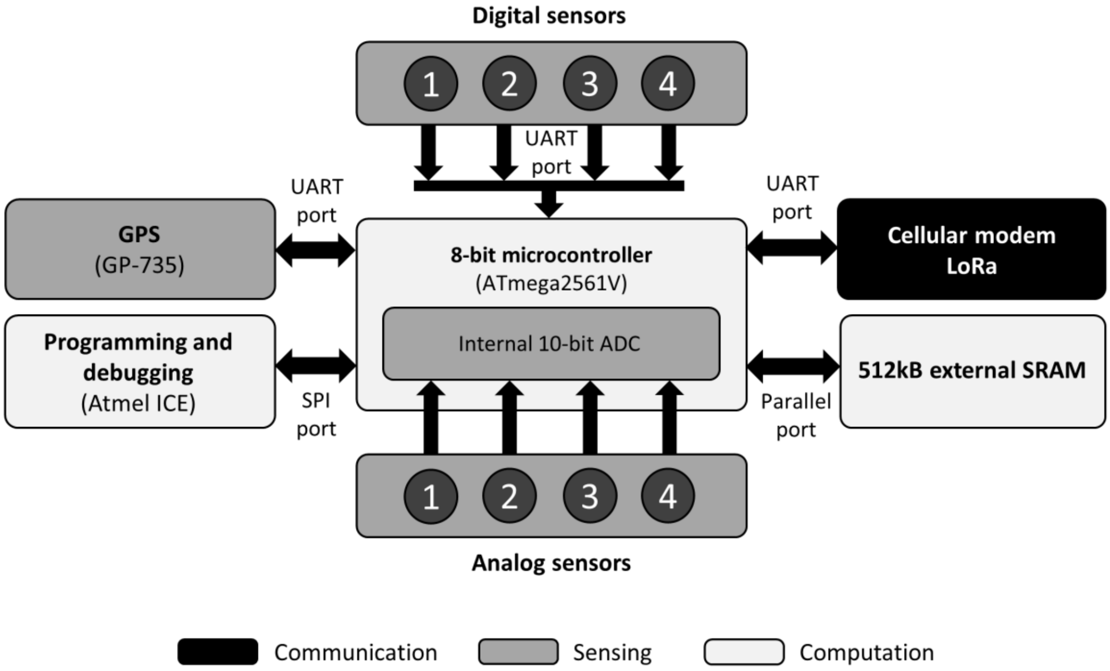


Figure 6-2. Hardware design for sensing interface, computational core, and wireless communication.

packaging of the *Urbano* wireless sensing node to demonstrate how the hardware satisfies the needs identified in Section 6.1. The general architecture of the hardware design is shown in Figure 6-2.

6.2.1 Introduction to Wireless Sensing

Wireless sensing has emerged as a major platform for collecting and transmitting data both within and outside of cities over the past several decades. Wireless sensing has been particularly prominent within the context of SHM, in which wireless sensing systems are installed on infrastructure in order to detect damage and use data to help guide decisions to repair, rehabilitate, or replace a structure (Straser et al. 1998). Wireless sensing technologies for monitoring infrastructure originally emerged as an alternative to existing wired systems, which require high upfront costs due to material procurement, and labor intensive installation, not to mention extensive installation times (Swartz et al. 2007).

Using commercial off-the-shelf electrical components, researchers have successfully developed and deployed wireless systems for use in monitoring infrastructure (Lynch and Loh 2006). However, there remain several limitations to current wireless sensing platforms. Specifically, limited ranges of wireless transceivers and the dependence of wireless sensing nodes on connections to wireless communication infrastructure, such as local base stations (that house single-board computers), hinder low-cost dense and rapid deployments over large areas, and the ability to have mobile sensors.

The limitations associated with current wireless sensing platforms used for monitoring single-asset infrastructure are particularly important to recognize as more attention turns toward monitoring multi-asset infrastructure, physical systems, and environmental parameters within urban environments where connectivity and automation are inherent features. Due to the need for a wireless sensing platform for which stakeholders with sensing needs in a city can deploy dense networks (possibly up to hundreds of sensors) of stationary and mobile sensors rapidly for data collection, wireless sensing nodes must be completely autonomous with no dependence on additional communication infrastructure (such as base stations).

While smart city IoT applications continue to emerge that are reliant on short-range technologies such as Bluetooth, Zigbee, and Wi-Fi (Mehmood et al. 2017), it quickly becomes less feasible for citizens to engage in and initiate smart city applications in under-resourced areas of cities, such as in neighborhoods outside of city centers, where power and access to additional communication infrastructure necessary to enable these means of communication, are scarce. On the other hand, with typical coverage of 5-30 km (Mehmood et al. 2017), long-range technologies such as cellular modems and LoRa are much more feasible and autonomous methods of communication that give city governments, local organizations, and citizens the flexibility to sense a diverse array of parameters throughout a city. A drawback to using cellular communication for data transmission is that cellular modems have high energy consumption during active communication. However, as outlined in the following subsections, the hardware and software architectures of the *Urbano* IoT cloud-based flexible sensing platform are designed to minimize power consumption such that each node is able to use solar energy harvesting with a small solar panel to realize full autonomy from access to power and communication infrastructure aside from a cellular network (applications in this chapter use various solar panels between 3.4-10W).

6.2.2 Hardware Design

6.2.2.1 Computational Core

The computational core is programmed to carry out three main functions: the operation of the hardware, data interrogation and on-chip processing, and network communication. At the center of the computational core of the *Urbano* node is an Atmel AVR ATmega 2561V 8-bit microcontroller with 8MHz system clock, which operates with a 3.3V supply voltage. An 8-bit microcontroller was selected to avoid the significantly higher power consumption and costs associated with 16- and 32-bit microcontrollers. In addition, an 8-bit internal data bus is sufficient for the required on-chip processing. The ATmega 2561V is a low power microcontroller that has a current consumption of 7.3mA in active mode. Strategic software manipulation of the microcontroller's sleep modes allows for reduced power consumption when the node is not actively collecting data, as the microcontroller consumes

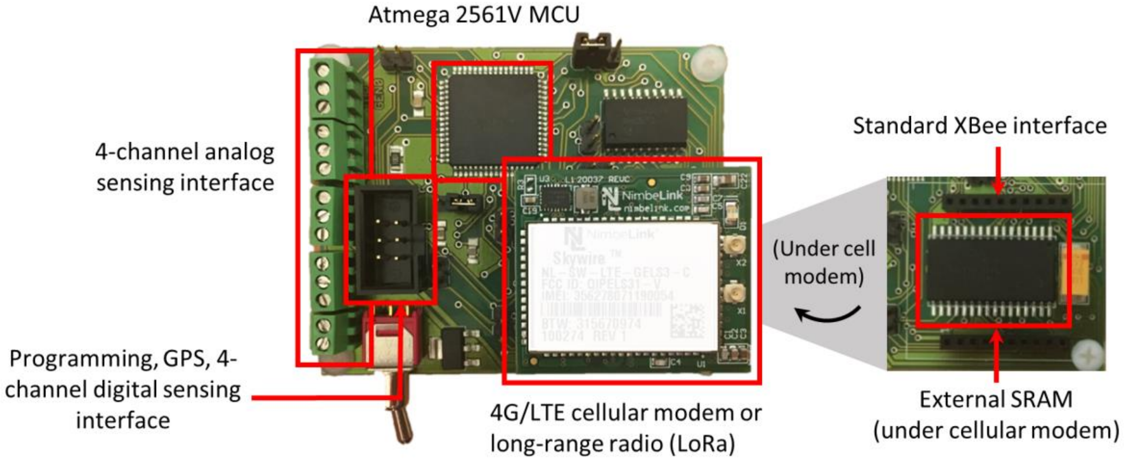


Figure 6-3. *Urbano* node with key components highlighted.

4.5μA in power-save mode, and 0.18μA in power-down mode. As a result, the microcontroller is able to perform data processing computations and operate the sensing and communication subsystems using very little power.

The ATmega 2561V has more than sufficient read-only memory, with 256kB of flash and 4kB of EEPROM, which reduces constraints on data interrogation and on-chip processing. The microcontroller has 8kB of internal SRAM. Since 8kB of internal SRAM is not sufficient when large amounts of data need to be stored, the microcontroller is configured to include an extended 512kB of external SRAM to augment the internal memory using the Cypress CY62148EV30. Since *Urbano* nodes do not require a continuous power source and the cellular modem consumes the most power in the system, additional SRAM is valuable so that it is possible to execute embedded data-processing algorithms on large amounts of data and only transmit necessary, pre-processed information. Using the solar harvesting configuration, the *Urbano* node is capable of the periodic transmission of both raw continuous time-series data and pre-processed data. In the case where the node has access to a power source (i.e. solar harvesting is not necessary), there is no limit on the frequency of data transmission.

6.2.2.2 Analog and Digital Sensing Interface

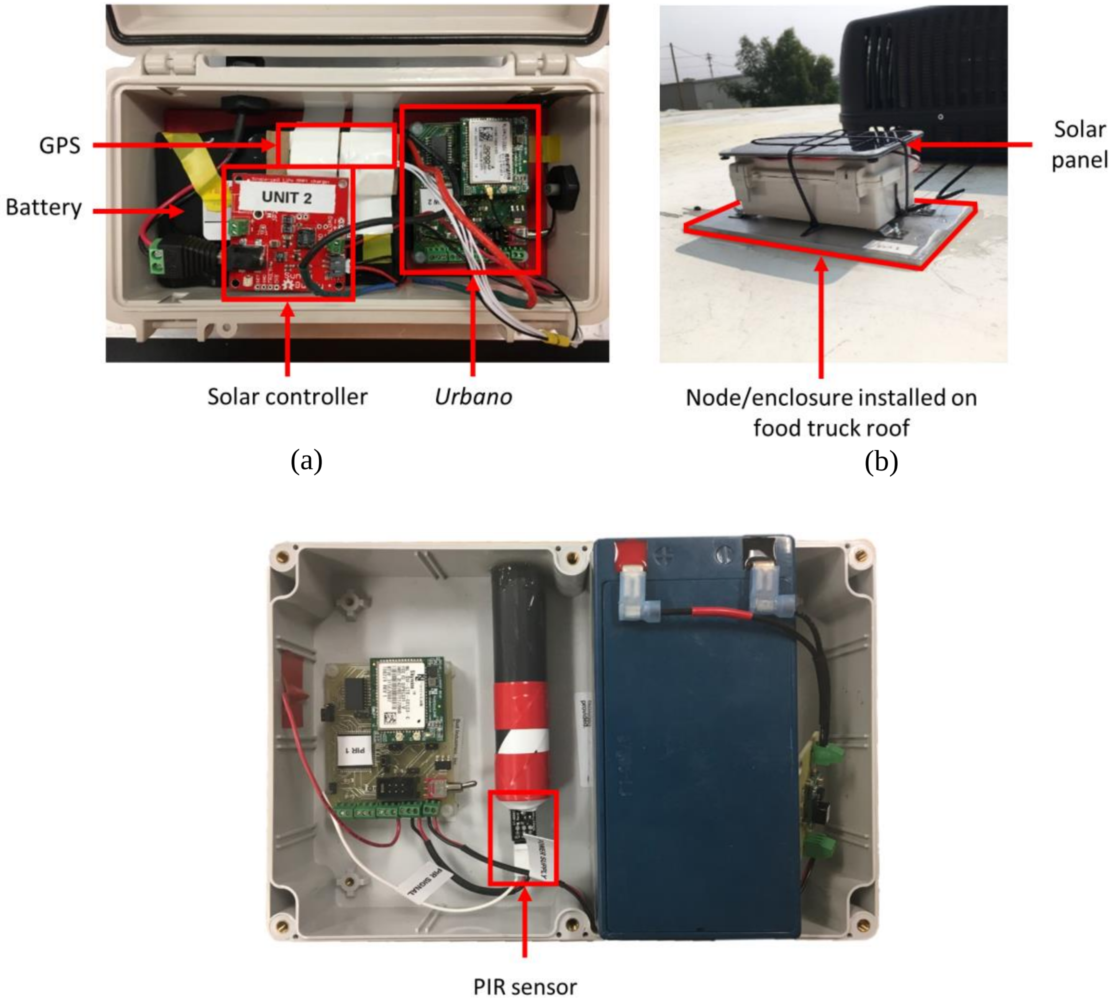
Controlled by pre-programmed data collection schemes in the microcontroller, the sensing interface includes four analog and four digital sensing channels. The sensing channels support heterogeneous sensing transducers so that each node can sense a diverse array of parameters, such as environmental *and* physical parameters. An internal 10-bit 200kHz analog-to-digital converter (ADC) in the microcontroller is used for digitizing analog signals to enable embedded processing and transmission of data to the cloud using the cellular modem. For sensing infrastructure systems such as bridges and buildings (i.e. structural health monitoring), a higher resolution 16-bit ADC would be more appropriate. However, for sensing urban parameters such as mobility (e.g. vehicle and pedestrian movement), and various environmental parameters, 10-bits is sufficient. While an external 16-bit ADC can be easily integrated into the *Urbano* node, it is desirable to avoid the additional power consumption and cost that accompany higher resolution external ADCs.

6.2.2.3 Cellular Wireless Communication

Wireless communication is achieved using the Nimbelink Skywire 4G/LTE Cat 1 Cellular Embedded Modem. A standard XBee hardware interface is used to connect the cellular modem to the printed circuit board (PCB) (Figure 6-3). The cellular modem connects to Verizon or AT&T’s 4G LTE network and is the lowest power fully developed LTE technology available in the market (Nimbelink 2017). The Nimbelink Cat 1 modem consumes 616mA of current during active cellular communication and 48mA when idle, but the utilization of sleep modes can reduce current consumption to 8.6mA in low power mode and 44µA when it is off. To minimize the amount of power consumed by the cellular modem, the computational core is programmed with robust timing and interrupt schemes to ensure the cellular modem remains off whenever active read and write transmissions are not necessary. In addition, the modem is Federal Communications Commission (FCC) and end-device pre-certified, meaning that it does not require carrier certification, which significantly reduces the cost and eliminates the time associated with the certification process (which can take up to months in the United States). In addition, the cell modem supports multiple LTE bands (B4(1700) and B13(750)) with fallback capabilities, and has a small U.FL port for antenna flexibility. The Nimbelink Cat 1 is based on the Gemalto ELS31 chipset and achieves excellent speeds of 10Mbps download and 5Mbps upload. The Gemalto ELS31 is designed for power optimization and speed which make it an excellent candidate for machine-to-machine and IoT applications. In addition, the Nimbelink Cat 1 has a commercially available development kit that allows users to connect the cell modem to a PC via a USB-to-UART converter, and send AT commands to the modem through any serial terminal application. This direct and simple method of communication enables rapid development and debugging.

6.2.3 Packaging

To emphasize its ability to support community uses, *Urbano* has been assembled in a user-friendly packaging (smaller than the size of a shoebox). All components are integrated into a single enclosure, and the *Urbano* node is programmed with a variety of libraries of data processing blocks that support the different sensing applications. Without making significant changes to the physical structure of the node, community members can simply



(a) (b)

(c)

Figure 6-4. (a) *Urbano* node assembled for studying food truck curbside management; (b) *Urbano* node installed on the roof of a food truck with solar panel mounted to top of packaging; (c) sensing kit for pedestrian sensing.

deploy, turn on the device, and see the data stream to a data portal of their choice (visualization portals discussed in Section 6.3). For the three applications discussed in Section 6.4, two different methods of packaging are used. For deployments by the City of Grand Rapids planning officials and the Detroit Riverfront Conservancy, these two stakeholders desired a compact assembly. As seen in Figure 6-4(a), the *Urbano* node, sensing components, rechargeable lithium ion battery, solar controller, and a small 3.4W solar panel were assembled in a small 20.32cm x 10.16cm x 7.62cm weatherproof box. Figure 6-4(b) shows the *Urbano* node installed on the roof of a food truck in Grand Rapids, MI, complete with solar panel. On the other hand, an alternative packaging scheme (Figure 6-4(c)) is used for pedestrian sensing (Section 6.4.3), which allows for more room to house the PIR sensor. For educational purposes, a clear lid is included with the kit so that students and community members can observe the contents of the box even when it is deployed and collecting data.

6.3 Software Architecture

The success of a diverse sensing solution for smart city applications relies on the implementation of a scalable cloud-based database system for the storage, processing, and analysis of sensor data transmitted from the *Urbano* nodes. In response to this need, *Urbano* nodes are designed to push collected data to a database server hosted in a commercial cloud environment. An additional requirement is that data be accessible and able to be interpreted by all stakeholders, including city governments, local organizations, and citizens, regardless of their education level or familiarity with IoT technologies and cloud computing. As a result, a variety of data portals are exposed. For example, pre-programmed code on *Urbano's* microcontroller (computational core) enables the Nimbelink Cat 1 cellular modem to send data and issue alerts in the form of SMS messages and Twitter posts to allow the nodes to tweet their notifications to users subscribed to their feed. This specific approach to data dissemination is well suited for presenting urban data and information using a user-friendly interface. The second approach adopts a more robust

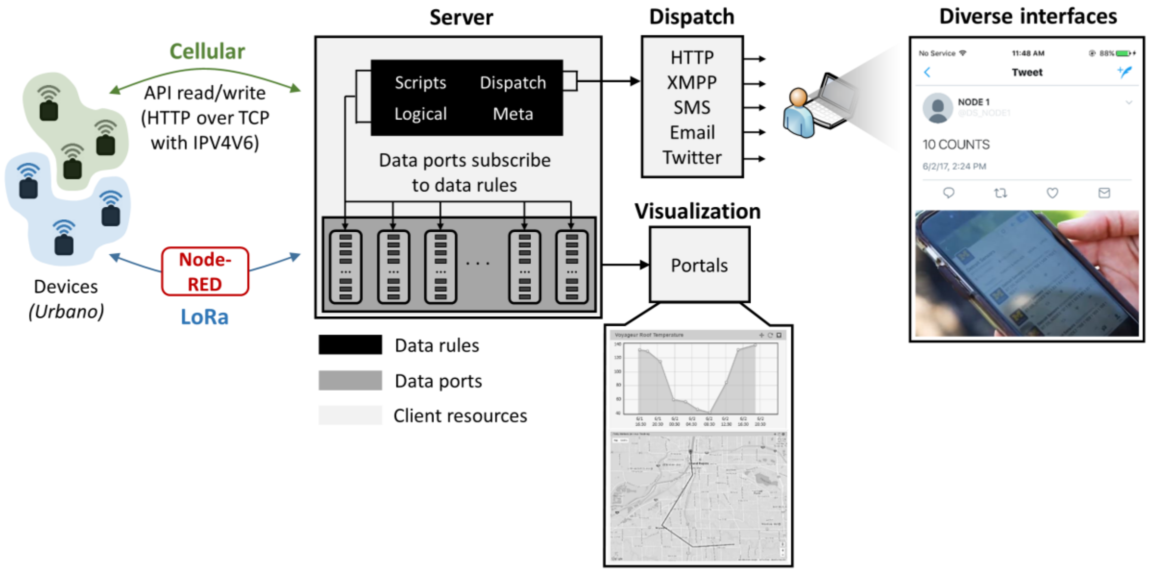


Figure 6-5. Data flow between *Urbano* nodes and the cloud.

data management web server platform that is well suited for storage, cloud-based analytics, and management of time history data.

Urbano's software architecture utilizes Exosite's commercially available One Platform as a time series database that is ideally suited to store and manage data collected and transmitted by each *Urbano* node. Exosite was selected due to its manageable system development, user friendly visualization tools, and primary focus as an IoT software platform. As seen in Figure 6-5, each user maintains a client, which consists of several resources such as data ports, data rules, meta, dispatch functions, and visualization portals. Each client is identified by an assigned Client Identifier Key (CIK) that is used for accessing application programming interface (API) calls to the One Platform and mapping to a client's resource identifier (RID). Data that is written from a device (*Urbano* node) is tagged with an alias that is used to map the data to a corresponding data port in the One Platform database. The data is then stored in the appropriate data port's data-store. Each data port can subscribe to user-defined algorithms, known as data rules, which are scripts and logical statements that are written in the Lua language to process data. Data rules are also used to call dispatches, which are outputs from the One Platform in forms such as HTTP, XMPP, SMS, email, and Twitter. This means that data can even be seamlessly integrated to another robust server or cloud service, such as Amazon Web Services, or to a user-friendly interface such as Twitter.

The software that is embedded onto *Urbano's* microcontroller (C language) is programmed to issue and receive AT commands through the Nimbelink Cat 1 cell modem to connect and interact with the server. For data transmission, the *Urbano* network utilizes unconstrained protocol stacks where requests and responses between the devices and web server are managed by HTTP application layer protocols. Under the HTTP protocol, TCP is used as a transport layer protocol to handle the HTTP traffic. For the network layer protocol, the Nimbelink Cat 1 is an IP capable device and the microcontroller is programmed to utilize an IPV4V6 dual stack PDP context that simultaneously supports both IPv6 and IPv4 using the cellular modem.

Graphical representations of the data are provided using Exosite's internal standard and customizable visualization portals so that data can be easily interpreted by all stakeholders. The applications described in this chapter (Section 6.4) leverage several of these portals. For example, Exosite's GIS tools are leveraged to track the movement of food trucks throughout the city of Grand Rapids, numerous data time series associated with environmental parameters are plotted, and Twitter dispatches are enabled to tweet data to subscribers. In addition, portals are leveraged to show the current air quality associated with several pollutants (e.g. NO_2, SO_2, O_3, and PM), which are automatically colored green, yellow, or red, to indicate the value's relationship to pre-defined "Safe Level," "Warning Level," and "Alert Level" thresholds as defined by the Environmental Protection Agency (EPA).

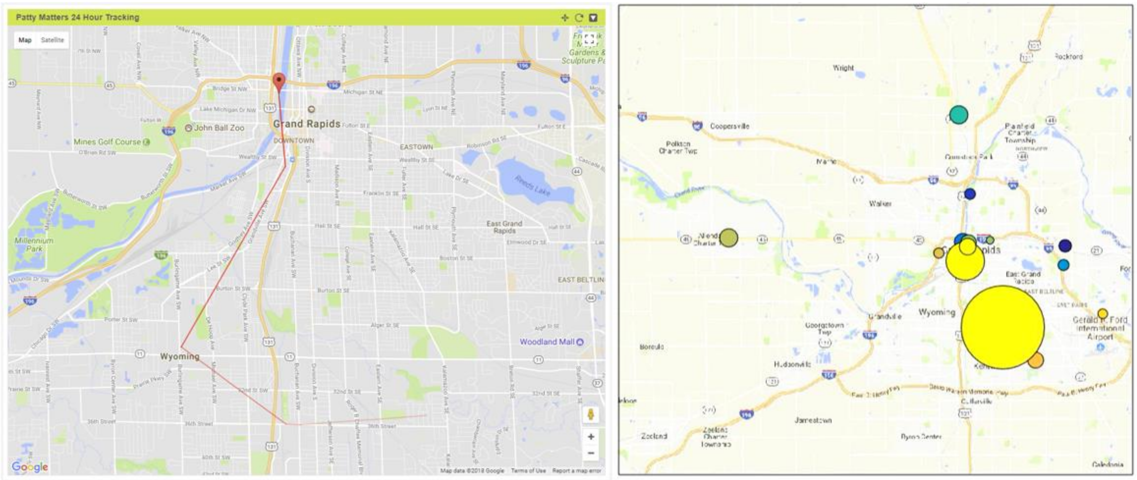
6.4 CPS Applications

The *Urbano* sensing node has been deployed in a number of smart city engagements in Michigan. The three applications described in the following subsections incorporate a wide variety of stakeholders including city governments, local organizations, urban youth, and communities. These diverse applications demonstrate the flexibility of the *Urbano* platform, as a wide variety of analog and digital sensors are interfaced on both mobile and stationary sensing nodes using only the single *Urbano* platform.

6.4.1 Food Trucks as Mobile Sensors

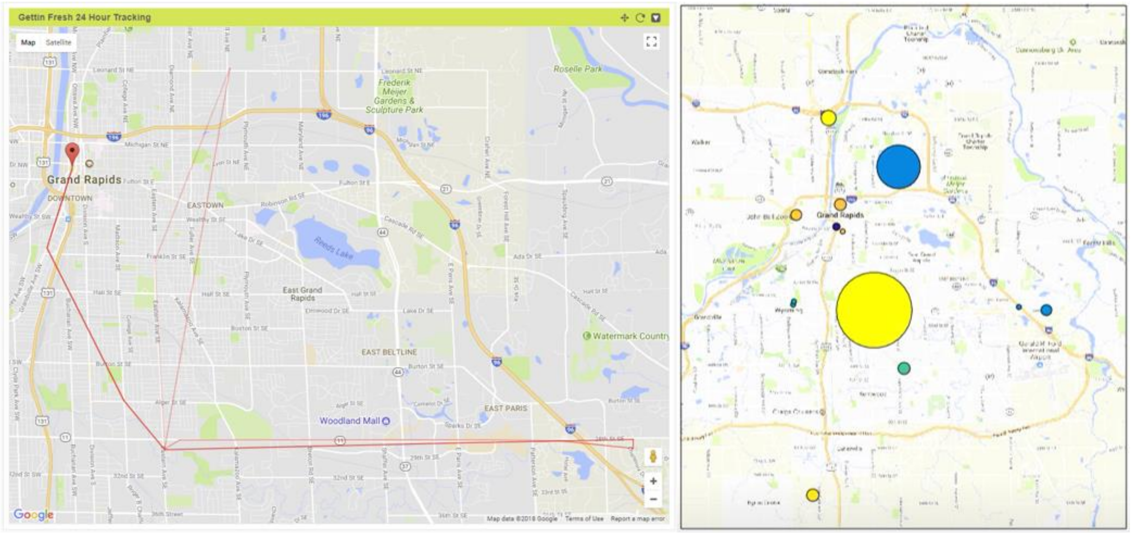
Using the *Urbano* platform, an array of GPS-enabled mobile *Urbano* nodes were deployed on food trucks (Figure 6-4(b)) in the city of Grand Rapids in order to observe the behavior and locations of these mobile assets for two months. This first test case was desired by the city's planning and transportation officials in order to assess compliance with permit rules and to explore novel curbside management models. Monitoring the food trucks was mutually beneficial for both the city's officials *and* the food truck vendors. Food truck vendors readily volunteered for this pilot programs, as it allows them to provide information in real-time to their customer base regarding their current location.

In addition to storing GPS output data in the One Platform's database, an Exosite trail map portal is used to visualize each truck's movement within the past 24 hours (Figure 6-6(a), Figure 6-7(a)). In addition, Matlab is leveraged to help visualize the duration of



(a) (b)

Figure 6-6. Patty Matters food truck (a) 24-hour tracking period; (b) duration spent parked at curbside locations for one month.



(a) (b)

Figure 6-7. Gettin' Fresh food truck (a) 24-hour tracking period; (b) duration spent parked at curbside locations.

time each truck spends parked at various curbside locations. The time that a food truck spends stationary at a known curbside location is proportional to the area of each circle associated with that location (Figure 6-6(b), Figure 6-7(b)). In addition to leveraging the One Platform's server, database, and user-friendly visualization tools, the *Urbano* platform is flexible enough to be directly integrated into the City of Grand Rapid's existing GIS platform. In upcoming engagements with the City, dispatch scripts will be configured to output data from the relevant data ports (corresponding to various food trucks) to an existing ArcGIS server that is currently used by the City. Collected data showed that the food trucks spend a considerable amount of time at a wide range of locations, including colleges, community parks, downtown business areas, and residential neighborhoods. As a result, a diverse array of air quality sensors, including NO_2, SO_2, O_3, and PM, will be added to each mobile *Urbano* node to monitor air pollution throughout the city. This allows for air quality to be monitored in areas outside of the city core that are often neglected because power and access to additional communication infrastructure necessary to enable existing smart city technologies, are scarce.

6.4.2 Pedestrian Counting Along the Detroit Riverfront

Depopulation of the residential sectors of Detroit has left many public spaces and parks neglected. Of the city's 307 parks, the city is quietly closing some of them (NPR 2015). As the city decides on its plan for public spaces, various stakeholders are interested in observing and assessing the use of their public and park spaces. For the second application, *Urbano* nodes are deployed along the Detroit Riverfront to monitor pedestrian traffic in this popular park area along the Detroit River. While the data collection approaches described in this study are applicable to any public space, the author partnered with the Detroit Riverfront Conservancy (DRFC) for demonstration. The DRFC manages the development, operations, and improvement of the Detroit Riverfront along the Detroit River and the Dequindre Cut greenway, which runs perpendicular to the Detroit River and provides pedestrian access to the Riverfront (DRFC 2013). Around three million people

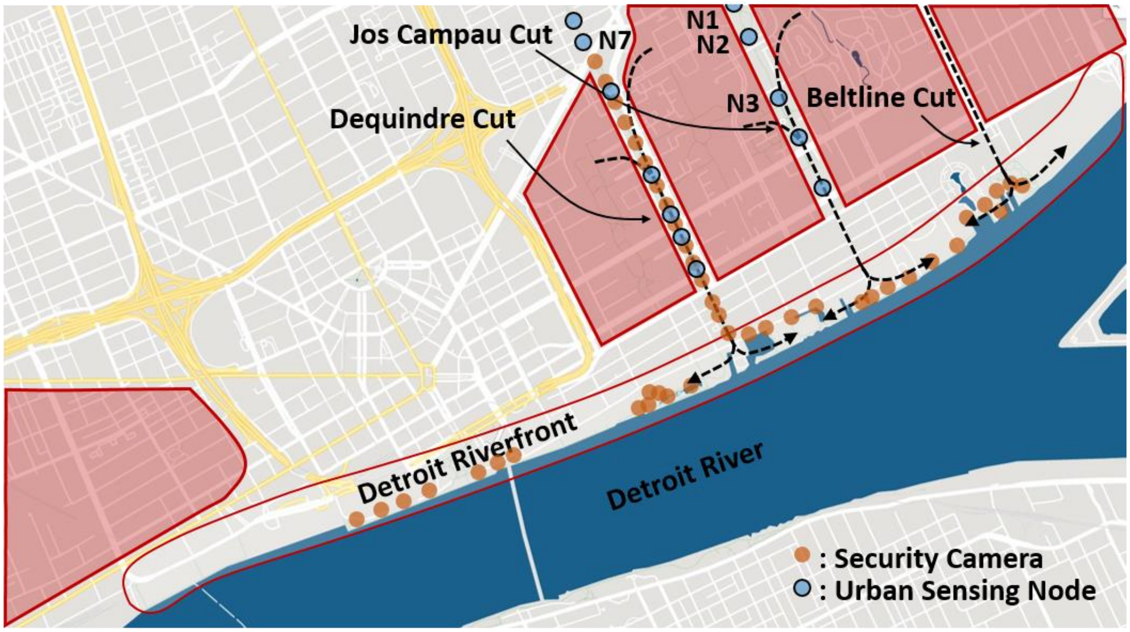


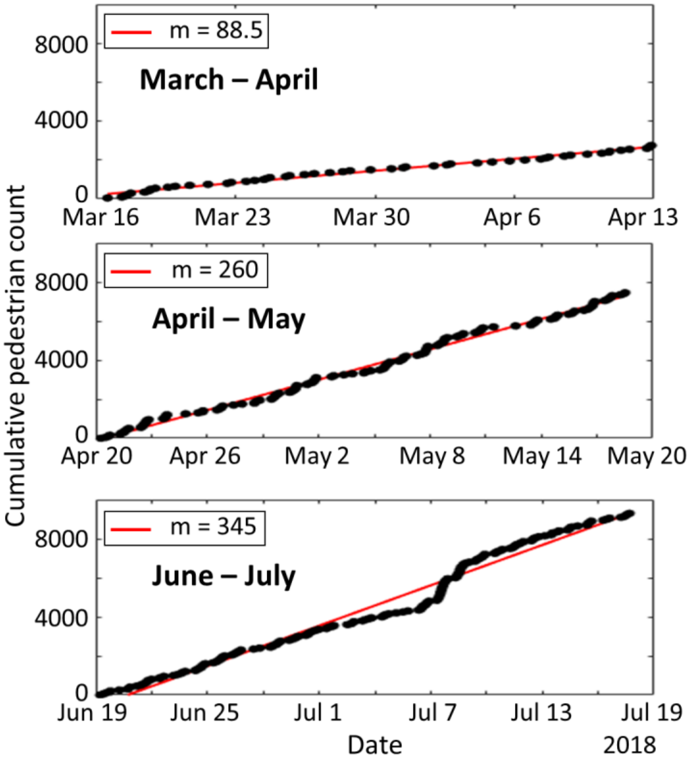
Figure 6-8. Map of the Detroit Riverfront including adjacent vulnerable neighborhoods, Dequindre Cut, Jos Campau Cut, location of security cameras, and location of urban sensing nodes.

visit these two areas annually (DRFC 2018). In addition to serving as a social space where patrons can exercise and enjoy natural experiences, these parks are assets that serve as engines of economic growth in the residential communities north of the river.

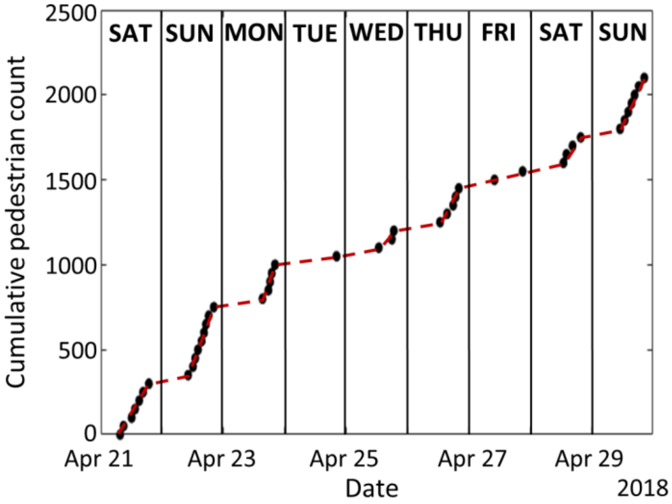
The DRFC is currently making major investments to create two new greenways parallel to the Dequindre Cut: the Jos Campau Cut and the Beltline Cut (Figure 6-8). The majority of the insight into park usage patterns is based on sparsely collected surveys and visual observation. The DFRC previously used commercially available pedestrian counting technology to monitor the use of public spaces in several parks throughout Detroit (Eco Counter 2017). However, the costs associated with the existing commercial technology, in particular, the cost to continuously replace the device-specific battery, were too high. Consequently, the DRFC would like to understand how these spaces are used

Figure 6-9. *Urbano* sensing node measuring pedestrian traffic along the Dequindre Cut in Detroit, MI.

and where to invest future resources to make these areas more accessible. With the Jos Campau and Beltline greenways nearing development, the DRFC wants to understand how people from communities use these new connections to reach the waterfront. In collaboration with the DRFC, twelve *Urbano* nodes with PIR sensors are installed along the Dequindre Cut and the Jos Campau Cut in order to gain quantitative insight into patron use. Low-power PIR sensors connected to autonomous *Urbano* sensing nodes are installed on light poles to record, timestamp, and transmit the number of pedestrians passing key park fixtures while maintaining full anonymity (Figure 6-9). The One Platform's Twitter output dispatch is used to Tweet live pedestrian count updates for the nodes distributed along the Riverfront (Figure 6-5).



(a)



(b)

Figure 6-10. (a) Cumulative pedestrian counts from March 16, 2018 through July 18, 2018; (b) a one-week subset of this data illustrating the daily trends (i.e., weekend versus weekday).

The Dequindre Cut is a renovated greenway that is well-maintained and connects neighborhoods in Detroit's East Side to the Detroit Riverfront. The first phase of the

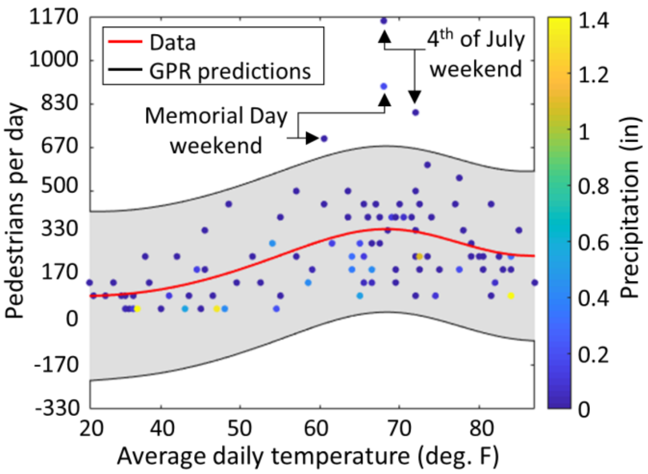
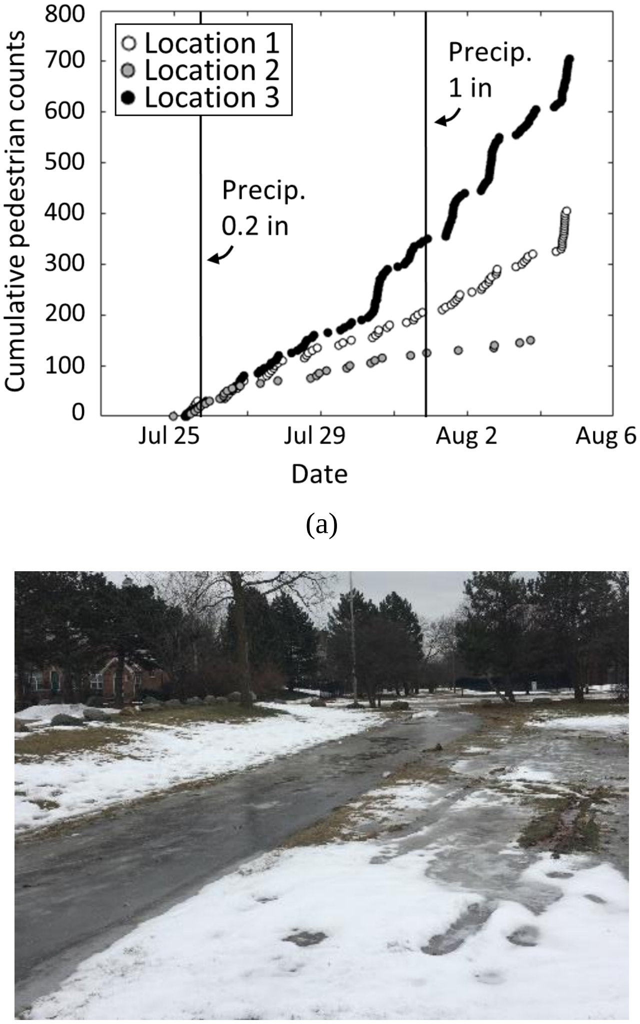


Figure 6-11. Comparison of the number of pedestrians per day, temperature, and precipitation.

installation was carried out along the Dequindre Cut during March 2018 and data collection continues today. Sample data for three months (from spring through summer) is shown in Figure 6-10(a) for Node 7 (denoted as N7 in Figure 6-8). The slope, m (pedestrians per day), is calculated for each one-month period (Figure 6-10(a)). Intuitively, as the weather improves from March through July, the number of pedestrians using the greenway increases drastically. The data in Figure 6-10(a) quantitatively supports this intuition. During the same three months, the number of pedestrians using the Dequindre Cut each day increases as the weather becomes warmer and precipitation decreases, and the pedestrians per day decreases when the temperature becomes too high (Figure 6-11). At a one-week resolution, the data shown in Figure 6-10(b) illustrates the increase in usage during the weekends compared to weekdays. At a one-day scale, holidays (e.g., Memorial Day weekend, fourth of July weekend) correspond to increased park usage (Figure 6-11).

Unlike the Dequindre Cut, the Jos Campau Cut has not been renovated and is nearing preparation for further development. The second phase of the installation was carried out along the Jos Campau Cut in July 2018 and data collection continues today.



(a)

(b)

Figure 6-12. (a) Cumulative pedestrian counts along the instrumented Jos Campau Cut corridor; (b) flooding along the Jos Campau Cut renders the pathway unusable during the winter and after heavy rainfall.

Sample data for a one week period is shown in Figure 6-12(a) for Nodes 1, 2, and 3 (denoted as N1, N2 and N3 in Figure 6-8). Comparing Figure 6-12(a) and Figure 6-10, even though the these two greenways run parallel to each other and are located less than a quarter of a mile apart, nearly ten times the number of people use the renovated Dequindre Cut per day than the Jos Campau Cut. Further inspection of Figure 6-12(a) indicates that even though N1, N2 and N3 are on the same path, the three locations experience different

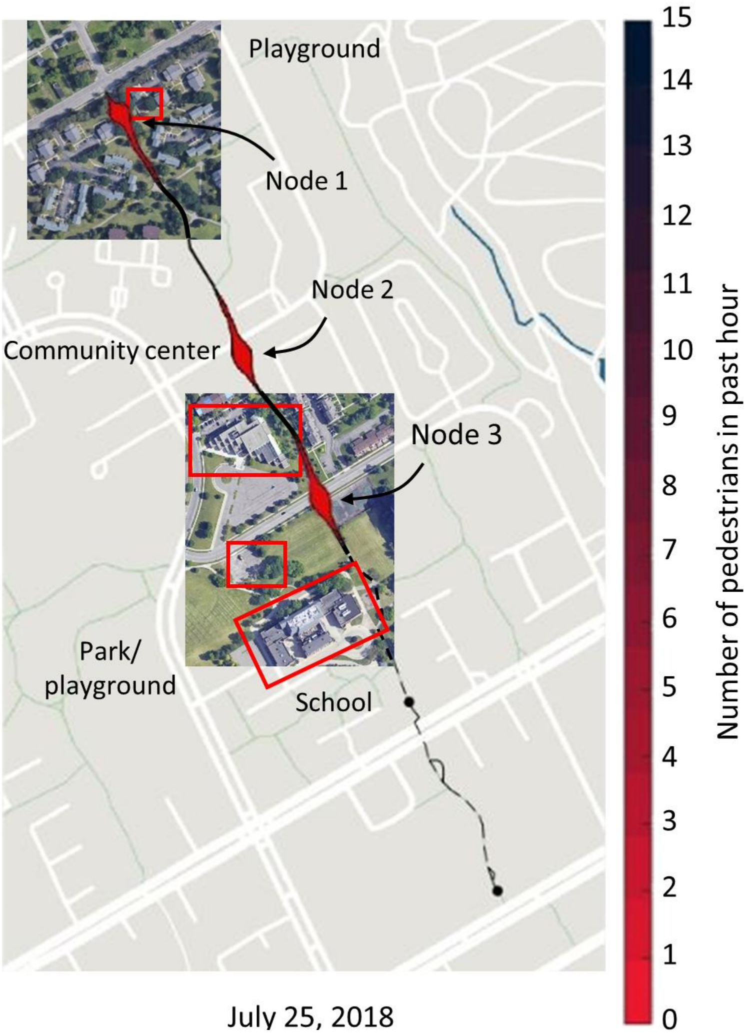


Figure 6-13. Instrumented pedestrian corridor along the Jos Campau Cut in Detroit, MI.

pedestrian use. In order to better understand this phenomenon, prominent landmarks along the Jos Campau Cut are highlighted in Figure 6-13. N3 sees the highest pedestrian traffic and is located next to a school, community center, park, and playground. N1 experiences the second highest amount of pedestrian traffic and is located near a park within an apartment complex. N2 is not located next to any significant social spaces, but it monitors the segment of the greenway between N1 and N3. The author returned to the location of N2 and observed that the entire path in this segment of the greenway is covered in mud and dirt. As Figure 6-12(a) indicates, pedestrian counts decrease after it rains because water several inches deep collect on the path. Figure 6-12(b) provides a picture of the pathway at N2 during the winter. Pooled water freezes into ice about three inches thick and renders the path unusable.

While assessing the use of public park spaces is of particular importance to the DRFC, the same platform can easily be implemented in other parts of the city to help quantify community mobility more generally. As Detroit communities begin to fortify and rebound, communities would benefit from understanding the mobility of their neighbors. In particular, quantitative data on the utilization of roads and pedestrian pathways could inform them in the quest to secure transportation infrastructure investments.

6.5 *Sensors in a Shoebox*: CPSS Application to Strengthen Community Resilience

In order to move towards the realization of smart cities, there is a need to support citizens and, importantly, urban youth in developing the basic skills necessary to engage with sensing technologies so that they can be connected to their communities in a meaningful way. This section provides an overview of the design and development of a community-governed, place-based educational program called *Sensors in a Shoebox* that is implemented at full scale in Detroit. This program aims to address issues related to equity and governance over decision making in communities. In contrast to engineering processes that initiate from technology and sensing, the engineering process within the *Sensors in a Shoebox* program centers on understanding and designing sensing strategies based on the needs of the community and human users. This program leverages the *Urbano* IoT platform and expands a traditional CPS architecture to directly integrate human action by

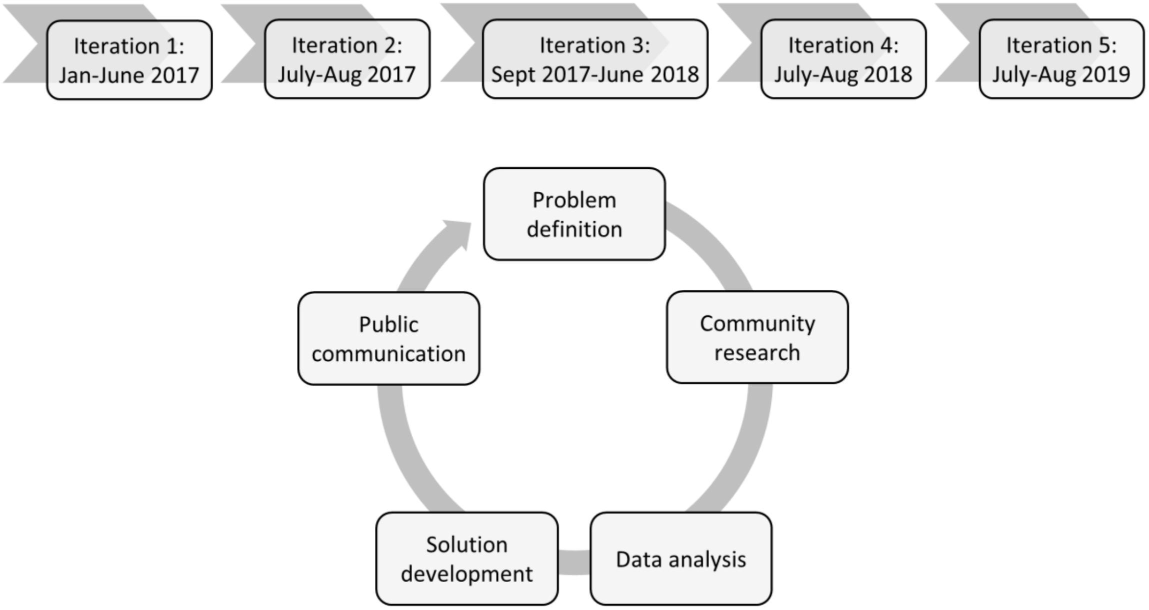


Figure 6-14. Overview of the educational programming using the *Urbano* IoT platform.

taking into account a citizen's ability to observe and take action in response to CPS elements.

The overarching goal of this program is to support young peoples' existing connections to their communities through problem solving and to empower them as citizen scientists. Guiding the design of the educational programming is an emphasis on participatory action—or meaningfully working with youth and communities in the engineering process—to promote youth autonomy in engineering work and to engage them in engineering-rich work in community places with tangible outcomes. Given the goal of developing a sustained CPSS, this work draws on human-centered design models situated in community to guide the content of the educational programming. A historical design tradition within engineering, human-centered design is "an ongoing search for what can be done to support and strengthen the dignity of human beings as they act out their lives in varied social, economic, political, and cultural circumstances" (Buchanan 2001). It is a constant engineering process that centers on understanding and designing for the needs of human users, rather than starting from technology (Cross 2001). It establishes the purpose of engineering as understanding and solving problems that humans have, rather than using technology for technology's sake. As an engineering process, human-centered design requires designers to fully scope and define their engineering problem as understood by the full stakeholder community, conduct research to better understand the community problem space, and continue to bring human uses into each aspect of the solution design work (Cross 2001, Dym et al. 2005, Zoltowski et al. 2012). To support this work, the program leveraged an array of research methods alongside the *Urbano* IoT platform, including survey, observation, and interviewing methods. A schematic outlining how engineering is conceptualized in the programming is shown in Figure 6-14.

To show how a CPSS was enabled through the *Sensors in a Shoebox* programming, the remainder of this section illustrates the project specifics for one iteration of the educational programming. The focus of this explanative analysis is the fourth iteration of the program, which ran from July through August of 2018. The researchers worked with a youth summer camp program that is part of a well-known community organization serving a predominately Latinx population in Detroit. The partnership with this organization was paramount because 1) it continues to be a trusted community organization that serves a population of youth historically marginalized in traditional engineering contexts; 2) it has a pre-established summer program for youth with flexibility in the schedule to allow for projects; and 3) it has an interest in developing youths' science, technology, engineering, and math (STEM) skills. During the summer camp, youth selected a programming option that they attended for five weeks. The group met Monday through Wednesday for three hours a day. Together, the youth and research team defined and refined a problem the youth felt compelled to address in their community using engineering and research. During this iteration of the program, the youth decided to explore why a local park located across the street from the organization was under-utilized by the community as compared to surrounding areas, as well as how the space might be improved to address the community's needs. Additionally, the youth measured air quality at the park because residents in southwest Detroit reside in one of the most polluted regions of Michigan due to the presence of heavy industries including steel mills, oil refining, and coal fired power plants. As a result of the poor air quality, there have been elevated rates of cardiovascular disease and youth asthma (EPA 2015); many of the youth have been directly impacted by this.

The *Sensors in a Shoebox* program comprises four primary stages: 1) youth identify and define problems; 2) youth conduct community research and observe these problems using observation data, interview data, and sensor data; 3) youth intervene and devise solutions; and 4) youth collaborate with diverse stakeholders to communicate their findings publicly and actuate change in their community (Figure 6-14).

6.5.1 Identify Problem

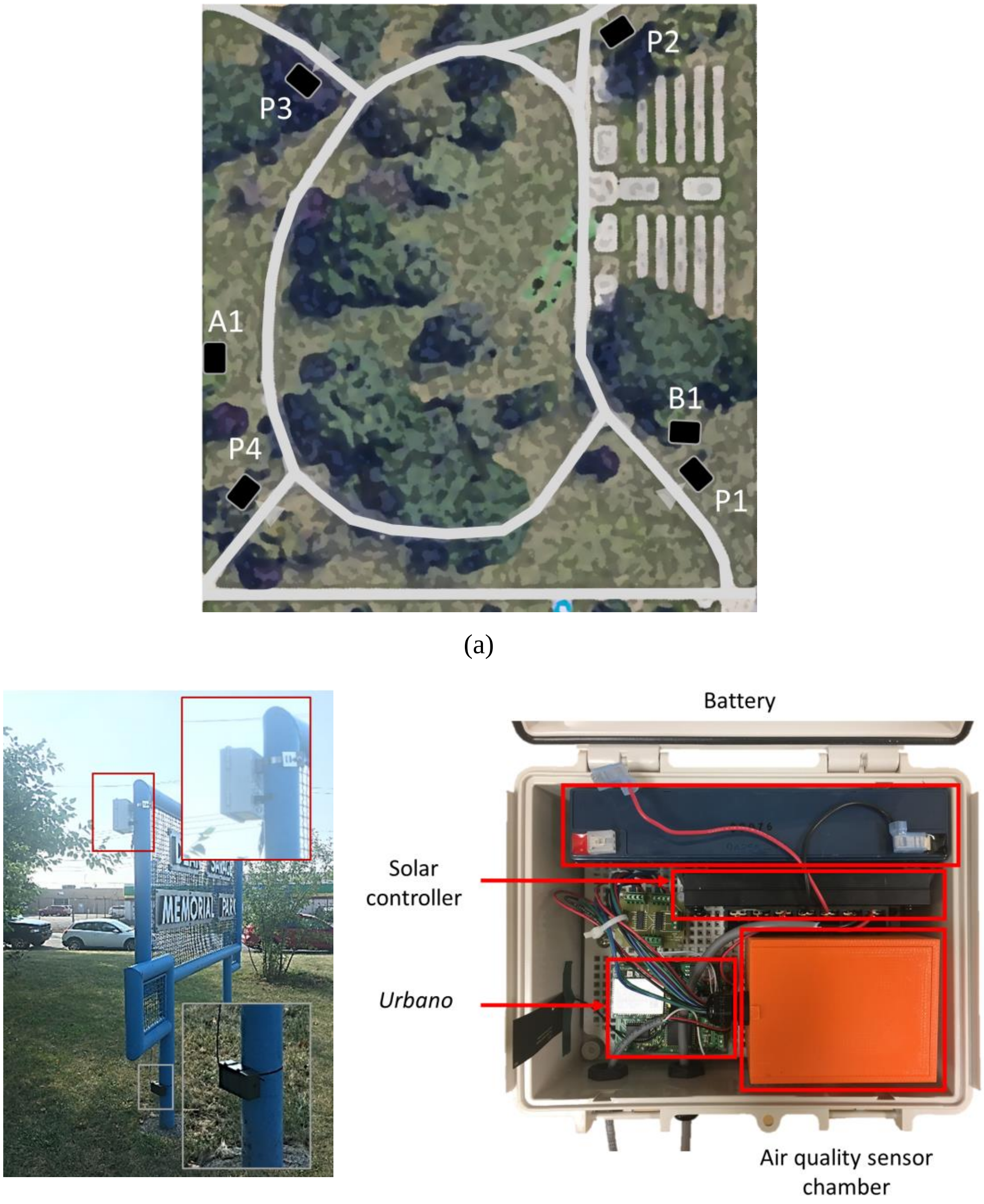
Drawing from principles of human-centered design, the youth in the *Sensors in a Shoebox* program and the research team first worked together to define and scope a problem they were interested in pursuing. The first day of the programming involved three key features: 1) a community walk in which the youth and research team physically explored the surrounding neighborhood and engaged in open observation; 2) an introduction to the *Urbano* IoT platform in which the youth were introduced to the innerworkings of the technology and some of its potential functions; and 3) a community problem or interest brainstorming session in which students drew or wrote problem spaces of interest for the summer work. The programming was intentionally designed in this order to facilitate natural connections between the community space and sensing work. This work generated long lists of potential problem spaces. In the subsequent days, the youth and research team worked to narrow down the focus by using the "five-whys" critical thinking strategy, asset scoping, and engaging in some preliminary conversations with stakeholders. During the fourth iteration of the *Sensors in a Shoebox* program discussed in this section, this resulted in the youth forming a collective interest in the park across the street because 1) some youth had previously explored it; 2) it is under-utilized for no easily discernable reason; and 3) the sensor technology offered some benefits to exploring and designing within the space. In this way, the purpose of the emerging CPSS was driven by community interest and shaped by access to progressive technology.

6.5.2 Community Research and Data Analysis

The scoping process then guided a targeted data collection process related to the identified problems in which the youth worked as a team to design a coherent research plan. To begin this process, the youth considered the benefits and tradeoffs between the different types of data collection tools, including surveys, interviews, observation, and sensing. This discussion was aimed at helping the youth to understand both the limitations and capabilities of social science tools and sensing technologies for learning about community needs and problems. The youth and research team then worked to construct a plan using all of these data collection methods and detailed the logistical considerations for each method. The youth wrote interview questions and developed survey tools, and then edited them accordingly before initiating the community data collection effort. The youth then determined what type of sensing data would be helpful for their work (namely, pedestrian traffic and air quality) and assembled the *Urbano* sensing kits accordingly. Using a satellite image of the park, the youth mapped where they wanted to install the sensing nodes to collect foot traffic data (P1 through P4 in Figure 6-15(a)) and to measure air quality (A1 in Figure 6-15(a)). The youth also built a paralleled observation protocol to track user behaviors in order to support and enrich the data they would collect with the sensors.

To collect data, part of the youth went to the park and surrounding establishments to observe and survey those who might be or become stakeholders in the park. Other youth called local business owners to interview them about their hopes for the space in the context of their business. Finally, other youth deployed the PIR and air quality sensors to collect data over the course of a week. Figure 6-15(b) shows one of the *Urbano* nodes installed at location P1 that is used for pedestrian tracking. Similar to the data discussed in Section 6.4.2, these sensing nodes transmit timestamped pedestrian counting data wirelessly to the cloud. Additionally, air quality sensors were installed at location A1. The air quality sensors were housed inside of a 3D printed chamber within the node that exposed the air quality sensors to the outside environment while keeping the node's internal circuitry weatherproof (Figure 6-15(c)). For this deployment, youth installed air quality sensors inside of *Urbano* nodes (measuring NO_2, SO_2, and O_3) and installed the sensing nodes on light poles at the park.

Given the timing constraints of the summer programming, social science data was collected for two programming sessions, a limitation that the youth would later articulate in their analysis with the team. The benefit of the sensors being relatively hands-off once



(a)

(b) (c)

Figure 6-15. (a) Sensing plan including pedestrian counting sensors (P1-P4), sensing bench (B1), and air quality sensors (A1); (b) pedestrian counting sensors installed at location P1; (c) *Urbano* node configured to sense air quality.

they were deployed was that it allowed for longer data collection and greater opportunity to establish the necessary social science skills to develop and sustain a community-driven

(a) (b)

Figure 6-16. (a) *Urbano* nodes built into the "Super Bench" are accessed by removing the back panel; (b) PIR sensors are strategically installed to measure up to two people sitting on the bench.

CPSS. That is, it was important to pair the sensing data with social science research to holistically analyze the needs of the people using the park. One of the aims of the programming was to highlight the use of sensor technology as a powerful tool among other tools that the youth can use to explore and understand their surrounding communities.

6.5.3 Solution Development

Analyzing the data collected in the research phase then informed the solution development process. The goal of the programming was to present solution development as both a creative process and data driven. As such, in the park project, youth and the research team worked together to analyze the community data collected, understand the limitations of the data collected, and think about the different types of data as a whole set. These conversations put the data collected by the *Urbano* IoT platform in conversation with interviews with local business owners or surveys from park-goers. Using the data, youth asserted that a lack of comfortable, inviting seating was one of the driving concerns in the park. The time constraints of the summer programming necessitated youth focusing on one focal project, although they also outlined others they were interested in pursuing. This prompted the development of “Super Bench,” a bench built in collaboration with a community leader from Sit On It Detroit who salvaged wood from demolition around the city to make temporary seating. Wanting to better understand how additional seating would be used and to see its impact on the park usage, the youth built a bench with the *Urbano* IoT system embedded into it to track how many people sat on the bench and to continue to monitor foot traffic in the park to see if the bench increases the park usage. As shown in Figure 6-16, the *Urbano* sensing nodes were secured in an enclosed compartment within the bench’s back, and solar panels were embedded into the top of the bench. Attending to community feedback indicating interest in color and art in the park, the bench was then colorfully painted. Working with the commitments of the community organization and the community leader, a small, enclosed library was built into the bench and stocked with books in both English and Spanish languages. In this way, the solution development was inspired by data collected from all sources, which guided the development of the bench’s features. Further, its features supported an ongoing research system, supporting the youth in further understanding how the community used the park seating.

6.5.4 Public Communication

To conclude their time in the programming, the youth presented their findings and their design work to community stakeholders and civic leaders in a format of their choosing, communicating both their process and results of their design. The youth developed their presentation materials and practiced the ways they might discuss their work with various types of stakeholders. Given time constraints, the youth developed a written memo and proposal to further elaborate data-driven suggestions for projects within the space. The youth chose to develop posters characterizing the narrative of their design process. They presented these posters outside at the park, near the bench and sensing network they had installed. Leaders of community organizations, members of the surrounding community, local business owners, and park-goers were invited to interact with youth around their posters at the park. The youth also developed a potential improvement plan for the park, with suggestions ranging from more permanent seating options, secured recycling and trash bins, more shaded areas and colorful landscaping. Ultimately, these final presentations created an opportunity for community members to engage with the research and design work the youth had done and to encourage meaningful action.

6.5.5 Results and Impact

Introducing and situating the IoT *Urbano* platform within a larger community engineering educational program offers one potential map through which we can begin to develop a sustainable CPSS (Figure 6-17). Looking across this iteration of the *Sensors in a Shoebox* program, the youth were meaningfully engaged in the engineering design and research process, from beginning to end. By creating programming in which the youth were engaged in and drove the entirety of a design cycle situated in their communities, more opportunity was created for students to see engineering and technology as relevant to their lives and transformative in their communities. As a result of the fourth iteration of the program, the City renovated the park and a new bike repair and tire pump station, colorful

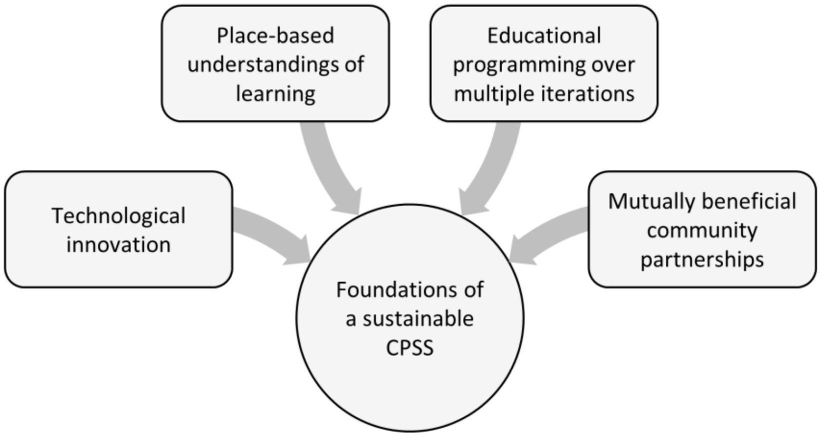


Figure 6-17. Integration of the *Urbano* platform within a larger community program offers a powerful mapping through which a sustainable CPSS emerges.

gardens, a covered picnic area, manicured trees, benches, recycle bins, trashcans, a pet waste station, and barbeque grills were added to the park.

Of the eight youth who participated in Iteration 4 of the programming, two youth were returning from the previous summer's programming exploring the park (Iteration 2, Figure 6-14). In Iteration 5 the following summer, four of the eight youth returned. On returning to the program, the youth highlighted the connection between their work with the *Sensors in a Shoebox* program and helping their community, which situates their participation as a contributing factor to positively impacting to their local area. These youth expressed interest in continuing their work on the park, suggesting that the data supported their work. This suggests that youth made connections between their research work, their design work, and city action that ultimately led to the park's renovation. Further, it suggests the need for more widespread structures that help to establish these connections by creating educational programming foundations (through clubs, community organizations, service centers etc.) that provide user-friendly technological innovation (i.e., the *Urbano* IoT platform) with the infrastructure through which to support meaningful change in a community.

6.6 Conclusions

The *Urbano* IoT cloud-based sensing node proposed in this chapter serves as a flexible sensing platform that can be densely and rapidly deployed throughout cities. The hardware and software architecture of *Urbano* nodes are designed to ensure ultra-low power consumption so that they can rely solely on solar energy harvesting to be self-sustaining. Since *Urbano* nodes are not reliant on continuous power sources (e.g. light poles or powered street furniture) and operate autonomously aside from connection to a cellular network, they can be deployed as stationary or mobile sensing units not only in city centers, but also in residential neighborhoods and areas where access to power and communication infrastructure can be scarce. This design architecture is based on the belief that urban sensing can play a major role in empowering communities to collect data on urban processes of interest, and can transform how communities engage with other city stakeholders to make decisions. To demonstrate the utility of the *Urbano* platform to a wide range of city stakeholders, three preliminary deployments studies were successfully carried out that engaged city governments, local organizations, urban youth, and communities. All three stakeholders were able to use a variety of analog and digital sensors to deploy both stationary and mobile sensing nodes, all using the same *Urbano* platform. While stakeholders have access to data in a time series database, a variety of user friendly visualization portals were leveraged to make data easy to interpret and accessible to community members. This platform offers a scalable and sustainable solution for enabling communities to help strengthen their city neighborhoods. In addition to helping to connect citizens to their communities, this platform provides a scientific and technological foundation to the extension of CPS to include humans. These CPSS human-in-the-loop systems have the potential to transform a variety of application areas including transportation, building energy management, among others.

6.7 References

AOT (Array of Things). Array of things, https://arrayofthings.github.io/ (2020, accessed 1 February 2016)

Buchanan R. Human dignity and human rights: thoughts on the principles of human-centered design. *Des Issues* 2001; 17(3): 35-39.

Cross, N. Designerly ways of knowing: design discipline versus design science. *Des Issues* 2001; 17(3): 49-55.

The Detroit News. Detroit population rank is lowest since 1850, www.detroitnews.com/story/news/local/detroit-city/2016/05/19/detroit-population-rank-lowest-since/84574198 (2016, accessed January 8 2018).

Detroit Riverfront Conservancy (DRFC). Economic impact study: Detroit riverfront. Report, DRFC, 2013.

Detroit Riverfront Conservancy (DRFC). DRFC, city of Detroit and Detroit economic growth corporation announce new plan to preserve and expand riverfront https://detroitriverfront.org/news/detroit-riverfront-conservancy-city-detroit-and-detroit-economic-growth-corporation-announce (2017, accessed 3 February 2019).

Dym CL, Agogino AM, Eris O, et al. Engineering design thinking, teaching, and learning. *J Eng Educ* 2005; 94(1): 103-120.

Eco Counter. Eco Counter, https://www.eco-compteur.com/en/applications/parks-recreation (2020, accessed 8 January 2018).

Environmental Protection Agency (EPA). Environmental issues in southwest Detroit, https://archieve.epa.gov/region5/swdetroit/web/html/index.html (2016, accessed 21 February 2016).

Exosite. Exosite, https://exosite.com (2020, accessed 15 January 2018).

Ganchev I, Ji Z and O'Droma M. A generic IoT architecture for smart cities. In: *25th IET Irish Signals & Systems Conference and China-Ireland International Conference on Information and Communications Technologies*, Limerick, Ireland, 2013, pp. 196-199.

Jin J, Gubbi J, Marusic S, et al. An information framework for creating a smart city through internet of things. *IEEE Internet Things J* 2014; 1(2): 112-121.

Luthar SS. Vulnerability and resilience: a study of high-risk adolescents. *Child Dev* 1991; 62(12): 600-616.

Lynch JP and Loh KJ. A summary review of wireless sensors and sensor networks for structural health monitoring. *Shock Vib Dig* 2006; 38(2): 91-128.

Mehmood Y, Ahmad F, Yaqoob I, et al. Internet-of-things-based smart cities: recent advances and challenges. *IEEE Commun Mag* 2017; 55(9): 16-24.

Mitton N, Papavassiliou S, Puliafito A et al.: Combining cloud and sensors in a smart city environment. *EURASIP J Wireless Commun Networking* 2012; 1: 1-10.

Nimbelink. Skywire LTE CAT 1, https://nimbelink.com/Documentation/Skywire/4G_LTE_Cat_1/30154_NL-SW-LTE-GELS3_ProductBrief.pdf (2017, accessed 13 January 2017).

National Public Radio (NPR). Shh! Detroit's little-known success story: its parks are getting better. http://michiganradio.org/post/shh-detroit-s-little-known-success-story-its-parks-are-getting-better#stream/0 (2015, accessed 22 February 2016).

Placemeter. The placemeter sensor, https://www.placemeter.com/ (2020, accessed 29 February 2016).

Ryan BD. *Design after decline: how America rebuilds shrinking cities*. 1st ed. Philadelphia: University of Pennsylvania Press, 2012.

Straser E and Kiremidjian A. Modular, wireless damage monitoring system for structures. Report, John A. Blume Earthquake Engineering Center, Menlo Park, CA, 1998.

Swartz AR, Zimmerman A and Lynch JP. Structural health monitoring system with the latest information technologies. In: *Proceedings of the 5th infrastructure & environmental management symposium*, Ube, Yamaguchi, Japan, 2007.

Wang FY. The emergence of intelligent enterprises: from CPS to CPSS. *IEEE Intell Syst* 2010; 25(4): 85-88.

Zanella A, Bui N, Castellani A, et al. Internet of things for smart cities. *IEEE Internet Things J* 2014; 1(1): 22-32.

Zoltowski CB, Oakes WC and Cardella ME. Students' ways of experiencing human-centered design. *J Eng Educ* 2012; 101(1): 28-59.

CHAPTER 7

Conclusions and Future Research Directions

7.1 Conclusions

In recent years, urbanization trends have imposed greater load demands on infrastructure and placed stressors on social urban systems. At the same time, naturally occurring hazards continue to threaten infrastructure and populations. Despite these challenges, opportunities are emerging out of the unprecedented proliferation of technologies enabling low-cost sensing, high-power computing, and actuation. Together, these advancements enable "intelligent," or cyber-physical, systems, which promise to greatly enhance the performance of the built environment. However, even with these advancements, the ability of professionals to "sense for decisions", which this dissertation defined as data-driven decision processes based on sensed data that have quantifiable returns on investment, remains a bottleneck across an entire class of problems such as infrastructure management, societal resilience, and smart cities.

The focus of the research presented in this dissertation was to build a broad and rigorous technological, analytical, and methodological background necessary to sense for decisions with respect to two of the field's prevailing challenges: decision making for infrastructure asset management and community-governed decision making. This necessitated developing deliberate pathways to sense for decisions in physical and social systems as well as in CPSS systems at the intersection of physical (CPS) and social systems. Sections 7.1 through 7.1.3 summarize the key findings presented in this dissertation.

7.1.1 Sensing for Decisions in Single-Asset Intelligent Infrastructure

First, this work aligned targeted sensing strategies with decisions governing infrastructure asset management by extending reliability methods—which have matured in civil engineering applications over the past several decades and are used to define structural design codes—to enable data-driven decision-making practices that reflect those used in practice today. First, the reliability index was used as a scalar measure of the safety of an asset (i.e., probability of failure) to track structural condition and trigger upkeep decisions in structures where fatigue is a controlling failure mechanism. As an illustrative example, a wireless SHM system was installed on the Harahan Bridge and long-term response data was used in a reliability framework to continuously track the fatigue life of critical eyebar assemblies. The proposed reliability-based SHM framework was then generalized to formally and more broadly link structural monitoring data with condition ratings, which are the starting point for infrastructure asset management decisions made in practice today. While reliability methods have historically quantified safety with respect to a single limit state (e.g., yielding), this dissertation demonstrated that there exist measurable reliability index values associated with "lower" limit states well below failure that can be described by each existing condition rating. This allowed for monitoring data to be used to assign condition ratings based on quantitative information encompassing the entire measurable domain of damage, as opposed to those informed only by visual inspection. The proposed methodology served as the first-ever SHM framework to explicitly map monitoring data to actionable decisions and was validated using a wireless SHM system installed on the Telegraph Road Bridge (TRB) (Monroe, MI).

Chapter 2 presented a reliability-based SHM framework that provides structural owners with actionable information associated with the safety of their assets given the current (potentially deteriorated) condition of their structures. This purely data-driven framework relies on long-term monitoring data to offer a quantitative assessment of structural safety with respect to remaining fatigue life. A significant contribution of this work was that it implemented a novel holistic approach to tracking fatigue accumulation in which the sensing design empowers full consideration of the dead and live loads and associated changes in boundary conditions. The contribution of changes in boundary conditions and secondary deterioration mechanisms on structural safety is critically important, yet often ignored due to the emphasis of existing sensing strategies on monitoring global structural behavior. Additionally, the proposed framework can be carried out without the need for approximate methods (e.g., train parameter estimates, FE model simulations, controlled loading tests) during the monitoring period. To augment this theoretical framework, Chapter 2 presented the application of a fully automated wireless sensing system on the Harahan Bridge that was used to assess probabilistic fatigue life coupled with relative tautness of a critical tensile eyebar assembly. The reliability index values for both the element (i.e., individual eyebars) and system (i.e., full eyebar assembly) reliability problems were assessed and indicate that, under the conservative assumption that progressive failure is brittle, first failure within the parallel eyebar system is generally equivalent to system failure. The proposed method also serves as an intervention strategy that can quantify the influence of eyebar realignment maintenance efforts on the future evolution of the reliability index.

Chapter 3 then drew upon the advancements made in the development of the decision-driven monitoring system implemented on the Harahan Bridge in Chapter 2 to develop a more generalized SHM framework to sense for decisions across diverse infrastructure assets. This work was motivated by the observation that while SHM has shown continued growth over the past several decades, there is a persistent chasm between SHM and the ability of structure owners to make asset management decisions based on SHM data in practice. This is in part due to the historical SHM paradigm cast as a problem of estimating structural state and detecting damage by monitoring changes in structural characteristic properties (namely, reduced stiffness). In reality, for most operational structures, deterioration does not necessarily correspond to changes in structural properties with structures operating in their elastic regimes even when deteriorated. For structures like bridges, upkeep decisions are based on federally mandated condition ratings assigned during visual inspection. Since condition ratings are widely accepted in practice, this dissertation proposed that condition ratings serve as lower limit states (i.e. limit states below yielding) with long-term monitoring data used to quantify these lower limit states in terms of the reliability index. In order to link structural monitoring data with existing condition ratings using reliability methods, a method was offered to quantify the reliability index values, β_i^r, corresponding to the lower limit states described by condition ratings (if there are ten condition ratings, $i = 0,1, \ldots ,9$). This explicitly assigns a quantifiable level of safety and performance to each decision that is currently made in practice after inspection. Once the reliability index threshold values are established for a set of condition ratings, the data-driven (measured) reliability index, β_m^r, of the in-service asset can be monitored continuously and explicitly mapped to a condition rating to track the structural condition of the asset at any point in time and to trigger upkeep decisions based on condition ratings.

The proposed framework was successfully carried out using long-term monitoring data collected from the Telegraph Road Bridge. The application of the proposed framework to one of the TRB's pin-and-hanger assemblies demonstrated that long-term monitoring is necessary to assess the influence that diverse types of damage have on the safety of a structural assembly. For hanger plates, these damage cases are historically tied to corrosion-induced locking at the pin-plate interface, and in-plane and out-of-plane bending. To illustrate the importance of this finding, the author compared the reliability index value and corresponding condition rating derived from the proposed data-driven method to the conclusions drawn from existing visual and ultrasonic-based methods of inspection when the net-section stress limit state is considered with both section loss and locking at the pin-plate interface. The results indicated that current inspection methods can fail to quantify and assess the impact that underlying damage and changes in boundary conditions can have on the safety and performance of the pin-and-hanger assembly. As a result, deterioration that increases the probability of failure of a structural assembly can go unnoticed during condition rating inspections.

7.1.2 Overcoming Technological and Analytical Barriers to Scaling Out

One of the foremost challenges faced by wireless SHM systems whose operation relies on energy harvested from the environment (such as the solar-powered SHM systems installed on the Harahan Bridge and TRB) is that such systems are subject to stringent energy constraints. There is an inherent tradeoff between the quality of a measured process' parameter estimates and the energy-constrained communication. Consequently, when decision-making processes rely on statistical estimations of performance, the utility of data should be considered to optimize the data collection and transmission processes given the system constraints. In order to enable sensing for decisions in applications where energy is an uncertain and limited resource, this dissertation derived a novel stochastic data collection and transmission policy for wireless sensing networks that minimizes the variance of estimated component parameters of a measured process subject to system constraints. Numerical results based on one year of data collected from the TRB illustrate the gains achieved by implementing the optimal policy to obtain response data used to estimate the reliability index.

Chapter 4 derives a stochastic data collection and transmission policy that minimizes the variance of estimated component parameters of a measured process subject to constraints imposed by a WSN node's energy and data buffer sizes, stochastic models of the incoming energy and event arrivals, the value of data, and temporal death. By controlling data collection within a transmission subsystem to optimize an objective governed by remote parameter estimation, data collection and transmission is automated to facilitate decision making. This work modeled a replenishable WSN node as a continuous-time Markov chain and derived a single unique threshold value governing an event-based policy that is independent of the energy buffer's SoC and places no restrictions on the size of the energy and data buffer sizes. The derived optimal threshold value produces the best possible estimate (i.e., minimum component variance) of the process parameters using MLE given the system constraints. Numerical results are presented to reflect the objectives of the theoretical results and to illustrate that the proposed framework is robust against uncertainty in estimates of the process' parameter estimates at the outset of the monitoring period.

This work differentiates itself from, and improves upon, existing literature in two notable ways. First, the proposed policy accounts for—and places no restrictions on the size of—a WSN architecture's data storage buffer. Data transmission is often the most significant source of energy consumption in a wireless sensing node; accounting for the storage of (potentially) large amounts of data in a buffer that are communicated to the remote estimator in batch transmissions—as opposed to transmitting each value individually immediately upon collection—leads to significant gains in the transmission rate. Consequently, the proposed policy can be used in diverse applications requiring wide-ranging hardware specifications, from sampling continuous measured processes at high rates (>100 Hz) in SHM applications to low-rate environmental applications. Second, this work represents the first effort to derive a stochastic policy that governs data collection and transmission with the purpose of explicitly linking optimal remote parameter estimation to a rigorous stochastic energy recharging model that accounts for all significant hardware constraints.

Chapter 5 then drew upon the primary result from Chapter 3: structural monitoring data can be directly incorporated into decision-making processes for asset management using reliability methods, where the response data at each sensor location is characterized by its statistical parameters. This suggests that there is a need for an automated data collection and transmission strategy that facilitates the characterization of the statistical parameters of structural response data with minimum variance so that bridge owners can increase the frequency with which they track structural condition without compromising accuracy. In the context of Chapter 3, this refers to the ability of bridge owners to update data-driven condition ratings based on the highest quality parameter estimates of the measured input processes.

The work presented in Chapter 5 drew upon the theoretical basis presented in Chapter 4 and extended the optimal data collection and transmission policy to an SHM application on the TRB, which was introduced in Chapter 3. The monitoring system on the TRB stands to benefit from the implementation of such a policy because the WSN utilizes sensing nodes that operate using harvested solar energy and are subject to stringent energy constraints due to the size of the solar panels and availability of incoming energy, geographic location, and battery size. Because the optimal policy was simulated based on historical data collected from June 2016 to June 2017—for which there was no measured data to inform the recharge and event arrival rates during the monitoring period—the recharge and event arrival rates were modeled based on continuous WIM station data collected along Interstate 275 to the north of the TRB as well as surface weather observations recorded by NOAA at a nearby station. In order to illustrate the gains achieved by implementing the proposed optimal policy, the transmit-all, best schedule-based, and optimal policies governing data collection and transmission processes were simulated over the one-year monitoring period, where the strain response measurements corresponding to each truck were drawn from the true process measured by bridge response data. The results confirmed that the implemented optimal policy minimizes the variance of the estimated component parameters subject to the system constraints. Consequently, bridge owners can update the reliability index more frequently and accurately.

7.1.3 Sensing for Decisions in Social Urban Systems

Finally, this dissertation extended the work performed in wireless sensing and sense-for-decision frameworks by exploring their role in community-based decision making. This dissertation posed societal engagement as a necessary entry point to urban sensing efforts in smart cities. Members of under-resourced communities are often stakeholders who are the most challenged with technology and the most vulnerable to lack of access to data and information. This dissertation offered a novel, low-power wireless sensing architecture that functions as a user-friendly urban sensing solution that communities can rapidly deploy to understand urban processes that are critical to informing their views and to guide their decision making. Applying this platform, this dissertation presented transformative work to "democratize" data by enabling members of vulnerable communities to easily use these generic urban sensors to collect data and generate insights on issues that tie directly to their unique needs and the diverse ways in which they use their cities.

Chapter 6 detailed the development of the *Urbano* sensing platform, which supports interoperability among diverse arrays of heterogeneous IoT devices, preserves privacy and trust among citizens, supports cloud-based analytics, has a user-friendly design, and supports low-power and low-cost sensing and communication to free nodes from a fixed power source. In order to illustrate the applicability and usability of the *Urbano* to diverse stakeholders, a number of smart city engagements using the *Urbano* sensor node were presented. These applications incorporated a wide variety of stakeholders including city governments, local organizations, urban youth, and communities. The first test case highlighted the use of GPS-enabled *Urbano* nodes to track food trucks in Grand Rapids, MI to assess compliance with permit rules and curbside management by city planning officials. The second application deployed *Urbano* nodes to monitor pedestrian traffic along the Detroit waterfront; this data is desired by the Detroit Riverfront Conservancy to understand utilization of public spaces to guide future investments.

Given these successful deployments, a program called *Sensors in a Shoebox* was developed and implemented at full-scale in Detroit, MI. By involving Detroit youth (grades 8-12), *Sensors in a Shoebox* aimed to encourage young, vulnerable populations to engage with their communities as citizen scientists. Students engaged in problem-based inquiry, devised sensing solutions, and interpreted collected data drawn from their own urban settings. Students collected their own data in their neighborhoods and cities using the *Urbano* platform and partnered with city governments and local organizations to address the challenges faced. Introducing and situating the IoT *Urbano* platform within a larger community engineering educational program offered one potential map through which we can begin to develop a sustainable CPSS. Looking across one iteration of the *Sensors in a Shoebox* program, the youth were meaningfully engaged in the engineering design and research process, from beginning to end. By creating programming in which the youth were engaged in and drove the entirety of a design cycle situated in their communities, more opportunity was created for students to see engineering and technology as relevant to their lives and transformative in their communities. This suggests that youth made connections between their research work, their design work, and resulting actions made by the City. Further, it suggests the need for more widespread structures that help to establish these connections by creating educational programming foundations (through clubs, community organizations, service centers etc.) that provide user-friendly technological innovation (i.e., the *Urbano* IoT platform) with the infrastructure through which to support meaningful change in a community.

7.2 Future Research

7.2.1 Intelligent Infrastructure

Consider the reliability-based SHM framework that was presented in Chapter 2. A key assumption is made to facilitate the development of the proposed framework: a conservative lower bound is placed on the parallel eyebar system's reliability index, which assumes that if one eyebar exceeds its limit state, then its load immediately redistributes to the other eyebars leading to a seemingly brittle failure mechanism. In reality, there remains axial load capacity in the cracked eyebar past the limit state resulting in a ductile failure mechanism. While lessening this conservative assumption may lead to economic gains, conservatism is ubiquitous across structural engineering due to the catastrophic consequences of failure. With this tradeoff in mind, future work should entail investigating the parallel eyebar system after crack initiation to more accurately identify the reliability of the parallel eyebar component. This will lead to an even more realistic approach to assessing the component reliability and will also offer some insight to the time scales over which progressive failure of the component would occur, allowing for inspection cycles to be optimally timed. Additionally, future work aimed at modelling the relationship between the reliability index and repeated eyebar realignment will help to inform optimally timed maintenance efforts that maximize the remaining fatigue life of the bridge.

In both the Chapter 2 and Chapter 3, the reliability index serves as a powerful metric because it is a direct indicator of the probability of failure of a component or system. The explicit linkage between structural condition and the reliability index has far-reaching implications enabling the translation of the proposed methodologies into widespread practice. The proposed data-driven condition ratings from Chapter 3 can be tied more concretely to owners' decisions by quantifying the cost implications of actions that are associated with each condition rating. This will extend the existing reliability-based framework to a more general risk-based decision-making framework where each element has to be assessed for risk. Here, risk is based on both the probability of failure tied to condition ratings, as defined in this dissertation, and consequences (i.e. cost of action or inaction). Such an extension will enable stakeholders to concretely understand what asset management decisions are considered financially optimal given an asset's current condition rating. While decoupled from condition ratings, a preliminary theoretical basis for such a decision-making process has been investigated by Cappello et al. (Cappello et al. 2016) and Thöns (Thöns 2018). The proposed method investigated in this dissertation sets the foundation for implementing better decision-making strategies based on risk—and even setting new, quantitatively informed criteria for condition ratings themselves—without needing to change the already established, and nationally mandated bridge management system that is founded on condition ratings. Future work will also include exploring the potential difficulties of extending the proposed methodology to other bridge components that require less straightforward analytical modeling. This is critically important because the method for assigning lower limit states to condition ratings relies on the development of a high-fidelity FE model.

7.2.2 Social Urban Systems

For WSNs within large-scale systems such as transportation and building energy systems that are subject to resource constraints, the optimal policy and theoretical framework derived in Chapter 4 should be extended from a single-hop to network-level scale. As the value of information becomes increasingly dependent on multiple wireless sensing nodes across a network, the embedded microcontroller will be used for extensive data processing as part of a larger edge-computing architecture associated with sensing in smart cities. This will necessitate moving the computation of the optimal threshold from a remote server to a WSN node's local embedded processor. To facilitate this transition, a surrogate model should be implemented to simplify the solution for the optimal threshold value to enable less complex on-board embedded processing on an 8-bit microcontroller. Additionally, the energy-aware sensing strategy should account for the fact that energy arrival and event arrival rates are rarely constant over long periods of time when WSNs rely on EH from the environment. This research would benefit from future work that incorporates predictive estimation of the recharge rate, β, and event arrival rate, λ. Short-term forecasting will also enable the integration of online supervisory control schemes that dynamically control whether the sensing architecture is accepting measurements or in a low-power sleep state in which all incoming data is rejected, even when the energy buffer is not fully depleted.

The CPSS platform presented in this dissertation offers a scalable and sustainable solution for enabling communities to help strengthen their city neighborhoods. In addition to helping to connect citizens to their communities, this platform provides a scientific and technological foundation aiding the extension of CPS to include humans. However, while CPSS holds tremendous promise in connecting urban citizens to their cities, there remain many challenges including how to integrate human-centric data and human-in-the-loop control solutions into CPS frameworks.

Moving forward, future work will necessitate understanding the role that sensing, data, and actuation can play in promoting equity and enabling social mobility by deeply rooting policy-making, urban planning, and governance within community-driven data collection. This will require the fusion of data produced by interdependent systems (including rigorous consideration of human action within these systems) and require addressing a number of technical and analytical challenges due to the increased demands on automated sensing, computation, and control in under-resourced areas. This ties directly to the idea of sensing for decisions because the field of civil and environmental engineering is becoming increasingly centered on understanding and designing engineering systems based on the needs of the human users and the ways they use infrastructure services and make decisions in their communities. Since human development is highly dependent on feedback mechanisms that are enabled by both social (e.g., education) and physical (e.g., public spaces) components, a primary extension of this work will focus on 1) identifying and quantifying critical feedback mechanisms that exist within the built environment that are either directly or indirectly related to factors impacting social mobility; and 2) using intelligent systems (i.e., sensing, data, and actuation) to close these feedback loops when existing mechanisms are not sufficient to enable a stakeholder (e.g., student, community) to reach its desired output.

7.3 References

Cappello C, Zonta D and Glisic B. Expected utility theory for monitoring-based decision making. *Proc IEEE* 2016; 104(8): 1647-1661.

Thöns S. On the value of monitoring information for the structural integrity and risk management. *Comput-Aided Civ Infrastruct Eng* 2018; 33: 79-94.