FREQUENTIST AND BAYESIAN APPROACHES FOR PROBABILISTIC FATIGUE LIFE ASSESSMENT OF HIGH-SPEED TRAIN USING IN-SERVICE MONITORING DATA

XIAO WANG

Ph.D

The Hong Kong Polytechnic University

2018

THE HONG KONG POLYTECHNIC UNIVERSITY

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

FREQUENTIST AND BAYESIAN APPROACHES FOR PROBABILISTIC FATIGUE LIFE ASSESSMENT OF HIGH-SPEED TRAIN USING IN-SERVICE MONITORING DATA

Xiao WANG

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

November 2017

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

Xiao Wang (Name of student)

To my family

for their love and support

Abstract of Thesis

Entitled

FREQUENTIST AND BAYESIAN APPROACHES FOR PROBABILISTIC FATIGUE LIFE ASSESSMENT OF HIGH-SPEED TRAIN USING IN-SERVICE MONITORING DATA

Submitted by

Xiao WANG (B.Eng., M.Sc.)

for the Degree of Doctor of Philosophy

at The Hong Kong Polytechnic University

in November 2017

Rapid development of high-speed rail has made high-speed trains one of the most favorable transportation means in recent decades. With the extension of in-service period, train bogie components experience more and more stresses due to the cyclic loadings during routine high-speed train journeys, resulting in cumulative fatigue damage and eventually leading to fatigue failure. With safety being the paramount factor for public transport, the fatigue life of a high-speed train needs to be properly and reliably assessed. As a paradigm of integrating on-board monitoring data into condition assessment of train components, the continuously collected in-service strain data from on-board monitoring systems is significantly useful for evaluation of the fatigue life of high-speed trains. However, the

complex operating environment of high-speed trains indicates that various types of uncertainties exist in the monitoring data and established models. This phenomenon brings forward a challenge in obtaining fatigue life assessment results with sufficient accuracy and precision. The research in this thesis is dedicated to developing a probabilistic methodology for integrating in-service monitoring data into fatigue life assessment of high-speed trains. It consists of two branches, as described in the following.

One branch leads to the investigations of monitoring-based fatigue life assessment of high-speed trains in the framework of frequentist probabilistic inference. An initial effort is made to address stress modeling technique to formulate probability density functions (PDFs) for both stress range and mean stress, which will be used for fatigue life analysis of train components in light of in-service monitoring data. For modeling of the multi-modal PDF of mean stress acquired from in-service stress history, different types of finite mixture distributions and hybrid mixture parameter estimations are employed. Using the Akaike information criterion (AIC) or Bayesian information criterion (BIC), the best fitted models for the PDFs of stress range and mean stress are obtained, respectively. Based on this modeling technique, the joint PDF of stress range and mean stress is derived through a frequentist modeling technique. Further, this technique provides a foundation for the establishment of a frequentist fatigue life assessment method. In this method, mean stress effect is quantified through a newly defined indicator Q. This indicator is then introduced to formulate a fatigue life assessment method based on the stress-life (S-N) theory. Then the PDFs of both stress range and Q are established and jointly determined with the Miner's

cumulative rule to evaluate the fatigue life of high-speed train components with the use of on-board monitoring data. For verification purpose, the method is adopted to assess the fatigue life using the monitoring data acquired from an in-service train running on a high-speed railway in China.

The other branch of the present study explores the development of a probabilistic fatigue life assessment approach for in-service high-speed trains in the framework of Bayesian inference. Firstly, Bayesian modeling technique for the stress spectrum of monitoring data is introduced. Analytical solution is used to estimate the derived PDF for the stress spectrum. By integrating prior knowledge and continuous monitoring data, the PDF of the stress spectrum can consider the uncertainties in the continuous monitoring data to predict and update the modeling results. The performance of the technique is shown through the estimated model parameters, model errors and the associated uncertainties. With the formulated Bayesian modeling technique, a fatigue life assessment method is proposed for high-speed trains under in-service environment. Uncertainties in the component material and in-service monitoring data can be considered in deriving the posterior PDF of the fatigue life. This method is verified with the monitoring data of a welded component of a train running on a high-speed railway in China. In addition, as stress concentration largely affects fatigue life of a welded joint, the hot spot stresses obtained from the monitoring data are also integrated into the method. Assessment and prediction results show capability and good performance of the proposed method. This Bayesian method is especially implementable to the fatigue life assessment of a high-speed train when the monitoring

period becomes sufficient that prior information and continuous monitoring data can be obtained.

LIST OF PUBLICATIONS

Journal Papers

Ni, Y.Q., **Wang, X.** and Zhou, L. (2017). "Bayesian approach for fatigue life assessment of high-speed train using in-service monitoring data", *International Journal of Fatigue*, under review.

Wang, X., Ni, Y.Q. and Zhou, L. (2017). "Analysis of influence of mean stress on fatigue life of in-service high-speed train bogies using monitoring data", *Mechanical Systems and Signal Processing*, under review.

Ni, Y.Q., **Wang, X.** and Ye, X.W. (2017). "Fatigue life assessment for high-speed train bogies using in-service monitoring data", *Fatigue and Fracture of Engineering Materials and Structures*, under review.

Wang, X., Ni, Y.Q. and Lin, K.C. (2015). "Comparison of statistical counting methods in SHM-based reliability assessment of bridges", *Journal of Civil Structural Health Monitoring*, 5(3): 275-286.

Patent

Ni, Y.Q., Sun, Q., Lai, S.K., Yuan, M., **Wang, X.**, Liu. X.Z. and Wang, J. (2017). *A New Design of Bumpers Embedded with High Damping Materials for Bogie Frame of High-speed Trains* (pending). Chinese Patent Application No.: 201710778502.0.

Conference Papers

Wang, Y.W., Ni, Y.Q. and **Wang, X.** (2017). "Detection of performance deterioration of high-speed train wheels based on Bayesian dynamic model", *Proceedings of the 11st International Workshop on Structural Health Monitoring*. Stanford, USA.

Ming, Z., **Wang, X.**, Wang, D. and Ni, Y.Q. (2017). "Towards a cloud-based data visualization system for online condition assessment of in-service high-speed train", *Proceedings of the 2017 World Congress on Advances in Structural Engineering and Mechanics*. Seoul, Korea.

Wang, X., Ni, Y.Q., Zhang, L.H. and Sun, Q. (2016). "Understanding of dynamic interaction of an in-service high-speed train via onboard monitoring", *Presented at the 1st International Workshop on Structural Health Monitoring for Railway System*. Qingdao, China.

Zhang, L.H., Ni, Y.Q., **Wang, X.** and Liu, X.Z. (2016). "Monitoring of dynamic behavior and ride quality of an in-service high-speed train under different operation conditions", *presented at the 1st International Workshop on Structural Health Monitoring for Railway System*. Qingdao, China.

Wang, X. and Ni, Y.Q. (2016). "Analysis of the influence of mean stress on the fatigue life of in-service high-speed train bogies using monitoring data", *Proceedings of the 3rd International Conference on Railway Technology: Research, Development and Maintenance*. Cagliari, Italy.

Ni, Y.Q., **Wang, X.** and Ye, X.W. (2015). "Probabilistic fatigue assessment for high-speed train bogies using in-service monitoring data", *Proceedings of the 13th International Conference on Railway Engineering*. Edinburgh, UK.

ACKNOWLEDGEMENTS

First of all, I would like to express my profound gratitude to my chief supervisor, Prof. Yi-Qing Ni, for his high sense of responsibility, valuable guidance, and constant support throughout the whole study. I thank him especially for his patience, constructive criticism and illuminating instruction. His conscientious attitude in scientific research is the most valuable lesson I have learned from this PhD. It is my great honor to have him as my advisor. I would also like to express my sincere appreciation to my co-supervisors, Dr. Dan Wang and Dr. Xinjian Jin, for their generous guidance and discussion.

The financial supports from the Research Institute for Sustainable Urban Development of The Hong Kong Polytechnic University and the National Rail Transit Electrification and Automation Engineering Technology Research Center (Hong Kong Branch), Hong Kong, are sincerely acknowledged.

It is with great sincerity that I thank Ms. Elaine Anson and Dr. Kai-Yuen Wong, for their great inspiration and encouragement on my English writing skills and system design for on-board monitoring. I also feel indebted to Prof. Otto Lin, Prof. Jan-Ming Ko, Prof. Jianchun Li, Prof. Derek Siu-Wing Or, Prof. Wenzhong Zhao, Dr. Joao Pombo and Dr. Ken Siu-Kai Lai, for their kindness in sharing their research experience which broadens my horizons about the academic world.

I would also like to thank Dr. Songye Zhu, Dr. You Dong, Mr. Wanglin Wu, Dr. Lu Zhou, Dr. Michael Shun-Yee Liu, Dr. Davis Cheung and Dr. Qiang Sun for providing technical guidance and resources for my research during this PhD study.

In addition, I am grateful to all of my friends and colleagues at The Hong Kong Polytechnic University, particularly Dr. Huaping Wan, Mr. Qi Zhang, Mr. Xiaozhou Liu, Dr. Cuidong Xu, Dr. Maodan Yuan, Dr. Junfang Wang, Dr. Youwu Wang, Mr. Chao Zhang and Ms. Siyan Liu for their very precious helps and suggestions.

Finally, I dedicate this thesis to my beloved parents with gratitude for their everlasting love, encouragement, and support.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

The main symbols used in this thesis are listed below:

Chapter 2

S	Stress range in *S-N* curve
N	Number of cycles to failure
σ_f'	Fatigue strength efficiency
b	Fatigue strength exponent
C	Constant in equation of *S-N* curve with exponent *m*
m	Exponent of *S-N* curve
D	Cumulative fatigue damage ratio
n_i	Number of applied stress cycles at *i*th stress range
S_i	Stress range of *i*th class in a stress spectrum
k	Number of classes in a stress spectrum
$\Delta\varepsilon_p$	Plastic strain
ε_f'	Ductility coefficient
c	Ductility exponent
$\Delta\varepsilon_T$	Strain value comprising elastic and plastic strain
A, r	Material factor in Paris power law
a	Crack length
ΔK	Range of stress intensity factor
ΔK_{th}	Threshold value below which no crack growth
σ_a	Alternating stress
σ_m	Mean stress
σ_{min}	Maximum stress in a stress cycle
σ_{max}	Minimum stress in a stress cycle

S_a	Stress range in stress spectrum

Chapter 3

$f(x\|\Theta)$	Predictive mixture density with parameter $\Theta = (w_1,\ldots,w_m;\ \theta_1,\ldots,\theta_m)$
$f_j(x\|\theta_j)$	Predictive component density with a given parametric family indexed by scalar or vector parameters θ_j
w_j	Weight for each component density
k_l	Number of component in mixture distribution
p, t, r	Parameter in constant life equation
S_e	Effective stress range for each stress cycle
S_u	Ultimate tensile strength of material
μ_j, σ_j	Mean and standard deviation for either Gaussian or lognormal mixture distribution in ith component
γ_j, β_j	Scale parameter for either Weibull mixture distribution in ith component
λ	Lagrange multiplier
l_l	Maximum likelihood
m_l	Number of estimated unknown parameters
n	Number of observations
R	Stress ratio
k_l	Number of classes in stress spectrum

Chapter 4

Q	Indicator to quantify effect of mean stresses
S_i	Stress range of ith class in a stress spectrum
k	Number of classes in stress spectrum
D_f	Fatigue damage ratio to failure
$S_{l,i}$	Integrated stress range for ith of k classes
S_l	Integrated stress range

$f(S_l)$	PDF of integrated stress range
$f_1(s)$	PDF of stress range
$f_2(q)$	PDF of indicator Q
n_{tot}	Cumulative factor relating to number of cycles
N_f	Number of stress cycles to failure
λ_i	Reduction factor to process stress cycles below knee point
$\Delta\sigma$	Knee point value according to IIW standard
D_m	Fatigue ratio for a set of measured strain time history
F	Fatigue life (unit: distance)
m_1	Exponent of *S*-*N* curve above knee point
m_2	Exponent of *S*-*N* curve below knee point
L_s	Design fatigue life of a typical high-speed train

Chapter 5

s	Stress spectrum of monitoring data
$P(s\|\Theta)$	Likelihood of observation dataset
μ	True response of stress range
ε	Noise in measured data
s_i	Stress range of *i*th moment in a stress spectrum
μ_i	True response of *i*th moment in a stress spectrum
ε_i	Noise of *i*th moment in a stress spectrum
$\overline{s}$	Sufficient statistics of μ
g	A parameter used to describe likelihood of stress spectrum
$F(x)$	Cumulative density function
$IG(x)$	Inverse-gamma distribution
$N(x)$	Normal distribution
$\Gamma(x)$	Gamma distribution
μ_t	Mean of *Student-t* distribution
σ_t	Scale parameter of *Student-t* distribution

v	Degrees of freedom
s_j	jth dataset of monitoring data
W	Number of monitoring datasets

Chapter 6

P	Maximum stress range obtained during a train trip
β, σ^2	Parameters used in logarithmic expression of S-N relationship
μ, σ_t^2	Parameters used to describe distribution of stress spectrum
T, M	Number of iterations using MCMC technique
s^*	Predicted distribution of stress spectrum
$P(F)$	Predicted fatigue life results using Bayesian inference
N^*	Predicted number of cycles to failure in S-N curve for S^* stress range

LIST OF ABBREVIATIONS

AIC	Akaike information criterion
BIC	Bayesian information criterion
BDM	Bayesian dynamic model
CFD	Cumulative fatigue damage theory
EM	Estimation-maximization algorithm
FCP	Fatigue crack propagation
FAT	Fatigue class according to IIW standard
FBG	Fiber Bragg grating
FEM	Finite element method
HCF	High-cycle fatigue
HSR	High-speed rail
IIW	International Institute of Welding
LCF	Low-cycle fatigue
MCMC	Markov chain Monte Carlo simulation
MH	Metropolis-Hastings algorithm
NDT	Nondestructive testing
PDF	Probability density function
PSD	Power spectrum density
SHM	Structural health monitoring

CHAPTER 1

INTRODUCTION

1.1 Overview

High-speed rail (HSR) is generally considered one of the safest and most efficient of transport modes, consuming far less energy per capita than other major means of transportation, while still maintaining a high travel speed (INFARS/IWW 2000). High-speed rail is also environmentally friendly. It is estimated that the carbon footprint of HSR can be up to 14 times less carbon intensive than conventional land transport and up to 15 times less than air freight even when measured over the full life cycle of planning, construction and operation (UIC 2011). The speed of HSR has been increasing since 1991 and has literally become the most preferred option of transport travelling between major cities, with the top speed reaching over 350 km/h. The rapid development of HSR over recent decades builds up its key role in social development and economic growth (Hotchkiss and Larney 2011; Platt and Steger 2014; Jackson 2015). HSR contributes profit growths in many fields including tourism, restaurants, hotels, property, rolling stock, railway infrastructure, car rental, and leads to unprecedented changes in every aspect of daily life. In the next decade, the HSR systems in China will increase the mobility of the population, affect over 700 million

people, contribute to generation of four new super-city clusters and further strengthening of the two existing super-cities, and exponentially boost central city traffic capacity (Lou et al. 2011). HSR is seen as bringing significant positive progress to social, economic and personal life on a global scale.

1.1.1 Background

High-speed rail is regarded as one of the most significant technological breakthroughs of public transport developed in the second half of the 20th century (Campos 2009). Givoni (2006) defined HSR as "high capacity and frequent railway services achieving an average speed of over 200 kph". More elaborate references relating to high-speed trains exist in the documents by the European council (2001), UIC council (UIC 2015) and Chinese National Railway Administration (NRAC 2014). Since the first high-speed train passenger service was launched on the 1st of October 1964, there has been a steady rapid growth, particularly in recent decades. As shown in Figure 1.1, it is expected that total length of the world HSR route will be over 50,000 km by 2030 (Leboeuf 2016). Taking the HSR of China as an example, the total length of railway tracks nationwide was only 370 miles at the beginning of the 20th Century (Smith and Zhou 2014), but it now possesses over 60% of the total HSR route length of the world (UIC 2017). As shown in Figure 1.2, it is also envisioned by the authorities that high-speed railways will cover the majority of areas of the whole countries of both China and the United States of America by 2030 (Platt and Steger 2014; NRAC 2016).

In the meantime, HSR also helps to open up new opportunities in many areas, through which cities are connected around the world, with trade promoted and experience improved by shortening time of travel (Urena et al. 2009). A case study by Morgan Stanley research group on the traveling time and fees by flight and high-speed train found that in comparison to traveling between intermediate cities by flight, the traveling time by high-speed train is similar in most cases, while the fares for high-speed train tickets are generally much cheaper (Lou et al. 2011). For instance, from Wuhan to Guangzhou, two major cities in China, a high-speed train ticket costs less than one third of the airfare (i.e., 490 RMB versus 1700 RMB).

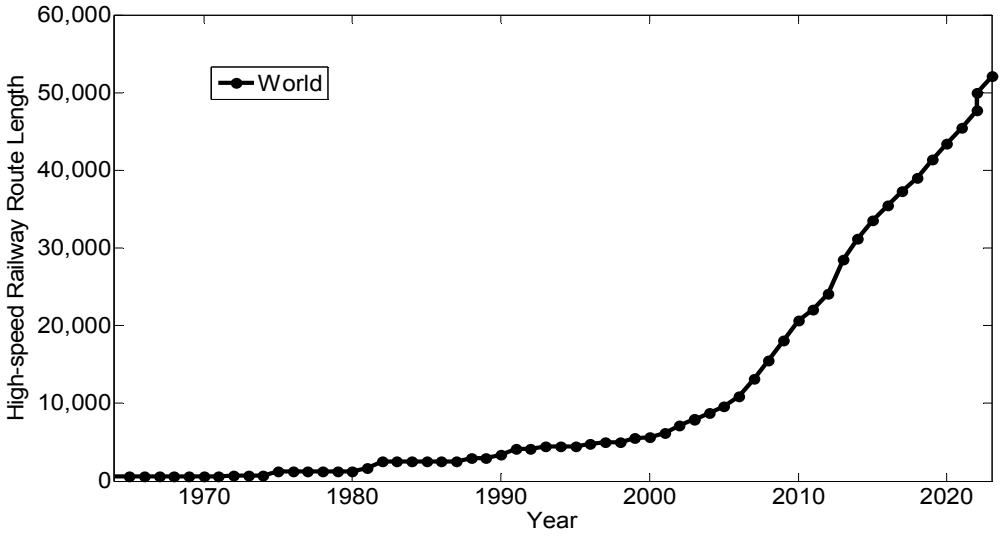
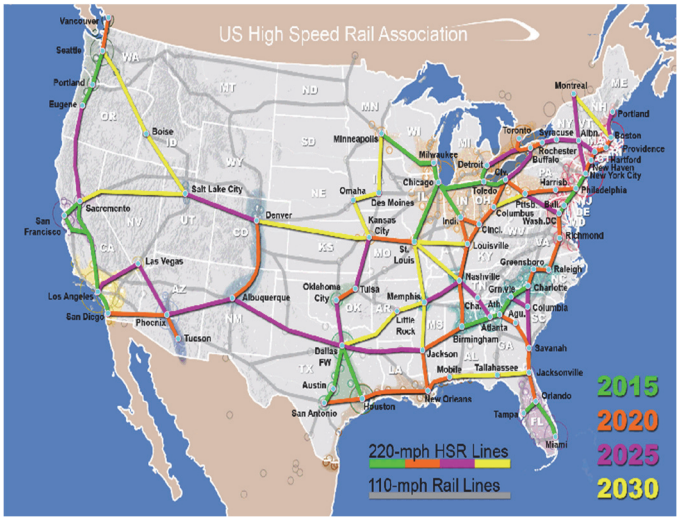


Figure 1.1 Worldwide HSR route length (Leboeuf 2016)

(a)



(b)

Figure 1.2 Plans of HSR routes by 2030 from authorities of (a) China (NRAC 2016) and (b) the United States (Platt and Steger 2014)

Nevertheless, due to the limited service life of high-speed trains, the cost of maintenance is a growing burden that affects the overall social benefit (Campos 2009). The total maintenance fee throughout the entire service life could be as expensive as the manufacturing cost of a high-speed train. For example, in European countries, the manufacturing cost of a high-speed train is approximately €30 - 32 million EUR. Given

that the life cycle for high-speed train rolling stock is 30 years, the annual maintenance cost for such a train, traveling roughly 500,000 km per year, can be €1 - 1.2 million EUR and the full-life maintenance cost may be equal to or higher than the manufacturing cost of a high-speed train (Barron 2016). Thus, improvement on high-speed train entire service life is crucial in that it can significantly reduce the maintenance cost and save the overall cost expenditure.

The entire service life of a high-speed train is determined by fatigue problems, particularly affected by insufficient fatigue life of train components under long-term operating conditions. Limited fatigue life of high-speed train results in higher maintenance cost over a shorter period. In Europe and China, the fatigue life of a typical high-speed train is 30 years, whereas this number is only 17 years in Japan (Barron 2016). One reason for the shorter fatigue life of Japanese high-speed trains is that the railway tunnel cross-sections for high-speed trains are narrower in Japan, which causes massive compression and decompression on train components resulting in more serious fatigue problems.

Analysis on fatigue problems of high-speed trains helps optimize the operating costs of high-speed railway line. Taking the early high-speed trains used in European countries as an example, the trains started to operate in the 1980s with an expected fatigue life span of 30 years. Up to now, their fatigue lives are reaching the maximum requirement and should be renovated. In order to reduce the cost of renovating these high-speed trains, railway authorities investigated to understand the fatigue problems

of these high-speed trains. The investigation showed some of the high-speed trains still own fatigue lives of 10 years or even more and do not need to be renovated. However, it was also revealed that there are still roughly 100 trains that started operation from 1981 and another 110 trains that started operation from 1989 in need to be renewed due to their train components suffering from fatigue problems. In forthcoming decades, more than 1600 high-speed trains of the world produced in recent decades will also have to be renewed, which would generate high costs to local governments and society (Barron 2016). It is stressed in Barron's report that, to improve economic benefits of HSR and its corresponding society issues, further research relating to fatigue life of high-speed trains is essential.

Another motivation to conduct research about the fatigue life of high-speed trains relates to safety concerns: after a long period of service life, the potential deterioration of the general performance of train components may be the results of metal fatigue problems, which may have devastating consequence, particularly in the case of in-service high-speed trains. It was reported that in a routine inspection in Jinan, China, a 7.11 mm (length) × 2.4 mm (depth) fatigue crack was discovered in a railway axle at a location near the gearbox (Ifeng 2011). The German Eschede train derailment in 1998, the world's most severe high-speed train disaster with 101 lives claimed and 88 injured, was attributed to cracks in the wheel rims caused by fatigue (500,000 wheel revolutions/day) (NASA 2007). In fact, it has been found that fatigue accounts for 80% or more of all "in-service" failures (Milne et al. 2003). Hence, with safety being the

paramount concern for public transport, the fatigue life of a high-speed train must be properly and continually assessed.

In summary, a proper and feasible fatigue life assessment method for high-speed trains is important for the improvement of train safety and the optimization of social and economic benefits from this emerging technology. In this study, an investigation into the development of fatigue life assessment approaches of high-speed trains under operating conditions has been pursued. A brief introduction relating to metal fatigue will be firstly given in the following subsection.

1.1.2 Understanding of fatigue phenomenon

Fatigue occurs when a metal component suffers from repeated loading and unloading cycles (Stephens et al. 2001). It is a progressive and localized process in which irreversible damage accumulates continuously due to the repetition of external loadings. These loadings, for example for high-speed train bogies, include wheel-rail interaction, traction and breaking forces transmitted to the bogies. The stress levels generated by these loadings may be much lower than the yield strength of train components (i.e., ultimate tensile stress limit, or the yield stress limit), but they cause the majority of in-service fatigue failures on metal components. Typically, a component suffers from fatigue problem when it receives a cyclic load, leading to initiation of a fatigue nucleus crack on a microscopical scale. The crack gradually grows to a macroscopic size under repeated loading till the component fails which

marks the termination of its fatigue life. The microcrack starts from a perfect free surface and can have a sub-micron crack length (ranging from 1 μm to 1 mm) which is very difficult to be detected by normal equipment. Only the macrocrack with a crack length larger than 1 mm is in general detectable.

From the view of material science, fatigue life consists of two periods: crack initiation and crack growth (Schijve 2003). In the context of engineering design practice, the crack initiation period dominates the fatigue life of a component, while the crack growth period is generally brief once the crack appears. During the crack initiation period, fatigue phenomenon is regarded as repetition of inhomogeneous stresses imposed on the components providing the source of micro cracks. Based on this, a life accumulation rule was established and was frequently employed as a fundamental principle for metal component fatigue life assessment (Buch 1988). More specifically, the fatigue life assessment using this principle for crack initiation period is also called *S*-*N* curve approach. The approach firstly establishes *S*-*N* curves of the relationship between the stress range (*S*) and the corresponding number of cycles (*N*) when failure occurs, so as to describe the fatigue life of a metal component under a series of constant-amplitude stresses. These *S*-*N* curves are usually obtained from laboratory experiments on the components. Secondly, a "relative Miner rule" (Buch 1979) is applied to evaluate the fatigue life of the component, in which the fatigue life under stress cycles of each amplitude level is separately counted and quantified as a value called "fatigue ratio". The fatigue ratios calculated under the stress cycles of all

amplitude levels together form an indicator of entire fatigue life. A detailed explanation of the *S*-*N* curve-based principle can be found in Stephens et al. (2001).

1.2 Problem Statement

Because fatigue analysis has mainly been based on deterministic approaches, particularly for fatigue life analysis in material science, much work has been done under well-controlled laboratory conditions. Fatigue life evaluation using *S*-*N* curve principle is established through laboratory experimentation. The test rig applies pre-configured stress levels on the test specimens. These stresses are repeatedly imposed on the test specimens until fatigue failure occurs on the specimens' surface. The total number of imposed stress cycles is counted to be the evaluated fatigue life for the specific stresses condition. It has been often observed that the results of the evaluated fatigue life for the same stress condition are different and scattered. In addition, the stress conditions of in-service situation also vary frequently. The phenomenon of scattered fatigue life results cannot be avoided. The randomness of fatigue lives is thought to be an inherent feature in fatigue problem, even when the conditions for fatigue life evaluation are well-designed in laboratory.

Some typical models based on *S*-*N* curve have been developed from previous studies, using laboratory-based fatigue resistance data for evaluating fatigue life more precisely (Bathias et al. 2001; Kim 2006; Song et al. 2016). However, these models are only

valid under laboratory conditions with a controlled environment. In addition, when the material of train components is tested in such an environment, it is difficult to truly simulate the variation of stress distributions on the in-service train components, because of insufficient information of all factors contributing to the imposed stresses. In addition, during the whole service life, high-speed trains are exposed to time-varying environments (Zhai et al. 2009), in which the factors affecting imposed stresses change time from time. It is crucial to understand the behavior of in-service stresses and integrate the information of in-service stresses with *S*-*N* curve for fatigue life assessment.

Advanced development in sensing technologies enables the monitoring of in-service component responses for different assessment purposes (Maurin et al. 2002; Barke and Chiu 2005; Balageas 2006; Ni 2015), and it is now practical to adopt a fatigue life assessment approach using in-service monitoring data. Monitoring data is obtained from an on-board monitoring system consisting of different types of sensors installed on train components. The monitoring data can provide authentic information about physical responses of the train during train trips (Maurin et al. 2002). Currently, monitoring data acquired from sensing systems has been increasingly used for problems analysis in railway engineering field (Lee et al. 2012; Hong et al. 2014; Leung et al. 2014). It has also been widely applied for investigations on fatigue related issues, including understanding of real-time stresses or vibration characteristics causing fatigue problems (Grubisic and Fischer 2006; Tam et al. 2007; Waston and

Timmis 2011; Hong et al. 2014), failure study (Morgado et al. 2008; Hu et al. 2017) and design optimization (Chong et al. 2010; Lee and Lee 2105). One of the primary goals in the studies applying on-board monitoring of train components is to construct relatively authentic stresses or load spectra and to obtain a more accurate estimation of the real operation condition.

Some classical fatigue life assessment methods using monitoring data are included in standards and codes and are widely applied in various industrial fields (BS7910 2005; Mancini and Cera 2008; Castillo et al. 2009; Hobbacher 2009; BS7608 2014). In particular, the procedure stipulated in IIW-2008 by the International Institution of Welding (IIW) is one of the common practices in use by the railway industry to assess the fatigue life of welded joints on trains (Hobbacher 2009). Fatigue life for a high-speed train is usually calculated by integrating an imposed stress spectrum into a fatigue analysis model with defined material property parameters as specified in the code.

On the one hand, the fatigue analysis model is based on *S*-*N* curve. Although various linear and nonlinear models have been used to construct fatigue analysis models (a good list of these models can be found in Castillo and Fernandez-Canteli (2009)), the most widely used one is the linear model. In the fatigue analysis model, the values of material parameters are selected based on appropriate welded joint criteria. One the other hand, for the imposed stress spectrum, IIW provides guidance on the application of measured stress data and a data processing procedure for establishment of the stress

spectrum. In addition, there are various types of measured stresses which can be adopted for fatigue analysis of a welded joint. Although most of the specifications are in favor of using measured nominal stresses, it is also a fact that stress concentration affects the fatigue life of a welded joint (Fricke 2002; Niemi and Marquis 2002). The hot spot stress, unlike the nominal stress, relates to the stress concentration at the critical point, i.e., the weld toe, where fatigue cracks are expected to initiate and expand. The selection of either the measured nominal stress or the hot spot stress obviously influences the fatigue life assessment results.

Over these years, the community on fatigue studies has become more aware of the effect of uncertainties in stress spectrum, in fatigue properties of component material, and in assessment models on altering the evaluated result of fatigue life. For this reason, conventional fatigue life assessment approaches in the International Standards quantify uncertainties with empirical solutions and even mitigate the influence of uncertainties with selected tools (e.g., standardized laboratory experiment, on-board measurement and detailed Finite Element (FE) modeling). However, these traditional solutions are not very useful and applicable (Grubisic and Fischer 2012). In particular, (1) in practice the imposed stresses and the material parameters naturally exhibit statistical variability (Schijve 2001), and (2) in-service monitoring data can be obtained, but the pattern of the data is inherently ever-changing and exhibits time-varying characteristics due to various types of changing conditions (Grubisic and Fischer 2012). These characteristics about monitoring stresses and materials

parameters have not been properly considered by conventional fatigue life assessment approaches in the International Standards.

In this context, it is more appropriate and effective to adopt a probabilistic fatigue life assessment method which considers uncertainties in the data for train components, including the statistical scatter and stochastic nature of material parameters and in-service stresses. One of the benefits is that, the material parameters and in-service stress spectrum which are considered deterministic in conventional approaches are treated as random variables in probabilistic approaches. In addition, a probabilistic model can make use of probability density functions (PDFs) to represent the 'distribution tail' of the in-service monitoring data effectively (An et al. 2011). This enables an effective description of the statistical scatter of material parameters and the variation of actual in-service stress spectrum. In probabilistic domain, there exist two theoretical frameworks to interpret probability: frequentist inference and Bayesian inference (Bishop 2006; Ambaum 2012). A brief explanation of the frequentist and Bayesian statistics is provided in Appendix II.

1.3 Research Objectives

The first objective of this study is to develop a probabilistic approach for conducting fatigue life assessment of in-service high-speed trains using on-board monitoring data. To achieve this goal, frequentist modeling technique used to interpret the in-service

monitoring data is selected as part of the method. This technique is able to describe with PDF the statistical characteristic of the in-service stress spectrum imposed on train components. The uncertainty and variation of the stress spectra of different trips are accounted for. Specifically, the tasks relating to this objective are:

1. Estimating the PDF of stress spectrum for fatigue life analysis of train components operating at in-service conditions; and
2. Constructing an efficient monitoring-based fatigue life assessment approach for in-service high-speed trains.

The second objective of the present study is to develop a probabilistic method to assess and predict fatigue life of train components by integrating continuously collected monitoring data, using Bayesian inference. The proposed probabilistic method can describe the uncertainties in in-service stress spectrum and component material property. This is achieved by employment of Bayesian modeling technique to formulize a novel fatigue life assessment approach. For this objective, the following tasks are addressed:

1. Modeling and updating the PDF of the stress spectrum of monitored stresses using both prior knowledge and updating data; and
2. Using the Bayesian theory as a fundamental principle to construct a probabilistic model for fatigue life prediction of in-service high-speed trains.

It is noted that in the proposed Bayesian approach for fatigue life assessment and prediction, parameters are regarded as random variables to reflect the uncertainties in different types of observations. The posterior distributions of the parameters not only depend on current evidence or observations, but also take into account previous information or so-called prior experience/information (Stone 2013).

1.4 Scope of Work

The scope of the thesis is to develop a practical methodology for monitoring-based fatigue life assessment of high-speed trains. The methodology considers frequentist and Bayesian approaches to model PDF of stress distribution and to address the objectives specified in Section 1.3. The fatigue damage accumulation mechanism is adopted to interpret fatigue behavior.

1.5 Thesis Outline

The research outline is given below:

Chapter 1 is an introduction.

Chapter 2 provides a literature review relating to the research problem.

Chapter 3 presents a frequentist statistical modeling technique for analysis of in-service stresses for fatigue life analysis. The formulated model can utilize in-service stress time histories efficiently for understanding the characteristics of stress ranges and mean stresses. Analysis of the effective stress ranges using frequentist approach is shown in this chapter. It is found that the distribution of in-service effective stress spectrum shows multimodality, which can be modeled by different mixture distribution models (Gaussian/lognormal/Weibull). The modeling process is demonstrated using the in-service monitoring data of a high-speed train.

Chapter 4 introduces a method based on a frequentist approach to conduct fatigue life assessment. Having considered the effects of both stress ranges and mean stresses, an approach for fatigue life assessment is developed using frequentist inference. In addition, an indicator Q is defined to quantify the influence of mean stresses in fatigue life evaluation. As the method conforms to the assessment procedures in railway industry standards, the fatigue life assessment method is flexible and feasible to implement, based on the information of in-service monitoring data. This method is verified through the in-service monitoring data of a train running on a high-speed railway in China.

Chapter 5 presents a modeling and updating method for stress spectrum considering prior knowledge and continuous monitoring data using a Bayesian technique. It is found that for fatigue life assessment, apart from the in-service monitoring data, there always exists prior knowledge about stress range. In real situations, such prior

knowledge also contributes vital information which influences the judgment on the results based on stress range, the probability relating to fatigue life analysis, and the final decision-making. As an alternative to the frequentist approach, when prior knowledge and new monitoring data are available, a Bayesian approach is more preferable and able to model and update the stress spectrum by using new monitoring data and other information (i.e., prior information, new data), to establish the updated stress spectrum for further fatigue analysis. A Bayesian statistical modeling technique is presented for deriving PDF of stress spectrum. This modeling process is demonstrated using the in-service stress ranges of a high-speed train.

Chapter 6 proposes a Bayesian approach for fatigue life assessment developed for in-service high-speed trains using monitoring data. The method is able to assess and update high-speed train fatigue life results by considering the uncertainties of material parameters and in-service train component stresses. The method is verified by assessing the fatigue life of high-speed train components, using fatigue resistance data and in-service monitoring data. Specifically, the monitoring data acquired from a train running on a high-speed railway in China is used.

Chapter 7 draws conclusions and gives recommendations about the future research directions.

CHAPTER 2

LITERATURE REVIEW

Literature review is presented in this chapter. Firstly, fatigue mechanisms, fatigue theories and the history relating to railway fatigue are reviewed. Secondly, an introduction about the research relating to on-board vehicle condition monitoring technology is presented. It is based on structural health monitoring (SHM) technique and provides a favorable and authentic way to understand the characteristics of component deterioration, especially fatigue-related stress variations, during the train normal operation. Subsequently, a research path towards the development of fatigue life assessment approaches is discussed. Finally, different types of approaches established for fatigue life assessment are reviewed.

2.1 Overview and Fundamental Concepts of Railway Fatigue Analysis

This section reviews the basic aspects of fatigue in railway history. Mechanism of metal fatigue occurrence is illustrated. In order to conduct fatigue analysis, previous research has developed several widely acknowledged and applied methodologies for

defining and quantifying fatigue life. These methodologies are explained. Finally, history of railway fatigue research and its significance in fatigue area are reviewed.

2.1.1 Mechanism of fatigue

Understanding fatigue mechanism is a prerequisite to conduct metal fatigue analysis. This is also one of the main focuses in railway fatigue research. The fatigue mechanism has been repeatedly investigated and has now obtained consensus about the general mechanism (Klesnil and Kukas 1980; Stephens et al. 2000; Pook 2007). Generally, cyclic applications of loads to a metal cause continuous changes inside the material, in which microcrack is expected to occur on the surface, along a direction of a so-called "slip band". This microcrack firstly initiates as a fatigue crack nucleus, which has a size at microscopic scale. It should be noted that the microcracks may occur at anywhere on the surface. Once the microcracks are presented and the cyclic applications of loads continue, the small crack nucleus then tends to coalesce, grow, and propagate to a macroscopic size. Finally, the crack becomes unstable and reaches the maximum affordable size to incur fatigue failure until the last cycle of fatigue life. This process of fatigue occurrence to failure has now been well explained by previous research (Schijve 1967; Forsyth 1969; Ritche 1986; Shang et al. 1998). Generally, the crack occurring process before long crack propagation is named as the "fatigue crack initiation" period (Stage I), while the process of long crack propagation is called the

"fatigue crack propagation" period (State II). A schematic illustration of Stage I and II of metal fatigue is illustrated in Figure 2.1.

Engineering practice has adopted different fatigue methodologies to analyze fatigue life of crack initiation and propagation periods. When fatigue of components is in high-cycle fatigue (HCF) regime, crack initiation period consumes most of the fatigue life; while in low-cycle fatigue (LCF) regime, most of the fatigue life is consumed during crack propagation period. For a welded component, the majority of the fatigue life is consumed during the crack initiation period (Schijve 2003).

Crack initiation theory is based on the assumption that the initiation of fatigue cracks is caused by local strains and stresses concentrating on the surface of a component due to geometric shapes (Ye 2010). While crack propagation and final failure states are analyzed using the theory that relates crack growth to the stresses in the component using fracture mechanics.

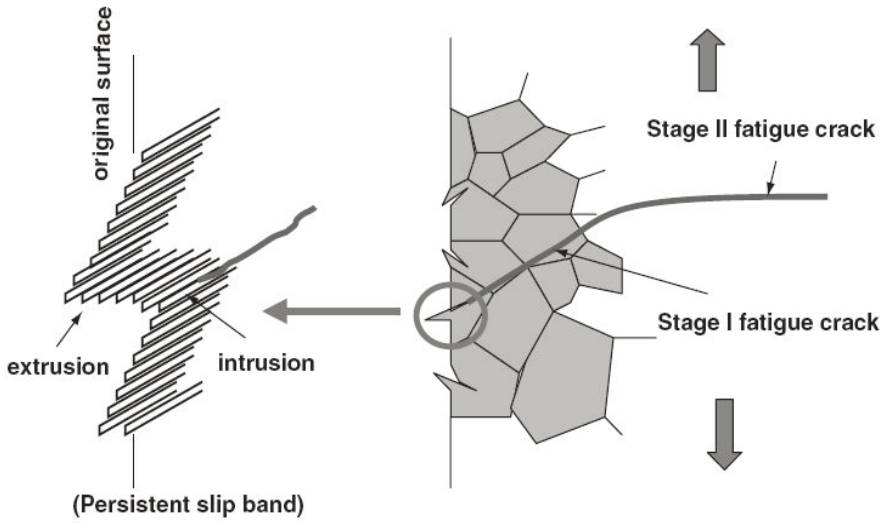


Figure 2.1 Illustration of two-stage mechanism of fatigue failure (Lee et al. 2005)

2.1.2 Overview of mainstream methodologies for fatigue analysis

Historically, two overriding desires about understanding of fatigue from industry have promoted the development of fatigue analysis methodologies: (1) the designing engineers and fatigue experts require methodologies to be practical, easily implemented and cost effective; (2) the methodologies should be reconciled with physical observations, i.e., the fatigue failure mechanisms.

These requirements have advocated the research and development in two major fatigue damage theories for fatigue analysis: cumulative fatigue damage theory (CFD) and fatigue crack propagation (FCP) theory. CFD theory is based on the framework for fatigue strength assessment, while FCP theory is based on fracture mechanics. Specifically, CFD theory believes that fatigue damage in a metal component increases with applied load cycles in a cumulative manner (Fatemi and Yang 1998). This theory is widely used to account for fatigue life during crack nucleation and early growth. FCP theory equates fatigue crack growth rate to a cyclic elastic stress intensity factor, which is relating to the crack condition at the tip of a long crack subjected to a low value of cyclic stress (Paris et al. 1961; Smith 1984). It is a theory to be applied for quantifying fatigue life during long crack growth.

These two theories lay a foundation for the majority of traditional and modern methodologies developed for fatigue analysis. Among them, the most applied and

representative methodologies are stress-life (*S*-*N*-based), strain-life and fracture mechanics methodologies.

The *S*-*N*-based is the earliest, but still the most frequently used methodology for fatigue life analysis (e.g., assessment and prediction). It regards the fatigue life of a component as the number of cycles *N* relating to specific applied stress range *S* experienced by the component; or the fatigue life is regarded as the cumulative number of cycles *N* relating to various amplitudes of stress ranges *S*, according to CFD theory or "linear fatigue accumulation rule" (Miner 1945). To illustrate the relationship between the fatigue life (*N*) versus the specific amplitude of stress range (*S*), a Basquin form is provided for explanation (Yung and Lawrence 1985):

$$S = \sigma_f^{'} (2N)^b \tag{2.1}$$

where *N* is the number of cycles to failure, *2N* is the number of load reversals to failure, σ_f' is the fatigue strength efficiency, while *b* is the fatigue strength exponent. In many industrial applications, the above equation can also be equivalently expressed as a so-called "*S*-*N* relationship" or "*S*-*N* curve":

$$NS^m = C \tag{2.2}$$

where *C* and *m* are empirical constants. *C* is certain for a given material and a specific class of structural details, and the exponent *m* is the material parameter which determines the slope of *S*-*N* curve.

For the stress cycles with variable amplitudes, the fatigue life is counted with fatigue ratio, *D*:

$$D = \sum_{i}^{k} \frac{n_i}{N_i} = \frac{n_1}{N_1} + \frac{n_2}{N_2} \ldots + \frac{n_k}{N_k} \tag{2.3}$$

where n_i is the number of stress cycles at the ith stress range (S_i), and N_i is the number of stress cycles to failure at stress range S_i, and k is the number of classes for stress ranges.

The second methodology is based on strain-life relationship. This methodology assumes that in most practical cases, the crucial locations for fatigue analysis would be notches. At these locations, plastic strains are burdened by surrounding elastic material. Similar to *S*-*N*-based methodology, the fatigue life in strain-life based methodology is expressed with the number of cycles to specific applied strain, and can be described by the Manson-Coffin relationship (Manson and Hirschberg 1964):

$$\frac{\Delta\varepsilon_p}{2} = \varepsilon_f^{'}(2N)^c \tag{2.4}$$

where $\Delta\varepsilon_p$ is the plastic strain, ε_f' is the ductility coefficient, and c is the ductility exponent. This equation can be integrated with the relationship of elastic strain resistance to be expressed as:

$$\frac{\Delta\varepsilon_T}{2} = \frac{\sigma_f^{'}}{E}(2N)^b + \varepsilon_f^{'}(2N)^c \tag{2.5}$$

where $\Delta\varepsilon_T$ is the strain value comprising elastic and plastic strain.

The third methodology is based on fracture mechanics. It is used to predict the fatigue crack propagation length using FCP theory. The methodology is able to describe the

fatigue crack growth rate to the cyclic elastic stress intensity factor, at the tip of a long crack subjected to cyclic stresses with small amplitudes (Paris et al. 1961):

$$\frac{da}{dN} = A(\Delta K)^r \tag{2.6}$$

where A and r are material factors, a is crack length, and da/dN is crack growth rate. ΔK is the range of the stress intensity factor during the fatigue cycles. The methodology is improved and mainly used for assessing fatigue life during crack propagation period (Donahua et al. 1972):

$$\frac{da}{dN} = A_c(\Delta K - \Delta K_{th})^r \tag{2.7}$$

where ΔK_{th} is the threshold value below which no crack growth occurs or fatigue is at crack initiation period.

Other methodologies proposed previously include the energy-based fatigue analysis methodology (Fatemi and Yang 1998; Pan et al. 1999) and the integrating methodology considering both CFD and FCP theories (Miller 1999; Cui 2002). But further and more profound research investigations on these methodologies are yet to be carried out.

The three major methodologies are adopted in industrial cases according to different situations and needs. The *S*-*N*-based is still the most frequently used methodology for fatigue life analysis, especially for designing engineers. The characteristic of fatigue behavior and fatigue life at crack initiation and early propagation period is feasibly described by *S*-*N* relationship and Miner's rule; the strain-life approach is useful when metal component has fatigue issues relating to low-cycle fatigue; while the fracture

mechanics methodology applies physical crack behavior to analyze fatigue, and is able to assess the fatigue life before the crack to rupture. This methodology is precise if the direction of the initial crack is known and the fatigue crack data is obtained.

2.1.3 Research history of railway fatigue and its significance in fatigue research area

In the field of railway engineering, research efforts have been continuously allocated into fatigue problems investigation since 18th century. The first recorded in-service metal fatigue failure was on train vehicle: it was discovered and investigated (Rankine 1842; Smith 1990) that the famous Versailles rail accident was caused by the fatigue failure of a locomotive axle. Same type of problems had repeatedly occurred and been observed on train components (Braithwaite 1854). In order to solve the fatigue problem, it was in the railway field in 1850s that for the first time scientific fatigue analysis was conducted: August Wöhler, a German railroad system technologist, carried out systematic fatigue testing with cyclic stresses (Schutz 1996). Ever after this, increasing amount of research effort has been spent on fatigue analysis throughout the past two centuries (Schijve 2003). However, it is surprising that catastrophic fatigue failures still occur despite long history of fatigue research. Many railway fatigue accidents have been publicized, but they are only the tip of iceberg (Smith 2005).

It was discovered from previous research that more types of components suffer from fatigue problems on train vehicle than on tracks (Smith 2005). Therefore, the research in this thesis is mainly targeted to investigate and consider the fatigue issues relating to train components. The continuous occurrence of fatigue problems on train components can be attributed to the complexity of the normal operating conditions encountered by an in-service train.

Firstly, complicated forces are imposed on train. Passenger loads and self-weight of train coaches and components are constantly supported by train bogie system. Apart from the static loads, dynamic loads including the stresses induced by wheel-rail interactions are the major sources of the forces imposed on trains (Iwnicki 2006). Noticeably, the wind loads (e.g., crossing tunnels) experienced by the train also influence the fatigue life of train components, such as initiating fatigue cracking at spot welds (Smith 2005). In addition, it is also known that the temperatures change throughout the journey of a train. The thermal loads due to temperature variations can change the characteristic of material and affect fatigue life (Tong et al. 2008).

Secondly, the train components are exposed to harsh mechanical conditions when a train runs in an open and ever-changing environment including urban, suburban and rural areas, as well as plains, mountains and bridges. The train is also intermittently accelerating, breaking and turning during the whole trip. All these conditions contribute to varying vibrations and responses of train components generated by the

motions of a train, making the interaction of the different loads on a train become even more complicated.

The complexity of loading conditions and operating environments results in fatigue problems occurring at various locations of train components. As is seen in Table 2.1, Smith (2005) summarized the significant areas of train components at which fatigue failures normally occur. The majority of welded components in train vehicles are affected by fatigue problems during in-service operation.

Table 2.1 Selected significant areas of fatigue at train (Smith 2005)

Zones adjacent to wheel-rail interface	**Train components**
Affected by forces generated at the wheel-rail interface	Bearings, axles, gearboxes, drive shafts, bogies, springs and suspension, components, brake components
Vehicles	Engine or motor components, body shells, couplings, internal components and fittings

It can be found that fatigue analysis of train component remains a significant research topic to be investigated and requires further development of advanced methods and approaches for fatigue analysis and fatigue life assessment.

2.2 In-service Train Condition Monitoring Technology

When the fatigue analysis methodologies were further utilized for fatigue design and life assessment of components of civil infrastructures, there is an increasing demand on more realistic and precise load and fatigue resistance information for more reliable fatigue design and life assessment (Beretta et al. 2011). Realistic information during in-service operation, about the stresses experienced by metal components and the fatigue resistance behaviors, can identify the characteristics of abnormal fatigue behaviors which significantly affect train component fatigue life. Accordingly, the recognition of the benefits of in-service monitoring data as well as the needs from industrial experts and engineers, have advocated the development and advancement in the research field of SHM for civil infrastructure. SHM technology involves condition monitoring, assessment and damage detection using the in-service monitoring data acquired from instrumentation systems (Balageas et al. 2010). With this technology sprouted and thrived in railway field, the investigation on on-board vehicle condition monitoring technology has further helped research advancement of fatigue life assessment in the fatigue discipline. This section reviews the general research and development in SHM technology, specifically on condition monitoring and assessment

for civil infrastructure based on measured data. Subsequently, literature review about the investigations on applying the technology for on-board train condition monitoring is also provided.

2.2.1 Development of condition monitoring technology

Almost all private operators and public authorities in industries desire to identify metal problems and detect component damages as precise and as early as possible (Charles et al. 2006). These pursuits have generally stimulated the generation and development of research in SHM. The SHM process is targeted to determine the current state of the system health under normal operating situation and also under extreme event. The goal is achieved through the observations of a structure or mechanical system for a time period using spaced measurements, the extraction of metal problem-related features, and the statistical analysis of these features (Worden and Barton 2004). Since 1970s, investigations on SHM have been carried out on various types of civil infrastructures for condition monitoring (Bentley and Hatch 2003), non-destructive evaluation (Shull 2002), statistical process control (Montgomery 1997), and damage prognosis on structural systems (Farrar et al. 2003). The initial and continuous research focus of SHM research is the development of effective SHM systems for condition assessment, mainly in the industrial fields of rotating machinery systems, aerospace, oil industry and civil engineering structures (e.g., buildings, bridges). Now the research on effective SHM system has obtained some practical achievements, such as the Shuttle

Modal Inspection System for aerospace vehicles (Hunt 1990); US Navy's Integrated Condition Assessment System (Benedettini et al. 2009); Fiber Optic Sensing Systems in both aerospace (Boller and Buderath 2007) and railway field (Tam et al. 2005); Long-term SHM System for long-span bridges (Ko and Ni 2005; Frangopol 2011) and high-rise buildings (Ni et al. 2012). These systems have generally adopted the statistical pattern recognition paradigms for the purpose of condition assessment. In SHM system, sensing devices are instrumented on structures to obtain various types of structure responses, to extract data features, and to conduct pattern recognition by operating feasible algorithms on digital computing hardware for aerospace, civil and mechanical engineering infrastructures (Chang and Tinoco 1997).

In the field of railway engineering, SHM concept is a relatively new research topic. The concept of condition monitoring technologies in railway emerged at the end of 20th century. It was firstly introduced for solving problems on track (Stone 1992) and later applied to train vehicle (Mastsumoto et al 1999; Mita and Yokoi 2000). The initial motivation of introducing this technology in railway engineering is to achieve predictive and preventive maintenance, reactive maintenance, and further detection and diagnosis of track or vehicle faults.

As the advancement of the technology is highly relied on the development of sensors and measuring techniques, many initial research studies relating to SHM in mechanical and railway engineering were on inventing better sensors and measuring technologies for instrumentation (Turner and Hill 1999). Later the breakthrough in the sensing and

measuring technologies has brought the research advancement on developing SHM-based instrumentation strategy and damage assessment or detection methods for the safety problems of railway tracks (Barke and Chiu 2005) and train components (Chong et al. 2010).

2.2.2 On-board vehicle condition monitoring technology

As mentioned above, the focus of the thesis is on the fatigue problems of train vehicles, because more types of components on train suffer from fatigue problems on vehicles than on tracks. In the subsection, the SHM technology particularly used for train vehicle is reviewed.

The increasing trend of using condition monitoring technology for train vehicle is due to the awareness of our insufficient knowledge on train condition during normal operation. This problem has continuously caused accidents and extra costs due to in-service problems of various train components for years (Federal Railroad Administration Office of Safety Analysis 2008). Employment of effective health and condition monitoring technologies for solving different problems on train vehicles has been an emerging engineering and research practice in many countries and regions, such as France (Maurin et al. 2002), UK (Weston et al. 2006), Germany (Beretta et al. 2011), Mainland China (Ni 2015; Zhao 2016) and Hong Kong (Tam 2008).

For different train components, vehicle condition monitoring technologies have been applied to solve specific mechanical problems experienced by train during normal operation. In previous studies, Tyrell et al. (2000) installed a condition monitoring system comprised by strain gauges, accelerometers and other different types of transducers on coach and bogie frames, to measure the impact forces on the train wall, elastic vibration motions of the train, gross motion of the bogie and the relative loads of the train components. For similar purposes, a fully automated and integrated track and vehicle condition monitoring system was proposed by the research group of University of Birmingham, UK (Weston et al. 2006). Sudden changes of the monitoring data were captured and analyzed using well-developed Kalman filter-based approach for faults detection and diagnosis. In addition to applying sensing device for condition assessment, development was also made on employing nondestructive testing (NDT) experiments for crack detections on gears, wheels and break seats (Zerbst et al. 2005; Rudlin et al. 2006). But there are still technical obstacles remaining unsolved in applying NDT for in-service detection and assessment.

With the development of advanced sensing technology, it has now been possible to employ vehicle condition monitoring system for investigations on the research topics relating to the regions of wheel-rail interaction, such as quantification of wheel surface conditions (Bladon et al. 2004), wheel profile detection (Tam et al. 2005; Lagneback 2007; Ni and Liu 2017), determination of impact forces between wheel and rail

(Stratman et al. 2007), quantification of axle load and vehicle weight gains (Barke and Chiu 2005) and detection of squats on rail (Li et al. 2015).

On-board monitoring, assessment and management of train condition have obtained a wide variety of industrial implementations (Chong et al. 2010; Ni 2015). It is now able to integrate and implement sensors for measuring and analyzing in-service train responses during train trips. From the perspective of sensing technologies, both electronic and fiber Bragg grating (FBG) have been employed to collect monitoring data (Mita and Yokoi 2000; Tam 2008), on different types of trains, such as subway (Tam et al. 2009), freight car (Donelson III et al. 2005) and high-speed train (Ni 2015). In the perspective of monitored objectives, on-board condition monitoring technology measures real-time accelerations, strain, and temperature fluctuations at various locations on train, in order to understand and quantify the characteristics of interactions among in-service train components. For examples, the wheel-rail interactions (Jin and Shen 2001), the patterns of temperature variation with respect to increased speed (Sneed and Smith 1998), the train motions detections in longitudinal, lateral and vertical directions when coupling, accelerating and braking of train (Donelson III et al. 2005), the train motions when crossing different subgrades (Zhang et al. 2016), the structural integrity of carbody (Tam et al. 2009), and the performance deterioration analysis of suspension system (Ho et al. 2006) and overhead line (Bosselmann 2005). Specifically, the research has also enabled the construction of stress or load spectrum

(Beretta et al. 2011) and fatigue life assessment of train components (Tam et al. 2009; Ni et al. 2015).

With the aforementioned research effort on implementing condition monitoring technology for different train safety issues, current investigation in this field has been focused on developing more applicable and feasible algorithms to analyze measurements for further understanding of more complicated train safety issues and prediction of train behaviors.

For the components of train vehicle, weldment is frequently used for connecting train components. Weldment is also one of the crucial locations where fatigue frequently occurs (Neale 2001). For a welded joint, the majority of the fatigue life is consumed during fatigue crack initiation period, which makes it suitable and practical to apply the *S*-*N*-based theory, as introduced in Section 2.1.2, to analyze the fatigue problems and conduct fatigue life assessment (Hobbacher 2008). Literature review relating to the research on *S*-*N*-based fatigue life assessment will be presented later.

Prior to the introduction of the *S*-*N*-based fatigue life assessment methods, some key factors should be considered to develop more applicable and feasible train fatigue life assessment methods; based on fatigue resistance data and in-service monitoring data. Research investigations on these key factors are briefly introduced as follows.

2.3 Key Points Considered in Fatigue Assessment

Some general factors and aspects are now well-recognized to affect the fatigue behaviors and are considered critical for fatigue analysis and assessing fatigue life. The factors of stress range, mean stress, fatigue limit and hot spot stress are normally considered in fatigue analysis and regarded as a guidance to collect data for fatigue life assessment. Their definitions and influences on fatigue are briefly reviewed in this section.

2.3.1 Stress range and mean stress

For typical railway vehicles during normal operation, train components experience diverse load histories, in which random stresses are repeatedly imposed on their surfaces. It is the stress variations (e.g. continuous fluctuations between small and large stresses), rather than absolute stress values, that consume fatigue life, accumulate fatigue damage and cause fatigue failure of train components (Stephens et al. 2000).

To quantify the influence of the stress variations on fatigue, specific notations have been proposed and used in fatigue analysis and life assessment for description of the stress variations. It is shown on a stress history versus time curve in Figure 2.2.

In Figure 2.2, the alternating (σ_a), mean (σ_m), minimum (σ_{min}), maximum (σ_{max}) stresses and stress range (S_a) are introduced. Their relationships can be expressed using the equations as follows:

$$\sigma_a = \frac{S_a}{2} = \frac{\sigma_{\max} - \sigma_{\min}}{2} \tag{2.8}$$

$$\sigma_m = \frac{\sigma_{\max} + \sigma_{\min}}{2} \tag{2.9}$$

$$\sigma_{\max} = \sigma_m + \sigma_a \tag{2.10}$$

$$\sigma_{min} = \sigma_m - \sigma_a \tag{2.11}$$

It is also shown in Figure 2.2 that a stress cycle is defined as the smallest segment of the stress versus time history that is periodically repeated.

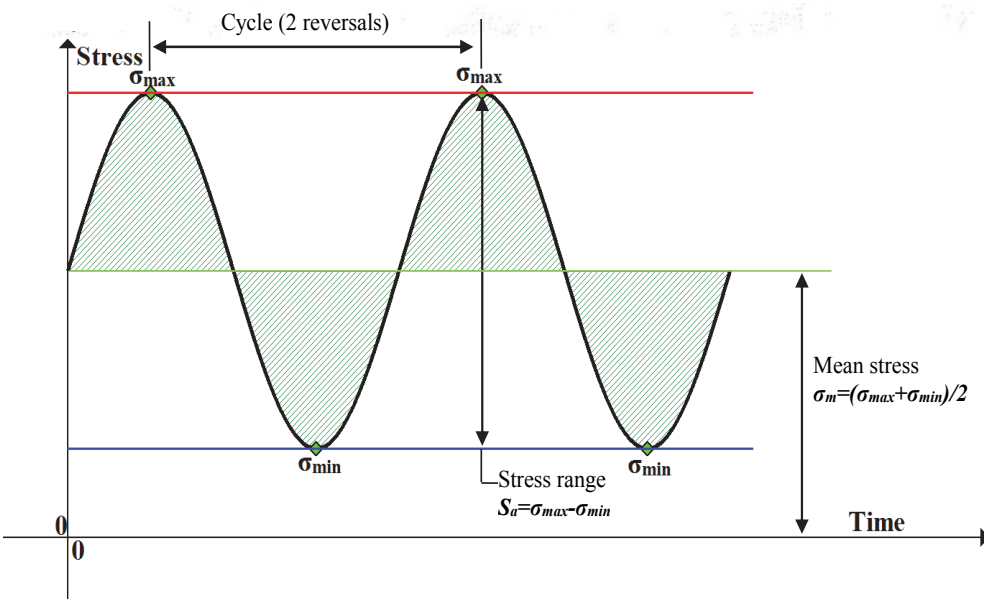


Figure 2.2 Definitions of stress range, mean stress and stress cycles in stress time history

2.3.2 Fatigue limit

A phenomenon is usually observed that fatigue failures do not happen after high number of cycles, when the amplitudes of the imposed stress ranges are very small (Beer et al. 2011). Previously, most metals were thought of having one "fatigue limit" (i.e., a corresponding stress amplitude) at around 10^7 cycles, below which it was assumed that the fatigue crack would not generate or propagate. This assumption is of practical value for the investigation on fatigue problems of in-service train components which are experiencing millions of stress cycles while ensuring no fatigue failure occurring during normal operation.

Further research breakthrough in this topic doubted about the existence of fatigue limit in metal fatigue. Experiments and observations revealed that fatigue ruptures could still occur even at very small stress amplitude but with large number of cycles (e.g., 10^9 cycles) (Bathias 1999), and the fatigue crack growth rates were similar for the different stress ranges with very low stress amplitudes (Taylor and Wang 1999). This concern on the characteristics of fatigue limit had advocated more practical treatments of fatigue limit for engineering practice, such as different solutions of fatigue life calculations at material, component and system levels (Smith 1984, Miller 1999), and using special reduction factors to count the fatigue life corresponding to the stress ranges lower than fatigue limit (Hobbacher 2008). Now it is a common practice in different engineering codes to adopt some suggestions to optimize the fatigue treatments of the stress ranges above and below fatigue limit.

2.3.3 Effect of stress concentration and hot spot stress

Train vehicles comprise numerous individual metal components which are integrated and connected together with welded joints. Geometrically, they can be regarded as plates with attachments or notches. The existence of the attachments or notches changes the stiffness in the welded area, resulting in stress concentration and nonlinear through-thickness stress distribution. This stress concentration is predominantly responsible for the fatigue damage accumulation at welded area. Past statistics about fatigue failures found that the welded joints are the locations where the potential fatigue cracks will probably initiate (Haagensen 1997). However, it is still a difficult and challenging task to determine a meaningful indication for fatigue analysis in complex welded structures.

Therefore, a structural stress termed as the "hot spot stress", was proposed for describing the stress amplitude causing fatigue at welded joints. The hot spot stress accounts for the effect of the global geometry of structure and the existence of weldment, but does not consider the micro-geometrical effects. The "hot spot stress" has been widely used in engineering industries (e.g., bridge, offshore structure, aerospace, and railway) for quantifying the effect of stress concentration and analyzing fatigue near the locations where stress may concentrate and distribute nonlinearly. Extensive research investigations have been conducted on stress concentration effect and hot spot stress. Some consensus were obtained about this issue, and major principles have been formalized to calculate hot spot stress in engineering practices:

International Standards such as BS (BS 7608; BS 7910), IIW (Niemi et al. 2006), ASME (2007), EN 13445 (2006) and CCS (2013) now provide similar solutions and hot spot-based fatigue *S-N* curves for defining, quantifying and measuring hot spot stress through theoretical calculation, finite element method (FEM) and experimental measurements. Theoretical values of hot spot stress are similar to the results from FEM analysis and in-situ monitoring data from sensors (Petershagen et al. 1991).

2.4 Features of *S-N*-based Fatigue Measurement Data

This section introduces the characteristics of different types of data to be employed for *S-N*-based fatigue life assessment. The fatigue resistance data, is obtained through experiments in laboratory environment; the in-service condition monitoring data, is acquired during normal operation of train components. Previous studies have also confirmed the inherent statistically scattering nature of the data, and investigated the possible causes of the uncertainties.

2.4.1 Fatigue resistance data obtained in laboratory

In modern theory of fatigue analysis, the primary focus is on the fatigue characteristics of the material, which is used for manufacturing different types of structures and metal components. The fatigue property of a material is described as the number of cycles N causing fatigue failures of a material under a specific amplitude of stress range S_a. The

information and characteristic about the material fatigue life is provided through a curve with the fatigue life N plotted in the horizontal direction against the corresponding stress range amplitude S_a plotted in the vertical direction. Such relationship is labeled as "Wohler curve" or so-called "*S-N* curve".

Ever since 20th century, numerous fatigue tests have been carried out to produce large amount of fatigue resistance data for understanding the material behavior under cycling stresses and also establishing *S-N* curves (Schijve 2003). In the beginning, rotating beam tests on unnotched specimens had been popular for a long period; it was not until 1940 that the fatigue research field introduced notched specimens to obtain fatigue resistance data in order to consider the effect of stress concentration. The next breakthrough was attributed to the advancement of fatigue testing equipment, which enabled the fatigue loads imposed on structures and components to be more controllable. This advancement boosted the research discoveries and investigations in terms of the fatigue behaviors being affected by different combinations of stresses ranges above and below fatigue limit. Meanwhile, in consideration of the importance of the research on welded joints fatigue, researchers also conducted a wide variety of experiments on welded joints according to different types of weld geometry characters (Haagensen 1997; Maddox 2001; Statnikov 2002; Habbacher 2008).

A common and normal phenomenon about the fatigue resistance data is the statistical scatter of fatigue resistance data: for the same type of specimens imposed with the same stress amplitude *S*, fatigue failures can be observed at different numbers of cycles

N, although these numbers are mostly close and similar. This is a natural characteristic of fatigue resistance data (Schijve 2001; Radaj et al. 2006), as the fatigue resistance data is in essence using statistics experiments to describe the fatigue property (Zhao et al. 2017). There are variations in the ways of different manufacturers producing same type of specimens, and there are also variations in the ways of different laboratories gripping specimens on test rigs before conducting experiments. The uncertainties in the observed numbers of cycles for the same type of specimens bring forward the necessity in development of statistical and probabilistic models to better describe the fatigue resistance data (Al-Rubaie 2008). Those models will be elaborated later.

2.4.2 In-service monitoring stress data obtained on-site

In addition to fatigue resistance data, in-service stress spectrum is also an indispensable input for fatigue life assessment and prediction. The in-service stress spectrum of a train component can be extracted from the stress history experienced by the component during operation, which is acquired through strain or acceleration sensors embedded in the aforementioned SHM system, such as on-board vehicle condition monitoring systems (Laruent et al. 2002). The techniques and systems for monitoring in-service stress-time histories have been developed to a decent and high satisfaction level for different types of structures (Schijve 2003).

In railway industry, the initial motivation in measuring in-service stress time history of train components is that the stresses in real situations are quite different from the

traditional assumptions used in train design. Although the techniques developed for monitoring in-service stresses and their application for fatigue life assessment had been introduced in other industries like aircraft at very early time (Seewald 1931; Gassner 1939), railway industry still followed classical design methods for fatigue analysis and life assessment, rarely considering in-service stress histories. It was not until the occurrence of many fatigue failure events during train operation (Smith and Zhou 2014), the gradual awareness of unrealities of some traditional design assumptions (e.g., fatigue endurance limit method) (Grubisic and Fischer 2006) and the more complicated operating environment encountered by newly invented high-speed trains (Lee and Lee 2015) that railway industry gradually introduced in-service monitoring data for fatigue life assessment and design optimization. Stress time history obtained from on-board vehicle condition monitoring system is used to conduct fatigue life assessment or realistic loads simulations of train components, to verify the safety of current status or the validity of new design.

An important development for fatigue analysis is the techniques proposed for extraction of stress ranges and mean stresses for varied stresses. It is realized through using the so-called rainflow counting method (Endo Murakami 1968; Downing 1982). The method counts the maximum and minimum stresses in its defined stress cycles into different stress ranges, in order to reduce a spectrum of varying stresses into a set of simple stress reversals (Frýba 1996). Many other methods were also proposed previously for different situations to achieve the same goal (ASTME1049 1998; Wei

and Dong 2010; You et al. 2011). Subsequent to the extraction of stress ranges and mean stresses, the fatigue life assessment can be calculated based on Miner's rule according to international standards. Although this is enough for engineering practice, more precise results for fatigue analysis should employ electro-hydraulic closed-loop systems to simulate service load-time history (Branger 1964; Grubisic and Fischer 2008). However, the latter one (i.e., regeneration of load spectra using in-service data) is mainly used for aircraft fatigue analysis, difficult to realize, and currently not a widely applied principle in train design against fatigue (Mancini and Cera 2008).

From the measurements information of sensors deployed on trains, it has been well understood in railway industry that the service stresses experienced by different train components have a time-varying feature (Hirakawa et al. 1998; Lagoda and Sonsino 2004; Qin et al. 2004; Waston and Timmis 2011; Lee and Lee 2015). This is a result of the varying force levels imposed on rolling wheelset, which are mainly caused by wheel-rail interactions during train running on railway (Grubisic and Fischer 2012). It is also required to employ appropriate methods to consider and quantify the effect of the inherent uncertainties in stress time histories, for fatigue life assessment and prediction.

2.5 *S-N*-based Fatigue Life Assessment Approaches

Fatigue life assessment approaches have been continuously optimized throughout the history of fatigue research. In the early years, the fatigue assessment approaches were formulated in simplified form in the aim of helping pioneers understand and address the fatigue problems more easily when they encountered fatigue problems. Later, the advancement of techniques in engineering research boosted the development of ways to conduct fatigue analysis and life assessment that more influential factors to be considered in the original approaches, enabling fatigue life to be assessed considering more complex conditions. Subsequently, development of modern theories in fatigue analysis gradually changed the view on fatigue life consumption from static to dynamic perspectives. It was found that the uncertainties in both materials and imposed stresses affect the fatigue life and the component performances. Therefore the fatigue life assessment approaches began to integrate the modeling techniques using PDF for fatigue analysis and life assessment.

2.5.1 Deterministic approaches

In railway industry, the traditional fatigue life assessment approach in the early years is based on *S-N* relationship. The approach judges whether the stresses experienced by train components exceed the defined "safety limit", and evaluates the fatigue life by applying the theoretically calculated loads according to specifications (Leluan 1994; Cera et al. 2008). Although this approach has been commonly applied for train

components since the middle of the 20th century, the fatigue life results calculated through this way were quite different from reality and hence could not reflect the real situations. It was discovered that the stresses from the theoretical calculation on train components were overestimated, which led to underestimation of fatigue life (Beretta et al. 2011). Meanwhile, fatigue failures frequently occurred prior to the termination of the design criteria for fatigue life, and phenomena of low fatigue life were usually observed at train components operating at different conditions (Lonsdale and Stone 2004; Smith 2005).

Acknowledged specifications in railway (Hobacher 2008; BS7608 2014) and railway experts (Grubisic and Fishcher 2012) began to advocate assessing fatigue life based on in-service measurement data obtained during train operating under different running conditions, for examples, straight, left/right turn, passage through station, bridge, tunnel, etc. To utilize the information of the measurement data, the data is firstly converted to in-service stress or load spectra using rainflow counting method, then to fatigue damage ratio through Miner's rule, and finally the maximum running distance or number of years in operation is treated as the fatigue life when the fatigue damage ratio reaches a threshold (Stichel and Knothe 2007). In addition, *S*-*N* curve is the basis used in the fatigue life assessment approaches, and its basic form has been extended to consider more factors influencing fatigue life assessment results (Al-Rubaie 2008), such as the influences of multi-axle stresses (Sonsino 2007), mean stress effect (Park et al. 2006), effect of stress concentration (Leander et al. 2010),

environmental effects (e.g., corrosion, temperature) (Chookah 2011; Nuhi 2011) and sensing performance (e.g., sampling frequency) (Luo et al. 2007). These research studies figured out the significance of different factors on fatigue damage accumulation, which findings can be introduced to optimize the approaches for fatigue life assessment of train components under different situations.

2.5.2 Probabilistic approaches

The investigations on statistical aspects of fatigue behaviors and measurement data introduced the concept of probability into fatigue analysis in the mid of 20th century (Stulen 1952; Miles 1954; Freudenthal and Gumbel 1956). This concept has been used in metal fatigue fields for constructing appropriate probability models of specific data for the analysis of fatigue resistance data (Wirching and Chen 1988), stress spectrum (Ye et al. 2012; Zhu and Collette 2016), fatigue failure data (An et al. 2011), fatigue limit (Sudret et al. 2003), and reliability assessment (Kwon and Frangopol 2010; Beretta and Regazzi 2016).

There are limited but pioneering work on developing probabilistic fatigue life assessment approaches by considering uncertainties in fatigue behaviors. Tunna (1986) introduced a probabilistic tool for fatigue life prediction method. The research showed predictions can be made directly from knowledge of the power spectral density of the strain history and conventional fatigue strength data. Mohammadi et al. (1986) and Dietz et al. (1998) utilized this probabilistic concept to conduct fatigue life assessment

for both freight and passage trains. However, their studies are simulation-based and do not mention the details about implementation process. Later many research studies in fatigue fields were on solving modeling issues to obtain probability density function for power spectrum density (PSD) of engineering measurement data (Petrucci and Zuccarello 1999, 2004; Ni et al. 2012). Grubisic and Fischer (2012) then evaluated the fatigue life by considering the probability of survival of the fatigue resistance data collected from the train axle. Beretta and Regazzi (2016) proposed a probability fatigue analysis approach for evaluating the fatigue rate of train axle during in-service operation. Probabilistic approaches are now possible solutions for fatigue life assessment of in-service train.

2.6 Remarks on Literature Review

In summary, fatigue is one of the major concerns for the railway safety. Investigations on fatigue stemmed from railway field and have been continuously conducted for more than 180 years. According to the fatigue crack initiation and propagation mechanisms, there are three types of methodologies developed for fatigue analysis, which are *S*-*N*-based, strain-life based and fracture mechanics methodology. They are used for fatigue analysis under different situations. The emerging SHM technology helps the advancement of fatigue analysis by providing authentic information of stress responses of train components during normal operation. Recent research effort has been devoted to the development of on-board vehicle condition monitoring techniques for in-service

trains. With the established fatigue theories, it is possible to establish a fatigue life assessment approach utilizing in-service monitoring stress time history for instrumented high-speed trains, to further achieve the vision of real-time monitoring-based fatigue life assessment. In order to develop a probabilistic fatigue life assessment approach, several factors should be considered regarding the fatigue problem of train components. One of the key components frequently affected by fatigue problems is the welded joint. The majority of its fatigue life is consumed during fatigue initiation period, and therefore it is appropriate to adopt *S*-*N*-based methodology to establish fatigue life assessment approach. The measurement data for fatigue life assessment includes both fatigue resistance data and in-service stress data. Both types of data contain statistical scatter, contributing to the majority of uncertainties observed in fatigue behavior. Uncertainties in fatigue resistance data have been well studied but limited studies have been made to investigate uncertainties in in-service stress spectrum. Previously developed fatigue life assessment approaches for train vehicles are mostly deterministic and do not systematically consider the uncertainties. Recent research starts to introduce the concept of probability to explore fatigue problems but still needs further investigations to develop more reliable and feasible fatigue life assessment approaches.

For a train instrumented with an on-board monitoring system, at early period, the monitoring data is the only source of data to assess fatigue life; as the monitoring goes on, continuous on-board monitoring data and environmental data are collected. For

practical consideration, the on-board monitoring techniques and fatigue life assessment methods for high-speed trains must be practically implementable in real situations. For fatigue life assessment approaches, at the initial stage the results are based on monitoring data, while in a longer period of time the predictive results can be provided by utilizing more information, such as prior knowledge and continuous monitoring data.

Based on the above, this thesis will focus on the first development of probabilistic fatigue life assessment approaches for railway systems, utilizing both frequentist and Bayesian techniques. The frequentist-based fatigue life assessment approach, which has been applied in civil structures such as bridges (Ni et al. 2011), will be integrated and optimized to enable fatigue life assessment for in-service high-speed trains. Another target for probabilistic fatigue life assessment is to assess the result with the information of both prior and continuous monitoring data, and predict the fatigue life based on history and current status. This goal can be achieved through utilizing Bayesian technique to formulate a procedure that carries out fatigue life assessment and prediction utilizing both prior knowledge and monitoring data.

CHAPTER 3

FREQUENTIST STRESS MODELING TECHNIQUE FOR FATIGUE OF TRAIN COMPONENTS UNDER IN-SERVICE ENVIRONMENT

3.1 Introduction

The *S*-*N*-based approach is the most frequently used tool to assess the fatigue life and fatigue failure occurrence of railway components (Esderts et al. 2012; Kassner 2012). The approach derives the stress spectrum for a train bogie over the stress ranges experienced under repeated loadings. Such an *S*-*N*-based fatigue assessment procedure is incorporated in many specifications in use by railway industry (JIS 2004; Hobbacher 2009; BS7608 2014). These specifications advise the adoption of the *S*-*N*-based approach for design against fatigue and fatigue life assessment of train components, using experimental data from laboratory-based loading test results. However, there are a variety of obstacles in obtaining representative stress spectra for train bogies using laboratory tests (Griffiths et al. 1971; Miller and O'Donnell 1999; Baek et al. 2008; Zhang 2008; Susmel and Tovo 2011). In addition, it is not reliable, in general, for laboratory tests to reflect the uncertain nature of the real stress situations for running

trains (Roux et al. 2008). For these reasons, it is highly desired to explore the stress spectra that derive from true life stress situations.

Real-time monitoring data acquired from an instrumented high-speed train provides uncorrupted and reliable data on the authentic health condition of train components (Barke and Chiu 2005; Seki 2012). Field monitoring data enables the investigation and estimation of the dynamic characteristics of train components during actual operating conditions (Madshus et al. 1996). Stress ranges extracted from the monitoring data are used to construct stress spectra for fatigue life assessment.

As mentioned in Section 2.4, the phenomenon of the statistical scatter of in-service stress ranges advocates the use of a probability distribution model to interpret the acquired stress spectrum. Meanwhile, over the initial periods of on-board monitoring of train, only relatively short measured stress time histories are available and it is necessary to model the stress range distribution in order to acquire stresses in the regions beyond the measured data. Hamid et al. (1983) found that the applied stresses on a vehicle could be interpreted by a common PDF. Mohammdadi et al. (1986), Tunna (1986), and Dietz et al. (1998) later extended the research by fitting Gaussian, Beta, and Rayleigh distributions to the normalized data, and showed that these distributions could be used for modeling applied stresses in train vehicles. Park et al. (2005) interpreted the stress ranges frequency for 400 block loadings using a lognormal probability distribution. Although various types of single PDF have been widely employed in stress spectrum modeling (Maljaars et al. 2012; Kang et al. 2012), their

performance is not satisfactory (Mclachlan and Peel 2000). This is because the stress spectrum exhibits multimodality characteristic and therefore it is more appropriate to adopt a mixture PDF to describe the stress distributions (Ni et al. 2011; Wang et al. 2015).

In addition, the influence of mean stresses on fatigue life must also be considered in this assessment. For in-service stress situations, therefore, the mean stress values in the stress history affect metal fatigue life assessments (Griffiths et al. 1971; Fuchs and Stephens 1980; Henrysson 2002). The mean stresses in stress history are attributed to several causes including fluctuating temperatures and mechanical loadings (Kihl and Sarkani 1999). The influence of mean stresses on fatigue life is described by constant life equations. These equations are formulated from past experimental results by applying artificial loadings on metals (Sendeckyj 2001). Thus, the train component stress spectrum models should be able to provide more reliable fatigue life estimation for train components if the mean stress effect is considered.

In this chapter, the stress ranges and mean stresses are analytically modeled by means of finite mixture distributions, which are relating to monitoring data collected from an operating high-speed train in China. The stress range and mean stress for each set of stress cycles are obtained from the measured strain time history using a rainflow counting method. The rainflow counting method recognizes the stress cycles from the original stress time history, obtains the maximum and minimum stresses from each stress cycle, and then calculates the stress range and mean stress for each of all stress

cycles, which will further be used for fatigue life analysis. Integration of the stress ranges and mean stresses allows an effective stress spectrum to be derived using the monitoring data and the constant life equations. Subsequently, the effective stress spectrum PDF is formulated and applied by means of a frequentist approach. The proposed modeling method is illustrated using one-month strain monitoring data to establish the stress spectrum PDFs.

3.2 Stress Spectrum Modeling Using Frequentist Inference

For fatigue life assessment purposes, it is valuable to know the characteristics of the tail behavior of stress spectrum. This can be realized by PDF modeling. When only relatively short stress time histories are available, it is necessary to model the distribution of stress ranges with a continuous PDF so as to adequately represent the stochastic stress states and extrapolate to the regions out of measuring areas. To realize a high-fidelity PDF representation for stress spectra, a finite mixture distributions technique, in conjunction with an estimation-maximization (EM) algorithm (Train 2008; Ni et al. 2011), is adopted in this study.

3.2.1 Theoretical basis

A fitted PDF is supposed to represent the observed probability distribution of stress spectrum adequately and accurately. Finite mixture distributions are capable of

modeling complex probability distributions using either single or multi-modal probability density functions, when a single distribution model fails to adequately describe the characteristics of the original observations. Generally, the basic structure of an arbitrary finite mixture distribution to depict the vector of observations, *x,* can be defined as the weighted sum of the component distributions:

$$f(x \mid \Theta) = \sum_{j=1}^{k_l} w_j f_j(x \mid \theta_j) \tag{3.1}$$

where the $f(x \mid \Theta)$ is the predictive mixture density with the parameters $\Theta = (w_1, ..., w_m;\ \theta_1, ..., \theta_m)$. The term $f_j(x \mid \theta_j)$ represents the predictive component density with a given parametric family indexed by the scalar or vector parameter θ_j., and w_j represents the weight for each component density and $\sum_{j=1}^{k_l} w_j = 1$. Thus, once the number of components k_l, the weight factors w_j, and the corresponding component parameters θ_j are determined, the mixture distributions model is formed for depicting the observations.

3.2.2 Preprocessing of monitoring data

For each stress cycle, the stress range S_a and the mean stress σ_{m} are obtained by calculation using the maximum stress $\sigma_{\max}$ and minimum stress $\sigma_{\min}$ measured in this stress cycle as:

$$S_a = \sigma_{\max} - \sigma_{\min} \tag{3.2}$$

$$\sigma_m = \frac{\sigma_{\max} + \sigma_{\min}}{2} \tag{3.3}$$

Stress ranges derived from stress cycles are used to form a stress spectrum. In order to quantify the contribution of mean stresses towards fatigue failure, an effective stress range must take into account the effects of both the actual stress range and mean stress through a series of constant life equations summarized in (Susmel et al. 2005):

$$\left(\frac{S_a}{S_e}\right)^r + \left(p\frac{\sigma_m}{S_u}\right)^t = 1 \tag{3.4}$$

where *p, t, r* are constants, and S_u is the ultimate tensile strength of the material. The value of the effective stress range is expressed in terms of stress based on the following expression:

$$S_e = \frac{S_a}{\left[1-(p\frac{\sigma_m}{\sigma_{UTS}})^t\right]^{1/r}} \tag{3.5}$$

Thus, the effective stress range S_e for each stress cycle is converted from the stress range S_a, in which all values are categorized into *k* classes as indicated in Equation (3.5). The effective stress ranges in the time history are then applied to construct an effective stress spectrum.

3.2.2.1 Stress ranges extracted by rainflow cycle counting technique

After collecting the strain data, the stress time history can be obtained by multiplying the measured strain values by the elasticity modulus of the material with the assumption of elastic strain.

To obtain stress ranges, the stress time history is converted from a variable-amplitude stress sequence into a series of equivalent constant-amplitude stress cycles. In this study, rainflow cycle counting method is adopted to extract stress ranges together with the corresponding number of cycles. A brief introduction of rainflow counting technique is as follows:

It is noticed that the fatigue damage is particularly affected by the alternate plastic deformations, causing irreversible changes (the so-called “hysteresis”) inside the material and finally leading to fatigue failure (Frýba 1996). Therefore, the evaluation of fatigue damage must be based on the information of the extracted stress-strain diagram from measurement stress responses rather than the original stress-time history, as shown in Figure 3.1.

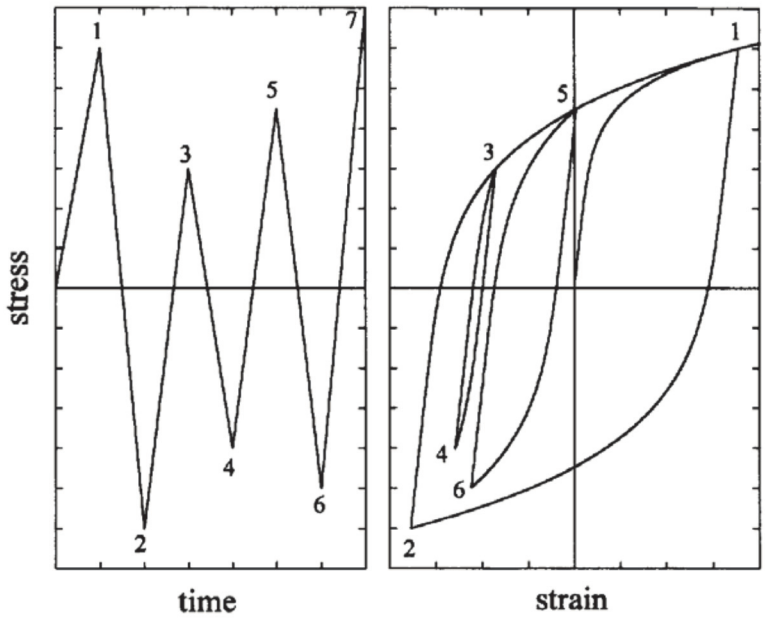


Figure 3.1 Stress time history and corresponding stress-strain diagram (Downing and Socie 1982)

There were several techniques developed for conversion of the stress-strain diagram from the stress-time history. Among them, the rainflow counting method is the most widely used technique, which is regarded as faithfully reflecting the behavior of the material and characterizing the aforementioned hysteresis inside the material. The fundamental assumption of the rainflow counting method is that the fatigue damage due to small induced stress cycles may be added to the fatigue damage due to large stress cycles. Based on this, the implementation of rainflow counting technique counts every time interval of stress only once and can obtain the same result by counting stress history from the opposite direction.

Specifically, As shown in Figure 3.2, the basic procedures of rainflow counting technique are as follows (Ye 2010): (i) reduce the stress time history to a sequence of tensile peaks and compressive troughs; (ii) imagine that the stress time history with lines joining the peaks and valleys is a template for a rigid sheet (pagoda roof); (iii) turn the sheet clockwise 90° (earliest time to the top) and the stress time history is plotted so that the time axis is vertically downwards; (iv) each tensile peak is imagined as a source of virtual rain that "drips" down the pagoda and a complete cycle or a half cycle of stress is defined by the path of the rain flow; (v) count the number of half-cycles by looking for terminations in the flow occurring when either: (1) it reaches the end of the stress time history, (2) it merges with a flow that started at an earlier tensile peak, or (3) it flows opposite a tensile peak of greater magnitude; (vi) repeat step (v) for compressive troughs; (vii) assign a magnitude to each half-cycle equal to the stress

difference between its start and termination; and (viii) pair up half-cycles of identical magnitude (but opposite sense) to count the number of complete cycles.

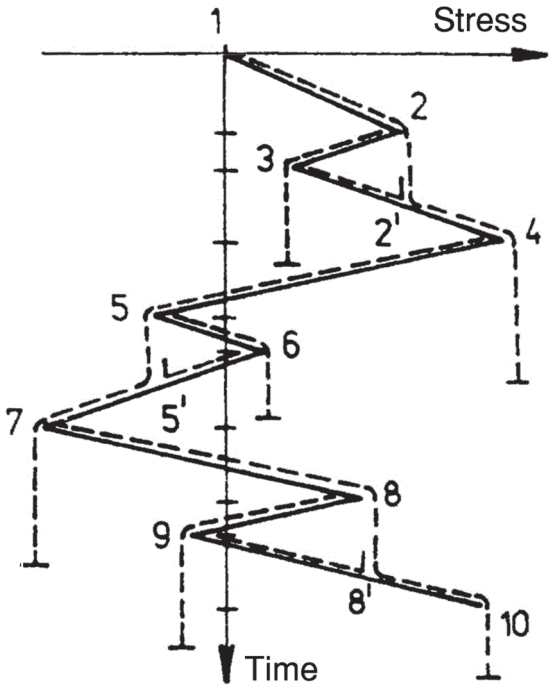


Figure 3.2 Process of stress-time history using the rainflow counting technique (Frýba 1996)

3.2.2.2 Mean stress effect extracted by empirical relation

An effective stress spectrum is derived to account for the combined effects of the two parameters: mean stress and stress range. A key limitation in the use of *S-N* curve for deriving stress ranges using rainflow counting method is the inability to account for the mean stress effect (Fuchs and Stephens 1980; Vassilopoulos et al. 2010). Previous studies have addressed this issue and proposed different constant life equations capable of considering the mean stress effect on materials fatigue (Susmel et al. 2005; Gedeon

2014), using the effective stress concept via a relationship combining stress range and mean stress. The effective stress range can be considered as a set of full stress reversals that would cause the same fatigue life as the combination of stress range and mean stress. These relationships satisfy Equation (3.4) by defining different values of *p, t,* and *r*. In particular, the effective stress range is obtained using a constant life equation.

Among the constant life equations, the Goodman relation is the most fundamental one used to account for the mean stress effect. In the design stage of a railway vehicle, it is also common to consider the effect of mean stress using the Goodman relationship diagram (JIS 2004; Kim 2006; Baek et al. 2008). In materials science, the Goodman relation is an equation used to quantify the interaction between stress range and mean stress, and also their effects on the fatigue life of a material. The Goodman relation postulates a straight line, on a two-dimensional coordinate system, connecting the alternating stress amplitude with mean stress to a fully reversed fatigue stress, S_e on the y-axis, corresponding to a specified number of cycles to failure, and with the ultimate stress, S_u, on the x-axis.

Because the Goodman relation constitutes the foundation for other advanced constant life equations (Sendeckyj 2001), it is used below in relation to the illustration and discussion of a constant life equation for obtaining an effective stress range. Based on Equation (3.4), the Goodman relation is expressed as:

$$S_e = \frac{S_a}{1 - \frac{\sigma_m}{S_u}} \tag{3.6}$$

where S_a is the alternating stress, σ_m is the mean stress, S_e is the fully reversed fatigue strength of the specimen, S_u is the ultimate tensile strength. S_e is calculated from the algorithm represents the combining effect by stress range and mean stress.

Firstly, with rainflow cycle counting method mentioned in Section 2.3.1, the stress ranges and mean stresses are generated for a structural component under stochastic loading conditions. Secondly, by making use of the constant life equations to consider both stress range and mean stress effects, the effective stress ranges and associated effective stress spectrum can be formulated for future fatigue life assessment.

3.2.3 Estimation of stress spectrum distribution

In general, the Gaussian, lognormal and Weibull mixture distributions (Xia 2012) are appropriate selections for finite mixture distributions modeling:

$$f(x \mid \Theta) = \sum_{j=1}^{k_l} w_j \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left\{ -\frac{1}{2} \frac{(x-\mu_j)^2}{\sigma_j^2} \right\} \tag{3.7}$$

$$f(x \mid \Theta) = \sum_{j=1}^{k_l} w_j \frac{1}{\sqrt{2\pi}\sigma_j x} \exp\left\{ -\frac{1}{2} \frac{(\ln(x)-\mu_j)^2}{\sigma_j^2} \right\} \tag{3.8}$$

$$f(x \mid \Theta) = \sum_{j=1}^{k_l} w_j \frac{\beta_j}{\gamma_j} \left(\frac{x}{\gamma_j} \right)^{\beta_j - 1} \exp\left\{ -\left(\frac{x}{\gamma_j} \right)^{\beta_j} \right\} \tag{3.9}$$

where μ_j and σ_j are the mean and standard deviation for either the Gaussian or lognormal mixture distributions, and γ_j and β_j are the scale and shape parameters for the Weibull mixture distribution.

Identification of the parameters in finite mixture distributions can be regarded as a latent variable problem that undergoes usual analysis by means of an expectation-maximization (EM) algorithm (McLachlan 2007). The EM algorithm finds the maximum-likelihood estimate of the distribution parameters underlying the given data set in an iterative manner (Train 2008). The EM algorithm comprises two repeated operating steps to estimate the optimum parameters for the model. The first step formulates the parameter expressions with the measured data for estimation purposes (E-step). The second step finds the maximum expected values for each of the newly optimized parameters (M-step, Bucar et al. 2004). In the E-step, the likelihood shows the relationship of current and next-step estimated results can be simply shown as (Bishop 1995):

$$\begin{aligned} L(\Theta^{t+1} \mid \Theta^t) &= \sum_{i=1}^{n}\sum_{j=1}^{m} f(j \mid x_i,\Theta^t)\log(w_j f_j(x \mid \theta_j)) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} f(j \mid x_i,\Theta^t)\log(w_j) + \sum_{i=1}^{n}\sum_{j=1}^{m} f(j \mid x_i,\Theta^t)\log(f_j(x \mid \theta_j)) \end{aligned} \tag{3.10}$$

where the probability of $f(j \mid x_i, \Theta^t)$ is expressed as:

$$f(j \mid x_i,\Theta^t) = \frac{w_j^t f(j \mid x_i,\theta^t)}{\sum_{j=1}^{m} w_j^t f(j \mid x_i,\theta^t)} \tag{3.11}$$

Once the relationship between current and next-step parameter states form the E-step, the next action is to estimate the next-step results for specific parameters by operating the likelihood maximizing process from the M-step. Since the parameters w_j and θ_j are not relating to each other, their next-step states can be updated independently, by solving the following equations separately (Xia 2012):

$$\frac{\partial}{\partial w_j}\left[\sum_{i=1}^{n}\sum_{j=1}^{m} f(j \mid x_i,\Theta^t)\log(w_j) + \lambda(\sum_{j=1}^{m} w_j)\right] = 0 \tag{3.12}$$

$$\frac{\partial}{\partial \theta_j}\left[\sum_{i=1}^{n}\sum_{j=1}^{m} f(j \mid x_i,\Theta^t)\log(f_j(x \mid \theta_j))\right] = 0 \tag{3.13}$$

where λ is Lagrange multiplier introduced with a constraint that $\sum_{j=1}^{k_l} w_j = 1$ in Equation (3.12). Analytical expressions of the optimized parameters w_j are derived as:

$$w_j = \frac{1}{n}\sum_{i=1}^{k_l} f(j \mid x_i, \Theta^t) \tag{3.14}$$

The optimized parameters θ_j can be obtained using the Newton-Raphson method (Nagode et al. 2006). In addition, various attempts have been made to speed up the calculation process, making it a very fast and effective algorithm for finding fitted models (Meng and Van Dyk 1997).

The EM algorithm can be used to obtain the best description of the observations using Gaussian, lognormal, and Weibull mixture distributions. Among the different descriptions, in order to select the best fitted model with an optimal number of components, the qualities of the Gaussian, lognormal, and Weibull mixture

distributions with different number components are compared and selected using the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) that measure the relative quality of statistical models for a given set of data.

3.2.3.1 AIC and BIC rules

In order to select the best mixture model for the distribution, a number of model selection criteria have been proposed (Bozdogan 1993). The criteria are applied separately to each of the class-conditional mixture models to find the model with the best performance. That is to consider the fit of each class-conditional distribution, rather than the overall discrimination performance.

A comparative study of different model selection criteria was made by Bozdogan (1993). One of the most popular rules is the Akaike Information Criterion (AIC) (Akaike 1974) given by:

$$\mathrm{AIC} = -2\log(l_1) + 2m_1 \tag{3.15}$$

where l_1 represents the value of maximum likelihood, m_1 is the number of estimated unknown parameters. The best model with optimal component number for the modeling of stress and effective spectra is the one with the lowest AIC value.

When making selection among competing models, the model with the smallest AIC is optioned. The AIC penalizes over-fitting using high-order models (which result in larger log-likelihood values). In addition to the popular AIC rule, another common

criterion is the Bayesian Information Criterion (BIC) proposed by Schwarz (1978). It is given by:

$$\mathrm{BIC} = -2\log(l_1) + m_1 \log(n) \tag{3.16}$$

where l_1 represents the value of maximum likelihood, m_1 is the number of estimated unknown parameters, and n is the number of observations for estimation. Similar to the AIC rule, when using BIC rule for model selection, the best model with optimal component number for the modeling of stress and effective spectra is the one with the lowest BIC value.

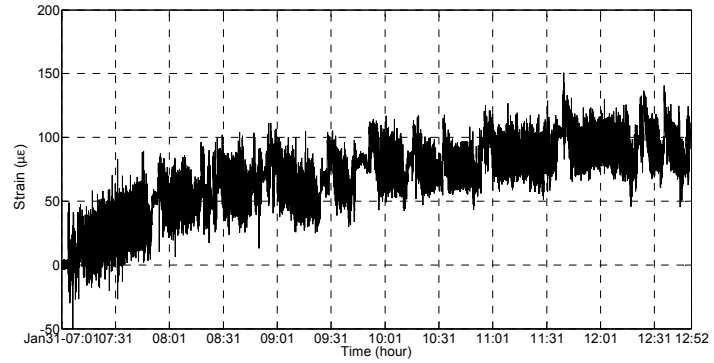
3.3 Demonstration Using Field Monitoring Data

3.3.1 Dynamic strain monitoring data

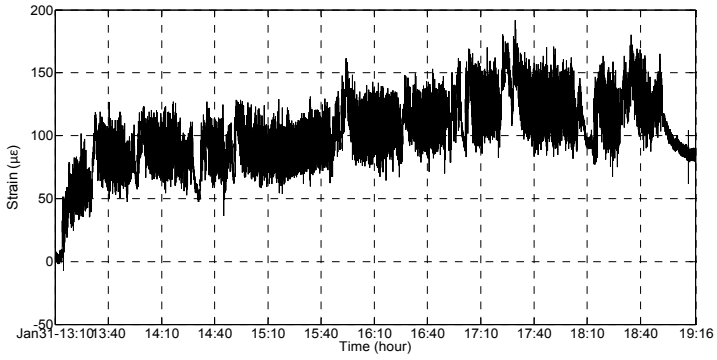
Sensors were implemented on one bogie of an operating high-speed train in China to enable the continuous monitoring of strain and temperature changes. To track the dynamic characteristics of high-speed train bogies under in-service loading, FBG sensors were deployed to monitor strain and temperature variations within the high-speed train bogie. The strain time histories from a deployed FBG strain gauge are used to verify the performance of the proposed modeling method in Subsection 3.2.

Figure 3.3 shows typical strain time histories acquired from the FBG strain gauge after temperature compensation for four train trips. The histories used for further

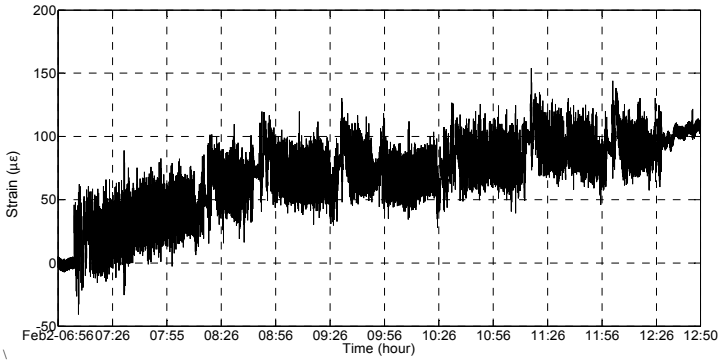
investigation are described in the next section below. In this chapter, "1st trip", "2nd trip", "3rd trip" and "4th trip" denote four monitored train trips.



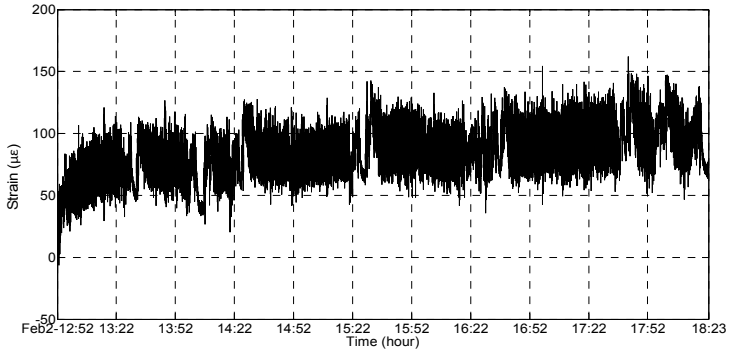
(a)



(b)



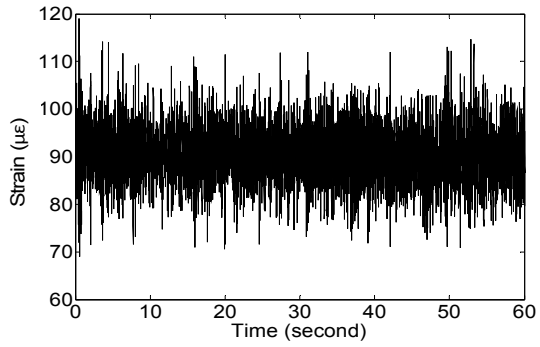
(c)



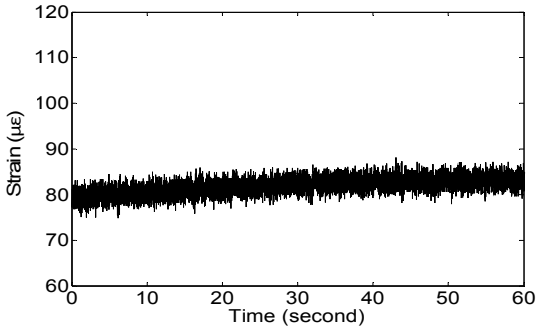
(d)

Figure 3.3 Strain time histories after temperature compensation collected during (a) 1st trip, (b) 2nd trip, (c) 3rd trip, and (d) 4th trip

It is observed that large-amplitude strain variations are generated during each trip when the train runs at normal speed as shown in Figure 3.4(a), whereas low-amplitude strain variations (lower than 7 με) appear in short time periods when the train is standing by at railway stations, as illustrated in Figure 3.4(b). As illustrated in Figure 3.4(a), the strain responses acquired during normal train travel present relatively stable variation, which are the most frequent condition when the train is running at normal speed for all the train trips, and are selected to represent strain variation during normal train operation. It can be seen from Figure 3.4 that there is an overall drift of the strain time histories. It might be caused by temperature variation. Further investigation needs to be made on this issue.



(a)



(b)

Figure 3.4 Strain time history: (a) train running at speeds around 300 km/h, and (b) train on standby at railway station

3.3.2 Computation of stress spectrum distribution

After obtaining the stress time histories by multiplying the measured strain data with the elastic modulus of the low-alloy steel (206 GPa), rainflow cycling counting algorithm is applied to obtain the number of cycles applying to different stress ranges and the corresponding mean stress data. Those stress cycles with amplitudes less than 1.3 MPa are discarded because most of the stress variations generated by noise and spikes of the FBG strain sensors are below this value. Figure 3.5 shows the histograms of one-hour standard stress spectra without the mean stress effect. The stress spectra have a resolution of 0.1 MPa for the stress ranges interval.

3.3.2.1 *Accounting of mean stress effect through constant life equations*

Figure 3.4 shows the scatter distribution of the obtained stress ranges and mean stresses. It is seen that the values for the area where mean stresses distribute are of higher amplitudes than the values of the area for stress ranges.

In analyzing the stress influence on metal fatigue, stress ratio R (the ratio of minimum stress to maximum stress in a stress cycle) is usually used to define the level of mean stress superimposed on the stress range. Figure 3.7 shows the relationship of stress ranges versus R values for the one-hour stress time history recorded by during the 1st trip. The figure shows that the R is mostly distributed from 0 to 1, with the corresponding stress ranges varying from 0 to 20 MPa. This phenomenon indicates that the mean stresses are mostly tensile.

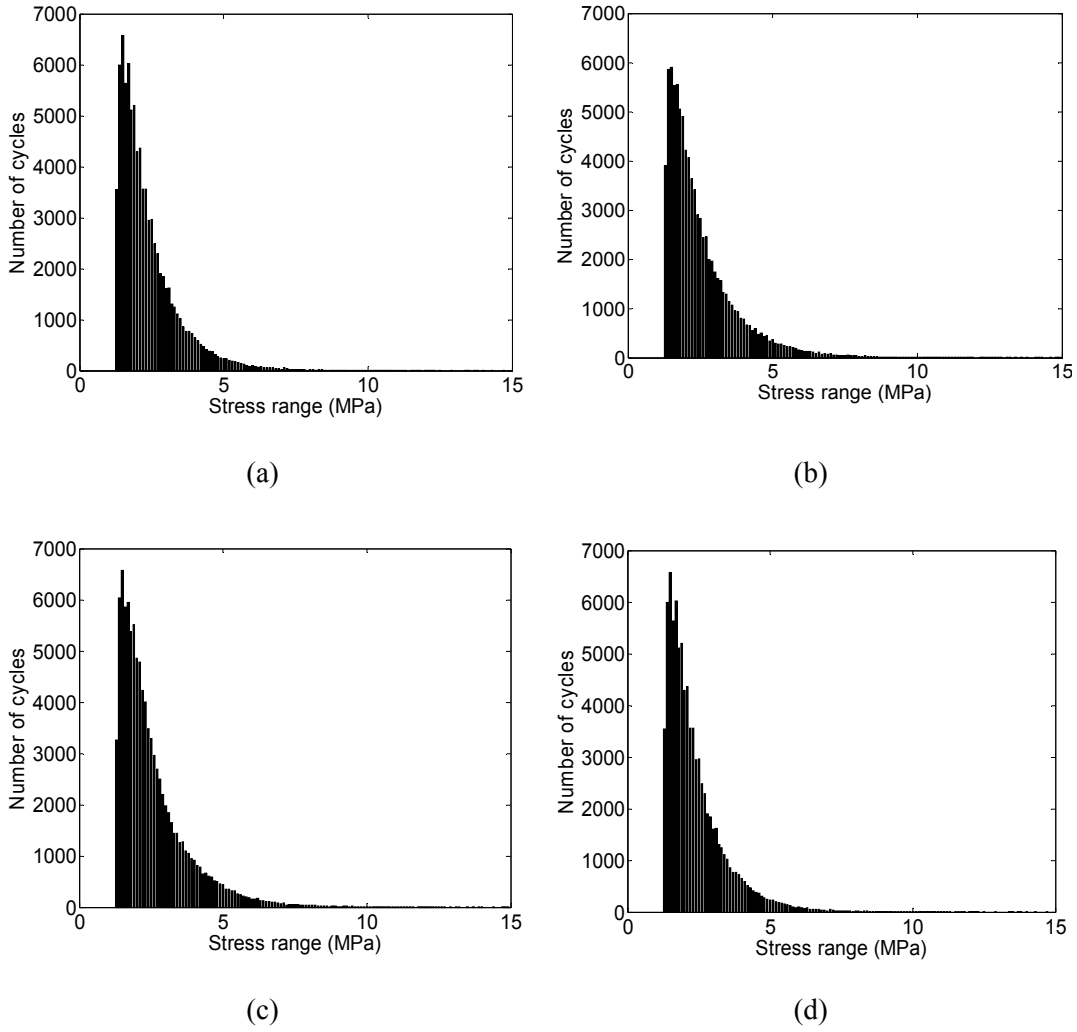


Figure 3.5 Histograms of one-hour stress spectra of stress ranges: (a) 1st trip, (b) 2nd trip, (c) 3rd trip, and (d) 4th trip

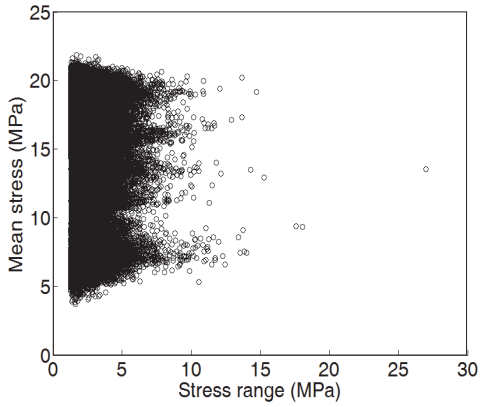


Figure 3.6 Scatter distribution of mean stresses and stress ranges

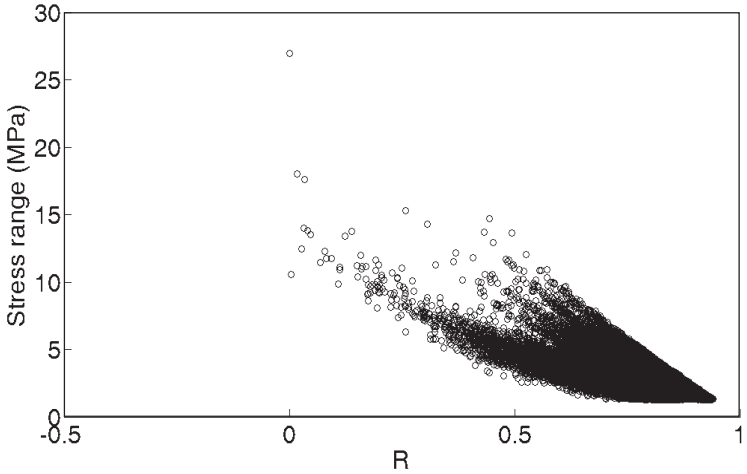


Figure 3.7 Relationship of stress ranges versus stress ratio, *R*

Subsequently, the constant life equations are applied to process the mean stresses so as to derive the effective stress ranges by using Equation (3.5). As discussed in Subsection 3.2.2.2, the Goodman relation is used for illustration and discussion purposes on applying a constant life equation to consider mean stress effects. Therefore, the Goodman relation shown in Equation (3.6) is selected for this study as the method to obtain effective stress ranges and associated effective stress spectra for defining the *S*-*N* relationship.

3.3.2.2 *Derivation of effective stress range*

The histograms of one-hour standard stress spectra, taking into account additional mean stresses (i.e., effective stress spectra) using the Goodman relation, are shown in Figure 3.8. Figure 3.9 compares a stress spectrum with an effective stress spectrum. In the effective stress spectrum, the numbers of cycles are mainly distributed in the higher

value region of effective stress range. In the stress spectrum, the numbers of cycles are mainly distributed in the lower value region.

A stress spectrum and an effective stress spectrum are then represented by PDF using appropriate modeling techniques. In the next subsection, the histograms of the stress spectra and effective stress spectra obtained during the 1st trip are used to determine the most appropriate PDF.

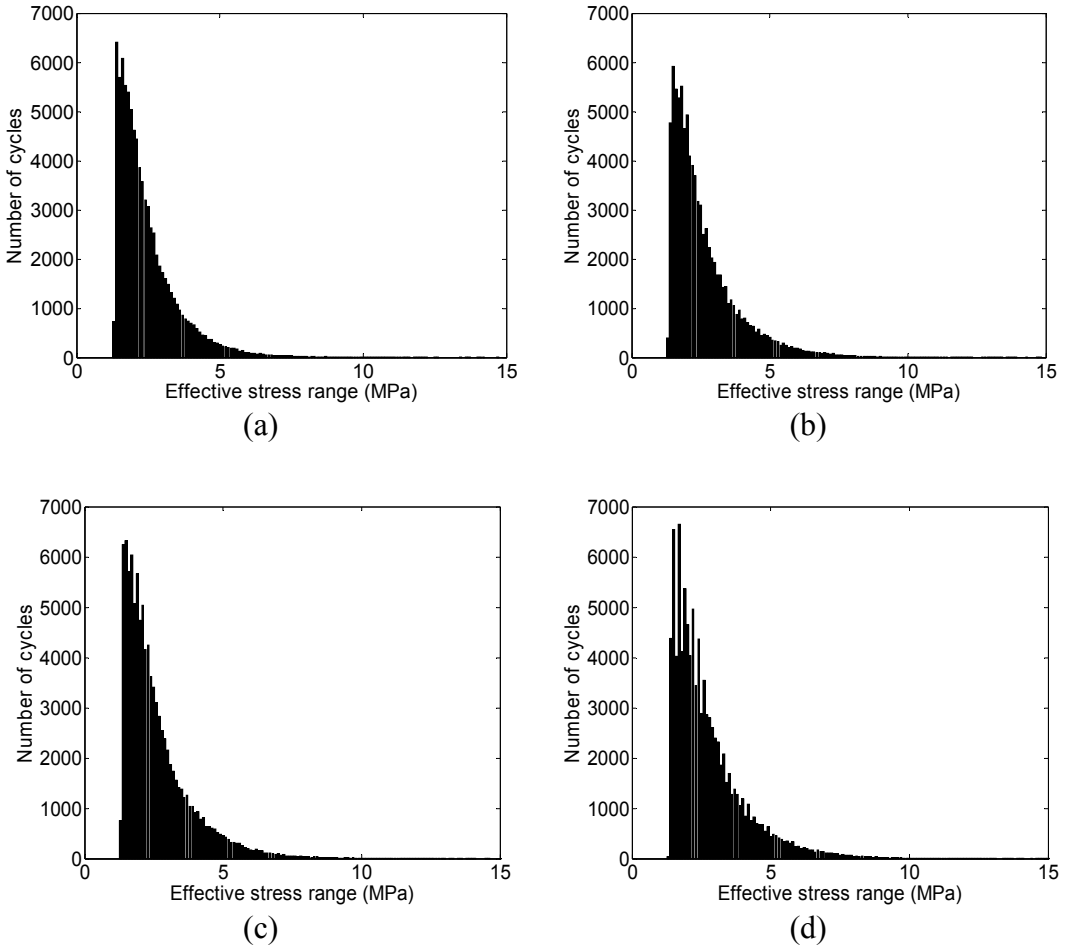


Figure 3.8 Histograms of one-hour stress spectra of effective stress ranges: (a) 1st trip, (b) 2nd trip, (c) 3rd trip, and (d) 4th trip

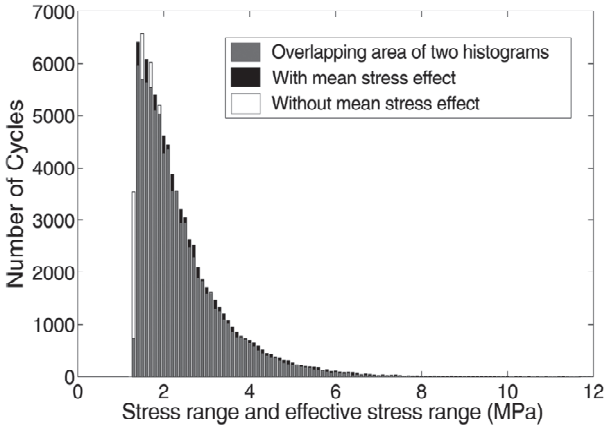


Figure 3.9 Comparison of histograms of typical one-hour stress spectra with and without considering mean stress effect

3.3.2.3 *PDF estimation*

As shown in Figures 3.5 and 3.8, the stress ranges and effective stress ranges are mainly distributed between 1.3 and 12 MPa. For both stress range and effective stress range, the total number of observations (n) of the stress cycles in formulating the one-hour standard spectra is 86,314, and thus, the number of classes obtained is k = 18 according to the Sturges rule (Scott 2009):

$$k = \lceil 1 + \log_2 n \rceil \tag{3.17}$$

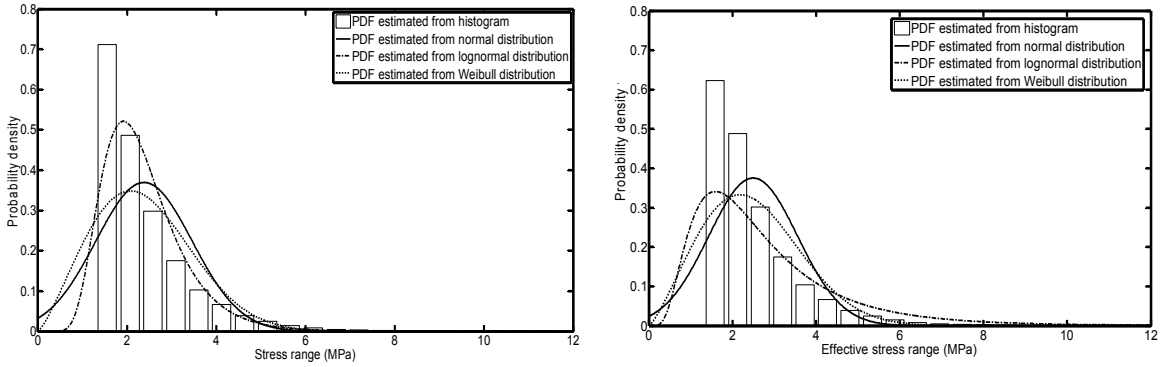
where $\lceil\ \rceil$ denotes the ceiling operator.

To formulate the PDF of stress spectra, the first attempt is to adopt commonly used standard parametric distribution models. Figure 3.10 illustrates the modeling results based on three single distributions (Gaussian, lognormal, and Weibull) for the one-hour stress spectra of both the stress ranges and the effective stress ranges obtained during the 1st trip. It is observed that none of the three distributions fit the histograms well, indicating that the single distribution models are incapable of adequately representing the histograms. In view of this, an alternative solution is sought using finite mixture distributions.

The Gaussian, lognormal and Weibull mixture distributions defined in Equations (3.7) to (3.9) are used to model the stress spectra, respectively. The EM algorithm is applied to identify the unknown parameters and the calculated AIC and BIC values are adopted and compared to determine the most suitable mixture distribution model and the

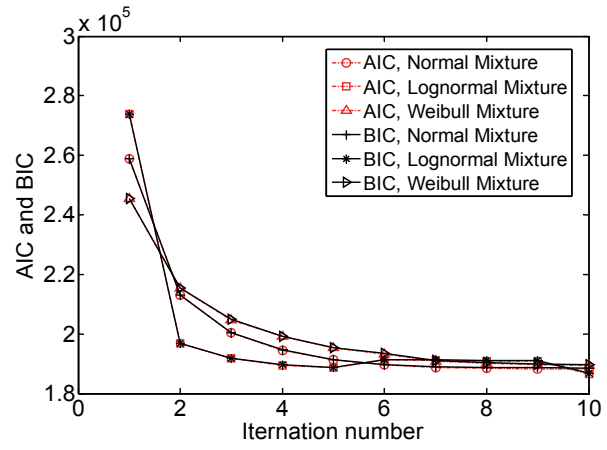
optimal number of components. Figure 3.11 shows the variations of AIC and BIC values for different numbers of components when using the respective three mixture distribution models. It is seen that the values of AIC and BIC converge rapidly with an increasing number of components.

According to the model selection rules of AIC and BIC, the best model has the lowest value (of either AIC or BIC). Fraley and Raftery (1998) suggested selecting the number of components where the values of AIC or BIC seem to converge. Based on the rules, the lognormal mixture distribution with five and four components is selected for modeling the stress spectra of stress ranges and effective stress ranges. Figures 3.12 and 3.13 show the modeling results of the one-hour stress spectra for the stress ranges and the effective stress ranges for the respective four trips. Table 3.1 gives the identified parameters of component distributions using the lognormal mixture distribution for modeling the stress and effective stress spectra from the 1st trip. It is seen that the lognormal mixture distribution is competent to represent both the stress and effective stress spectra. With the models formulated, the stress spectra information, beyond the range of measured data, can also be obtained by extrapolation.

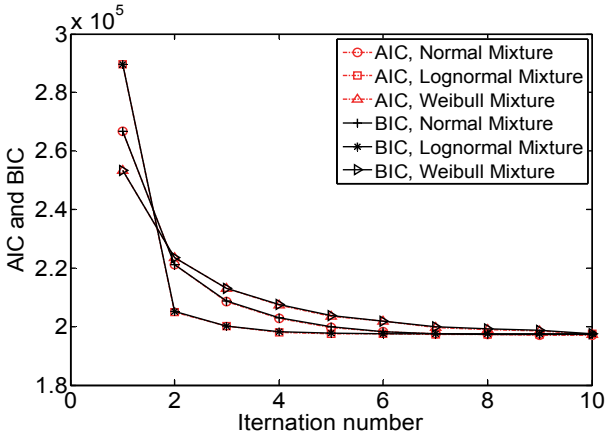


(a) (b)

Figure 3.10 Gaussian, lognormal and Weibull mixture PDF modeling for stress spectra: (a) stress ranges, and (b) effective stress ranges



(a)



(b)

Figure 3.11 AIC and BIC values for modeling of mixture PDFs of stress spectra: (a) stress ranges, and (b) effective stress ranges

Table 3.1 Identified parameters of component distributions using lognormal mixture distribution for modeling stress and effective stress spectra

Parametric family		Values of parameters		
		Weight (w)	Mean value (μ)	Standard deviation (σ)
Lognormal	Stress range	0.2629	0.6251	0.1255
		0.3163	0.9173	0.2117
		0.0808	0.3118	0.0329
		0.1946	1.2793	0.3437
		0.1454	0.4316	0.0678
	Effective stress range	0.3596	0.7190	0.0349
		0.3166	1.0825	0.0752
		0.0887	1.4317	0.1406
		0.2351	0.4433	0.0178

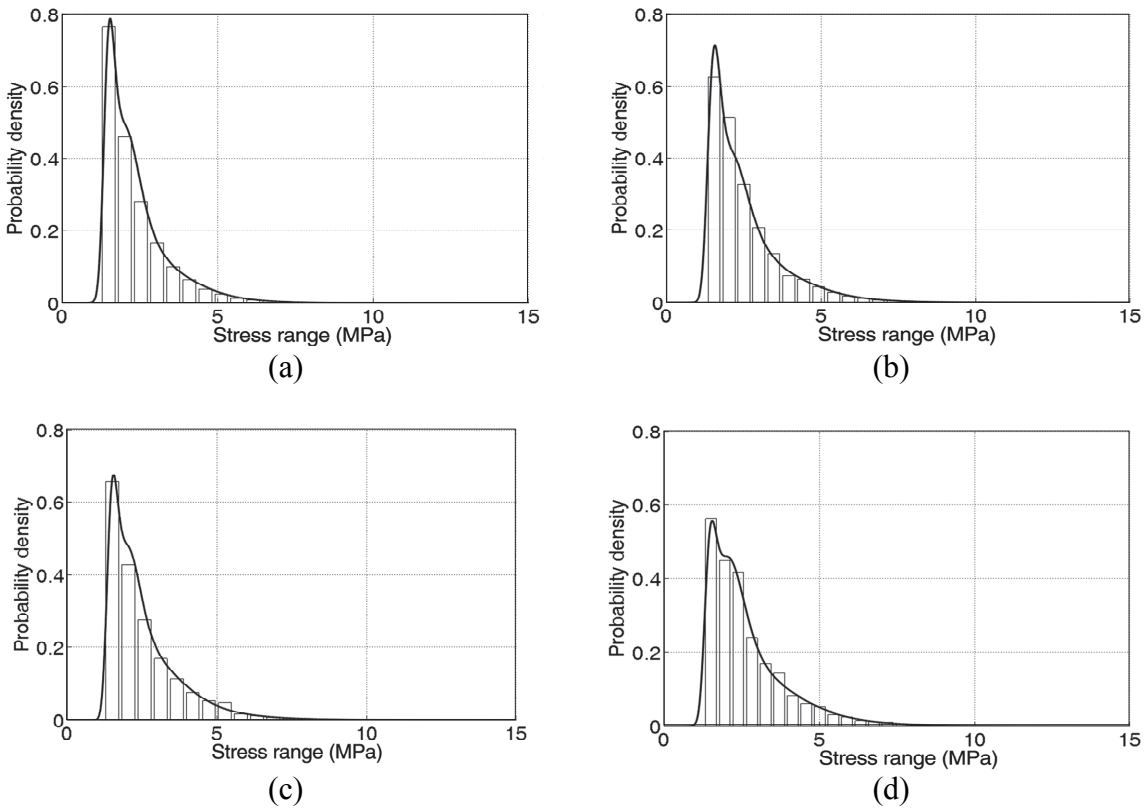


Figure 3.12 Lognormal mixture PDFs of one-hour stress spectra of stress ranges data from: (a) 1st trip, (b) 2nd trip, (c) 3rd trip, and (d) 4th trip

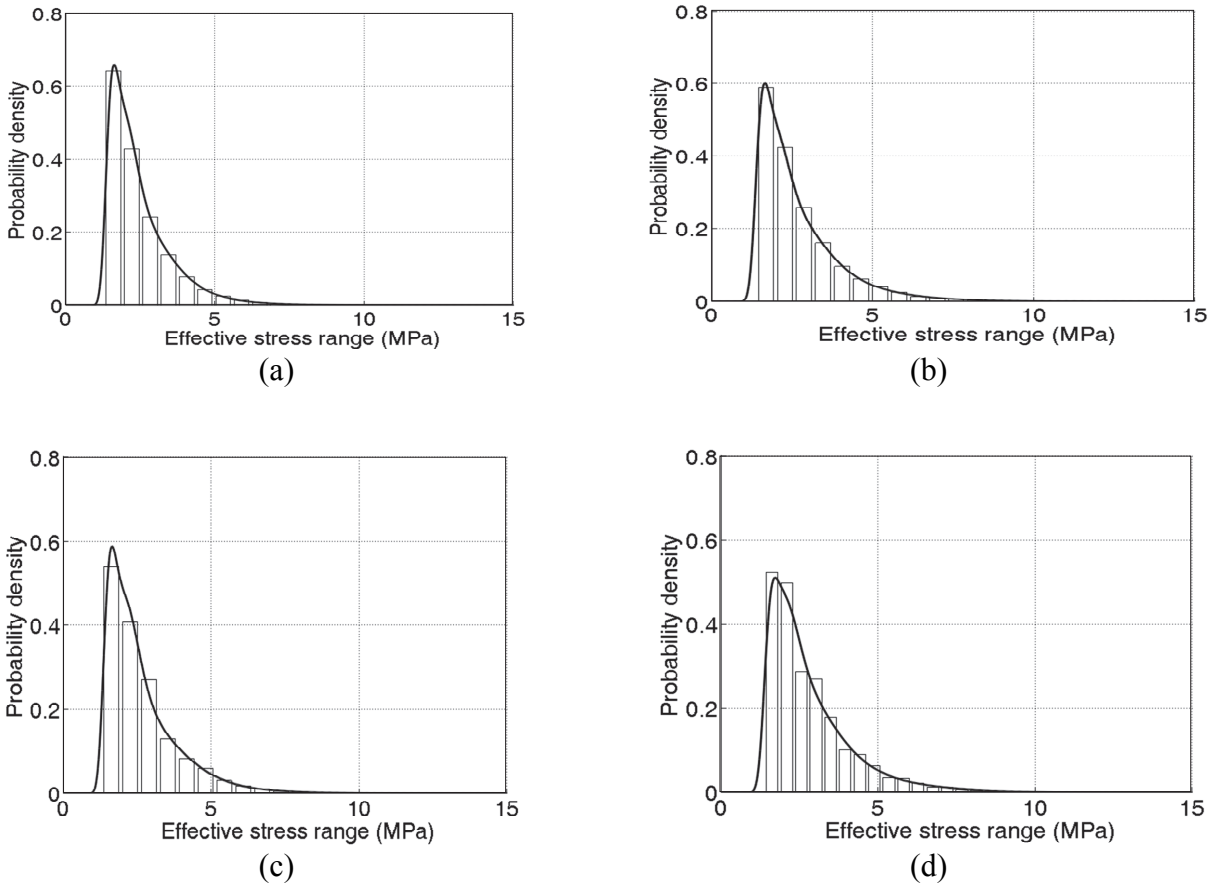


Figure 3.13 Lognormal mixture PDFs of one-hour stress spectra of effective stress ranges data from: (a) 1st trip, (b) 2nd trip, (c) 3rd trip, and (d) 4th trip

3.4 Summary

The monitoring stresses acquired from instrumented high-speed trains can be used for fatigue life assessment. The amplitudes of monitoring data commonly contain variations and show similar pattern for different train trips, in which the general pattern and the tail behavior need to be properly described.

In this chapter, a method has been developed for modeling the stress spectra acquired from instrumented high-speed trains. The method can describe the multiple effects in the model by the use of finite mixture distributions with frequentist EM estimation. This modeling procedure is applied to the in-service monitoring data of components of a train running on a high-speed railway in China. Both stress and effective stress spectra are formulated using the monitoring data. The finite mixture distribution modeling technique is used in conjunction with EM algorithm to formulate PDF models. The models adequately represent the distribution characteristics of the monitoring-derived stress spectrum. According to these research work accomplished, several points can be drawn as follows:

1. Similar patterns can be found for the stress ranges distributions of repeated train trips. Therefore, it is appropriate to formulate a representative PDF extracting general features and describing the tail behaviors of the stress spectra of repeated train trips;
2. As a single distribution with frequentist modeling technique is not adequate to describe the behavior of stress spectrum, it is more appropriate to employ mixture distributions to model the PDF of stress spectrum;
3. The best fit PDF for stress spectrum based on mixture distribution can be selected by using AIC or BIC rules. It is found that the distributions of stress spectrum and effective stress spectrum are well described by lognormal mixture distributions; and

4. The proposed frequentist modeling technique offers a feasible solution for describing the stress spectra acquired from in-service stress time histories. The fitted PDF is useful for further construction of a fatigue life assessment approach, which will be introduced in the next chapter.

CHAPTER 4

FATIGUE LIFE ASSESSMENT OF HIGH-SPEED TRAINS USING IN-SERVICE MONITORING DATA: FREQUENTIST APPROACH

4.1 Introduction

The past decades have witnessed rapid development in SHM technology, especially for railway applications (Chong et al. 2010). Different types of on-board monitoring systems have been invented and implemented to monitor and evaluate the conditions of trains running on railways. It is now possible for authorities to adopt in-service monitoring data for integration of stress or load spectra to conduct fatigue life assessment of train components (Beretta et al. 2011; Grubisic and Fischer 2012; Hong et al. 2014). The standard procedure for fatigue life assessment has been investigated in previous studies using measured data (Agerskov 2000, Sonsino 2007; Kassner 2012; BS7608 2014).

The fatigue life of a train component is mainly determined by the stress behavior imposed on its critical locations, such as welded joints (Kassner 2012). Therefore, an accurate estimation of the actual fatigue life of a welded joint by the in-service stress

histories is of great importance. It is known that the uncertainties in the stress histories that alter fatigue life. However, these uncertainties are not well considered by the current common practice for fatigue life assessment. Therefore, it is desired to develop a fatigue life assessment approach considering the uncertainties inherent in repeated stress histories.

Although most of the specifications are in favor of using nominal stresses acquired from the original measurements (BS7910 2005; EN13749 2005; Castillo 2009; Hobbacher 2009; BS7608 2014), the stress concentration that is not considered by nominal stresses affects the fatigue life of a welded joint (Niemi and Marquis 2002; Xiao and Yamada 2004). The hot spot stress as mentioned in Section 2.3.3, unlike the nominal stress, relates to the stress concentration effect at the critical point, i.e., the weld toe, where fatigue cracks are expected to initiate and propagate. The choice of either using the nominal stress or the hot spot stress can significantly influence the calculation results of fatigue life assessment.

For stress histories, in addition to well-known stress ranges which cause fatigue failure and life reduction, mean stress is also recognized to affect fatigue life (Stephens et al. 2000). The mean stresses of the cyclic stress responses on bogie components represent the initial static loadings, thermal changes, and loading changes during train operation (Kihl and Sarkani 1999). Vincent et al. (2012) pointed out that sufficiently high mean stress shortens fatigue life under an unchanged strain range level. The work by Vincent and his colleagues was conducted with strain-control load tests under different

constant mean stress values. However, Radaj and Vormwald (2013) declared that the mean stress effect on the fatigue performance of a welded joint is a secondary effect, and the amplitude of the equivalent fully oscillating stress range decreases in the case of a tensile mean stress and increases with a compressive mean stress. In addition, Moore and Booth (2014) reported that mean stress has no influence on the fatigue performance of a welded joint. These different opinions reveal that influence of mean stress on fatigue life needs further investigation.

The influence of mean stresses on the fatigue life of an in-service structure is not well understood. Many previous research articles on fatigue were generally based on the assumption that no mean stress is considered or mean stress has a constant amplitude with little variation. This is against the reality that mean stress amplitudes can vary frequently during in-service operation of a structure (Nicholas and Zuiker 1989). Because of the variability of in-service mean stresses, there is a need to develop an effective method for assessing the mean stress effect on fatigue life. There are several empirical algorithms known as “constant life diagrams” (e.g., Goodman, Gerber, Morrow, and Soderberg diagrams) to take mean stress effects into account (Susmel et al. 2005). The mean stress effect on fatigue is assessed using these equations. Although the approaches of the empirical equations need further solid theoretical support (Schijve 2001), reliable estimates of the fatigue life can be ensured.

The study on in-service mean stresses for fatigue can be further explored because of the advancement in measurement techniques and computing capacity in recent decades

(Schijve 2001). In the study presented in this chapter, it is found that the in-service stress time history reveals time-dependent statistical scatter, and also the associated distribution of in-service mean stresses exhibits multimodality. This behavior calls for good interpretation of the multimodality for fatigue life assessment. This requirement can be satisfied through employing different mixture PDFs to fit the in-service mean stresses. It has been shown in Chapter 3 that Gaussian, lognormal and Weibull mixture distributions are applicable and lead to different model fitting results. Selection rules such as AIC and BIC can help choose the best distribution model (Schwarz 1978).

In this chapter, a fatigue assessment method is developed that can consider in-service mean stresses by introducing an indicator Q. Analysis of mean stresses is initially conducted for the monitoring data at a bogie welded joint of an in-service high-speed train running for one month on a high-speed railway line in China. Monitoring data reveals the in-service mean stresses have multimodality characteristics, which cannot be interpreted by traditional probability models. To solve this problem, statistical analysis using mixture PDF is conducted on in-service mean stresses. This PDF of mean stresses is then integrated together with the PDF of stress ranges to formulate a method for fatigue life assessment. The method is verified using the monitoring data from a bogie welded joint of an in-service high-speed train.

4.2 Probabilistic Approach for Fatigue Life Assessment

4.2.1 Assessment procedure integrating monitoring data and PDF

The widely used method for fatigue life assessment is based on *S-N* curve. *S-N* curve is obtained from specimen cycling tests under constant stress amplitude. It describes the relationship between the number of cycles to failure, *N*, and the amplitude of the corresponding stress range, *S* (Hobbacher 2009):

$$NS^m = C \tag{4.1}$$

where *C* is a constant for a given material and corresponding to a specific class of structural detail, and the exponent *m* is a material parameter that determines *S-N* curve slope.

For stress cycles with variable amplitudes and, the fatigue ratio, *D*, can be expressed by Miner's rule as (UIC 615-4 1994; EN13749 2005; Hobbacher 2009):

$$D = \sum_{i}^{k} \frac{n_i}{N_i} = \frac{n_1}{N_1} + \frac{n_2}{N_2} \ldots + \frac{n_k}{N_k} \tag{4.2}$$

where n_i is the number of stress cycles at the ith stress range (S_i), N_i is the number of stress cycles to failure at stress range S_i, and k is the number of classes for the stress range. The nominal value of the fatigue ratio to failure D_f is 1. In many cases, however, the Palmgren-Miner theory leads to non-conservative life predictions and therefore the value 0.1 is suggested for D_f for the welded components of a bogie structure under normal service condition for more conservative assessment (Hobbacher 2009).

Considering the contribution of mean stress (σ_m) at the ith stress range (S_i) in causing fatigue failure, the number of ith class stress cycles in Equation (4.2) are found using the stress range and mean stress as:

$$n_i \equiv n(S_i, \sigma_m) \tag{4.3}$$

Equation (4.2) can then be expressed as:

$$D = \sum_i^k \frac{n(S_i, \sigma_m)}{N(S_i, \sigma_m)} \tag{4.4}$$

An integrated stress range, S_l, is defined as the cyclic stress level that will produce a specific damage D equivalent to that was caused by the combined effect of stress range and mean stress. Equation (4.4) can then be expressed as:

$$D = \sum_i^k \frac{n(S_i, \sigma_m)}{N(S_i, \sigma_m)} = \sum_i^k \frac{n(S_{l,i})}{N(S_{l,i})} \tag{4.5}$$

where $S_{l,i}$ is the integrated stress range for the ith of k classes. The integrated stress range can be obtained through a series of constant life equations summarized by (Susmel et al. 2005):

$$\left(\frac{S}{S_l}\right)^r + \left(p\frac{\sigma_m}{S_u}\right)^t = 1 \tag{4.6}$$

where p, t, and r are constants, S_u is the ultimate tensile strength of the metal, and the value of the integrated stress range is expressed in terms of stress:

$$S_l = \frac{S}{\left[1-(p\frac{\sigma_m}{S_u})^t\right]^{1/r}} \tag{4.7}$$

A new indicator Q is introduced and defined as:

$$Q \equiv \frac{1}{\left[1-(p\frac{\sigma_m}{S_u})^t\right]^{1/r}} \tag{4.8}$$

Equation (4.7) can then be simplified as:

$$S_l = S \times Q \tag{4.9}$$

Since stress range S and mean stress σ_m, are two independent random variables, S and Q are also independent of each other. Thus, the joint PDF of the effective stress spectrum for integrated stress ranges can be appropriately expressed as an explicit function $f(S_l)$:

$$f(S_l) = f_1(s) \times f_2(q) \tag{4.10}$$

where $f_1(s)$ is the PDF of stress ranges S, where s is the scope of the obtained stress ranges categorized into k classes; $f_2(q)$ is the PDF of variable Q relating to mean stress. Using the definition of Q cited in Equation (4.8), the fatigue cycles in Equation (4.3) can be re-sorted into stress range and Q, namely:

$$n_i = n(S_i, \sigma_{m,\,\mathrm{i}}) = n(S_{l,\,i}) \Rightarrow n(S_i, Q) \tag{4.11}$$

Accordingly, the Miner's rule in Equation (4.2) can be rewritten for the evaluation of fatigue ratio, with a total of n_{tot} stress cycles in the ith of k classes. It is corresponding to the integrated stress ranges, by defining the following equation:

$$D = \sum_{i}^{k} \frac{n_{tot} \times (S_l)^m}{N_i \times (S_l)^m} = \frac{\int_L n(S_i, Q) \times (S_l)^m \times f(S_l) dS_l}{N(S_i, Q) \times (S_l)^m} = \int_S \int_Q \frac{(S_l)^m \times n_{tot} \times f_1(s) \times f_2(q)}{N_f \times (S_l)^m} dsdq \tag{4.12}$$

It is simplified as:

$$D = \int_S \int_Q \frac{n_{tot} \times f_1(s) \times f_2(q)}{N_f} dsdq \tag{4.13}$$

where N_f is the number of stress cycles at failure corresponding to integrated stress ranges S_l. n_{tot} is an accumulative factor relating to number of cycles for each of k classes and is calculated according to the following rule: the conventional assumption is that *S-N* curve terminates at the fatigue limit point, below which failure will not occur, or become a horizontal line. For rail structures, the stress ranges below the knee point (also referred to as "constant amplitude fatigue limit," below which the slope of *S-N* curve declines) are included by introducing a reduction factor λ_i for n_{tot} (Kassner 2012):

$$n_{tot} = \sum_{i=1}^{k} \lambda_i \times n_i \tag{4.14}$$

and

$$\lambda_i = \begin{cases} (S_i / \Delta\sigma)^{m-1} & \text{if } S_i \le \Delta\sigma \\ 1 & \text{if } S_i > \Delta\sigma \end{cases} \tag{4.15}$$

where n_i is the number of stress cycles as defined in Equation (4.2), m is the material parameter as defined in Equation (4.1), $\Delta\sigma$ is the knee point value which can be obtained from the specification of IIW 2008, where the value of $\Delta\sigma$ is calculated and catalogued according to a FAT class defined in IIW 2008 according to different

geometric form of the welded joint. For monitoring-based fatigue life evaluations, the fatigue ratio, referring to a set of measured strain time historical data, is expressed as D_m. The fatigue life F is then calculated by (Baek et al. 2008):

$$F = \frac{D_f}{D_m} = \frac{D_f}{\int_S \int_Q \frac{n_{tot} \times f_1(s) \times f_2(q)}{N_f \times H} dsdq} \tag{4.16}$$

where H is the traveling distance of the monitored train during the data collection period of the measured strain or stress time history, and D_f is the fatigue ratio at failure.

It can be observed from Equation (4.16) that the value of fatigue ratio, shown as D_m, is determined by the number of stress cycles as well as the distribution of Q. As Q demonstrates the effect of in-service mean stresses on fatigue life, the proposed method is capable of exploring fatigue life which is affected by the combined influence of stress range and mean stress.

The aforementioned method requires the application of several statistical techniques, for example, (1) rainflow cycle counting method to derive the stress range spectrum and the distribution of mean stresses from the measurements of the train components (Kassner 2012), (2) the constant life equations to quantify mean stresses with indicator Q, and (3) the mixture distributions model to represent the respective distributions of stress ranges and mean stresses (Schwarz 1978). Based on the above derivations, the influence of mean stresses can be taken into account based on the method described in this chapter. A flowchart of the method is shown in Figure 4.1.

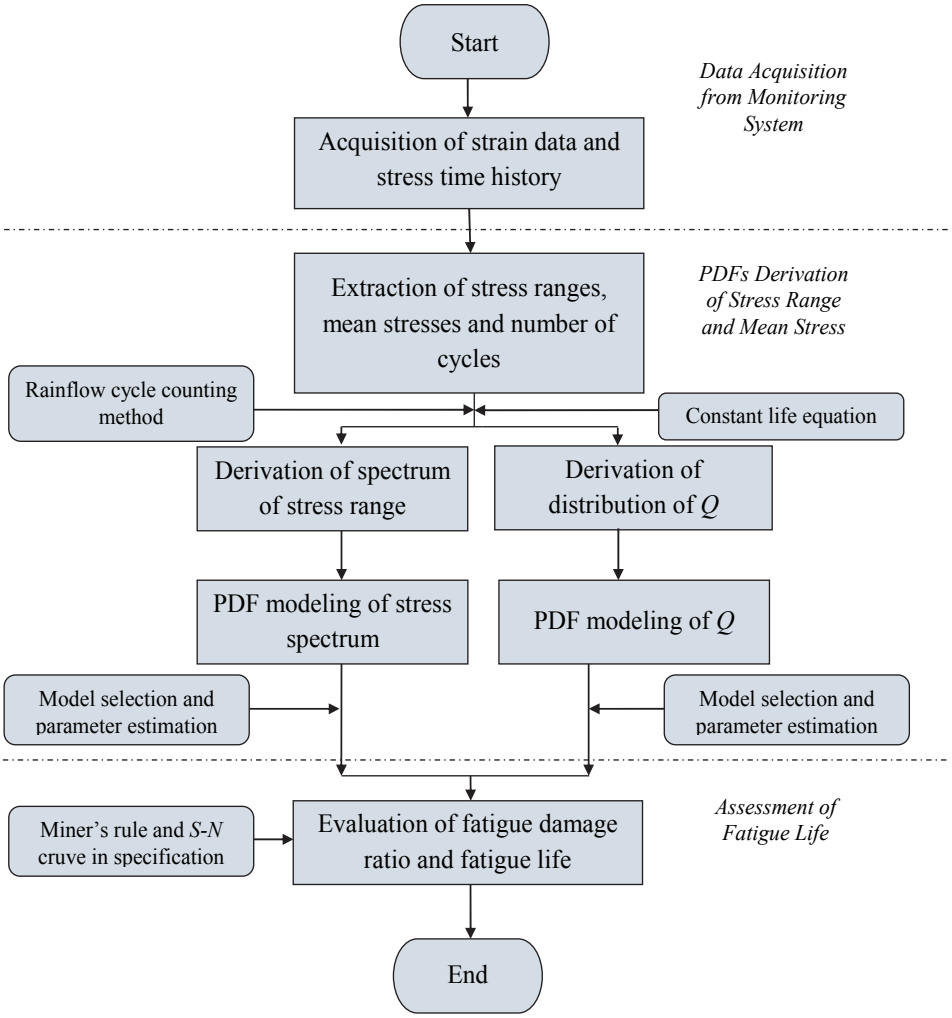


Figure 4.1 Flowchart of using fatigue life assessment method

4.2.2 Parameter used for defining mean stress effect

The effect produced by mean stresses in the fatigue life assessment method is represented by the indicator Q in Equations (4.6) to (4.9). The expression of Q for mean stress effect is based on the relationship defined by the "constant life equations" shown in Equation (4.6). With different values for constants p, t, and r, the Q expression changes for describing the mean stresses. Previous research proposed

several empirical algorithms to define the "constant life equations" expressions exactly, including the Goodman relation, the Soderberg relation, the Gerber parabola, the Dietman parabola and the so-called "elliptical relation" (Susmel et al. 2005). They define the constants *p, t,* and *r* in the "constant life equation" of Equation (4.6) with different values. Among the empirical algorithms, the Goodman relation is the most fundamental one in accounting for mean stresses (Sendeckyj 2001). Since other forms of empirical algorithms can also use the Goodman relation as a basis to define the indicator Q for fatigue life assessment, the Goodman relation is selected for demonstrating the application of the "constant life equation" to process the monitored mean stresses data to suit indicator Q, for the purpose of fatigue life assessment. According to Equation (4.8), the Goodman relation and associated expression for Q, are expressed as:

$$S_e = \frac{S_a}{1 - \frac{\sigma_m}{S_u}} \tag{4.17}$$

$$Q = \frac{1}{1 - \frac{\sigma_m}{S_u}} \tag{4.18}$$

4.3 Application: Fatigue Analysis of Welded Joint of High-speed Train

4.3.1 Description of on-board monitoring sensing system and data acquisition

Sensors were implemented on a bogie of an operating high-speed train in China to enable continuous monitoring of strain and temperature variations. The monitoring was conducted during normal operation period of the train for one month from December, 2015 to January, 2016.

The monitored bogie is motorized with a traction system, which makes the bogie not only crucial in guaranteeing the safety of the vehicle, but also responsible for driving high-speed train. Additionally, the bogie maintains the stability of the vehicle on both straight and curved tracks by minimizing the impact of centrifugal forces generated during the train turning. The whole bogie frame structure is a welded steel box made of high-strength low-alloy SJ335G steel with a thickness of 20 mm.

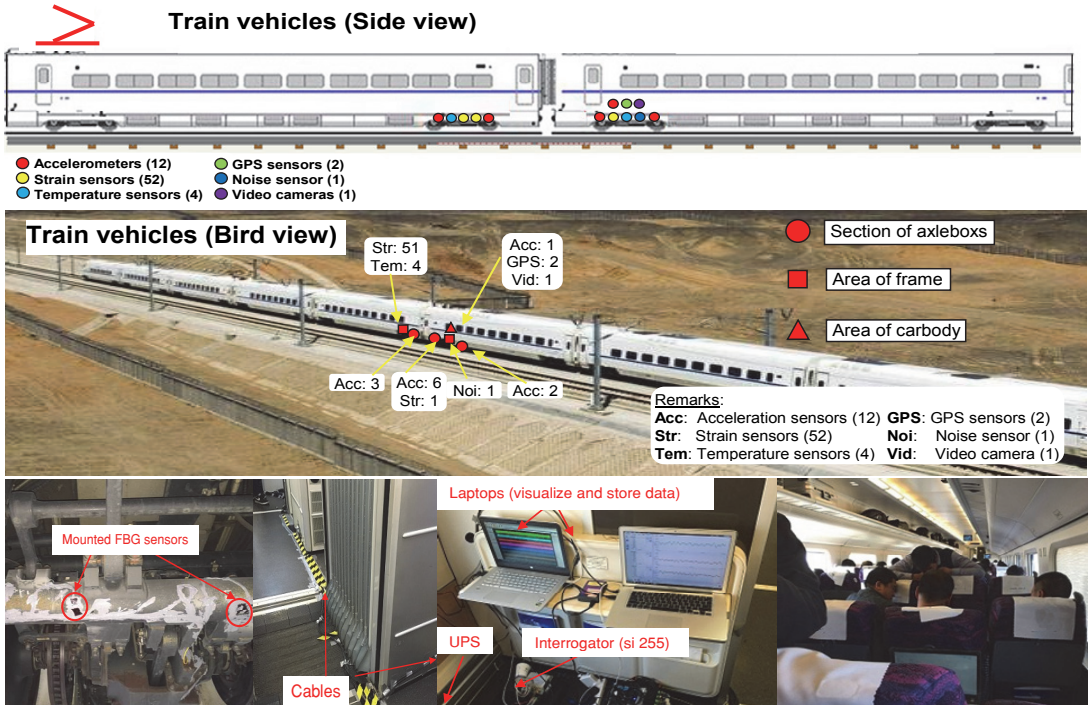
To track the dynamic characteristics of an in-service high-speed train bogie, FBG strain sensors, temperature sensors, accelerometers, and other types of sensors were employed separately, as shown in Figure 4.2. A total of 52 FBG strain sensors, 4 FBG temperature sensors, 11 accelerometers (multi-directional) and 1 noise sensor were installed on the high-speed train bogies. They were deployed at 6 measurement zones.

The 6 measurement zones are: (1) Zone near secondary suspension system, with 9 strain, 1 temperature and 1 noise sensors; (2) Zone of stresses transmission, with 21 strain and 1 temperature sensors; (3) Zone of crucial areas according to FEM calculation, with 6 strain sensors; (4) Zone of welded joints, with 12 strain and 1 temperature sensors; (5) Zone of bumper, with 1 strain sensor and 11 accelerometers; (6) Zone of gearbox, with 3 strain and 1 temperature sensors. In addition, there were also 1 FBG accelerometer (multi-directional), 2 GPS sensors and 1 video camera installed inside the carbody of the train.

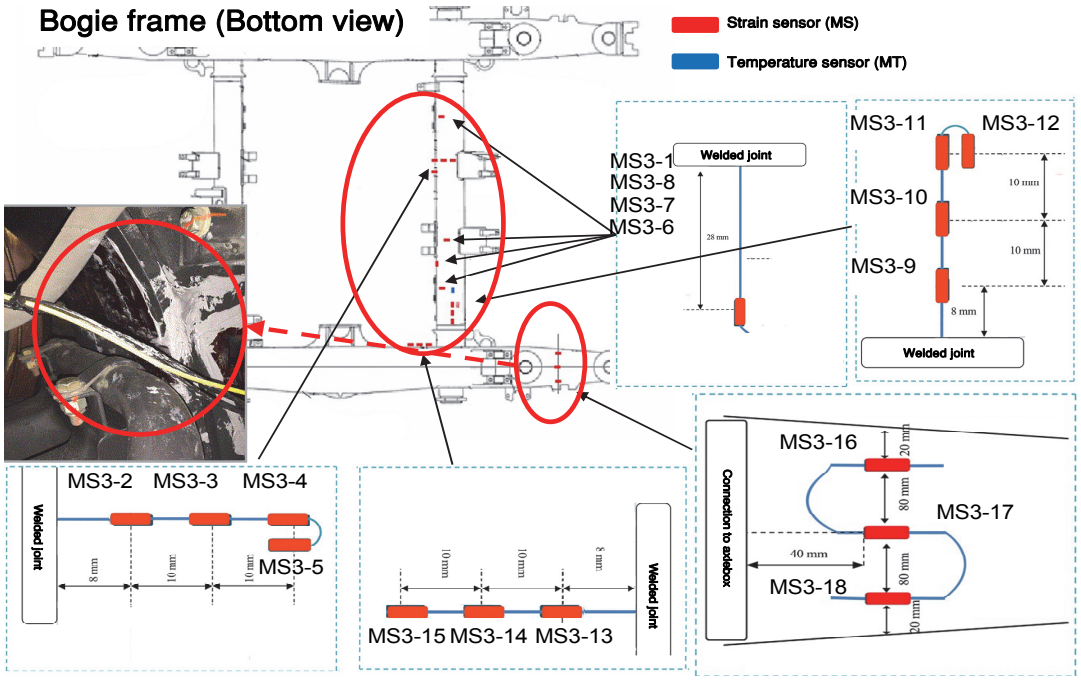
Among the sensors, all FBG strain sensors were installed at fatigue-prone locations of the bogie components including some welded joints. The FBG sensors were deployed and attached to the surface of the train bogie and other components with industrial glue and tapes. The FBG sensors were connected to the interrogator (si 255) positioned inside the carbody through optical fiber cables. The interrogator was connected to a laptop which had large storage and could real-time display the monitoring data from different sensors. From the interrogator, there is a broadband light spectrum transmitting through the connecting cable with numerous FBG sensors. In each FBG sensor, there is a gauge which has a length of 10 mm installed right above the measured point of the surface. The gauge of a FBG sensor will reflect wavelengths with very small bandwidths from the original broadband light spectrum. The small bandwidths will be received by the interrogator. As long as the grating does not change, the bandwidths received by the interrogator for the FBG sensor remain unchanged. Once

the position deployed with an FBG sensor experiences stress or load, the grating will compress or expand, varying the magnitude of the corresponding small bandwidth reflecting in the broadband light spectrum and being received by the interrogator. The interrogator will calculate the difference of the newly obtained bandwidth with the former bandwidth, and then multiply this difference with a specific factor (unit: με/pm) to convert as a strain value, which is regarded as the strain change at the measured location.

The FBG sensors were installed at the surface and at specific locations adjacent to the weld toe. The nearest gauge to the weld measured was at 8 mm near the weld toe. Seven locations near the welded joints that suffer higher stress responses according to a finite element analysis were selected as the points for deploying strain sensors. A total of 15 FBG strain sensors and 1 temperature sensor were installed at these seven measurement points. Detailed information on the locations of the sensors is shown in Table 4.1. The strain data were recorded for all of the train trips for one month, at sampling rates of 1000 - 5000 Hz, with a resolution of 0.01 με. Once the train started to run, the interrogator was triggered to collect data from the connected sensors and then the collected data were stored in the laptop.



(a)



(b)

Figure 4.2 In-service monitoring of high-speed train: (a) layout of all deployed sensors, and (b) locations of strain sensors deployed on welded joints of bogie frame

Table 4.1 Locations and serial numbers of strain sensors on bogie

Measurement point	**Location of welded joint**	**Serial number (distance from welded joint)**		
1	Tubular crosspiece and gearbox connection, A	MS3-1 (28 mm)		
2	Disk brake connection	MS3-2 (8 mm)	MS3-3 (18 mm)	MS3-4 (28 mm)
		MS3-5 (28 mm)		
3	Tubular crosspiece and gearbox connection, B	MS3-6 (28 mm)		
4	Tubular crosspiece and gearbox connection, C	MS3-7 (28 mm)		
5	Longitudinal connection to carbody	MS3-8 (28 mm)		
6	Tubular crosspiece edge	MS3-9 (8 mm)	MS3-10 (18 mm)	MS3-11 (28 mm)
		MS3-12 (28 mm)		
7	Connection of tubular crosspiece and frame	MS3-13 (8 mm)	MS3-14 (18 mm)	MS3-15 (28 mm)

Figure 4.3 illustrates the installations of four FBG strain sensors in two parallel lines at the connection of a tubular crosspiece and frame, which is the 7th measurement point. The stress responses of this location are complex, according to FEM analysis results. The strain time histories recorded at the 7th measurement point, therefore, are of primary interest.

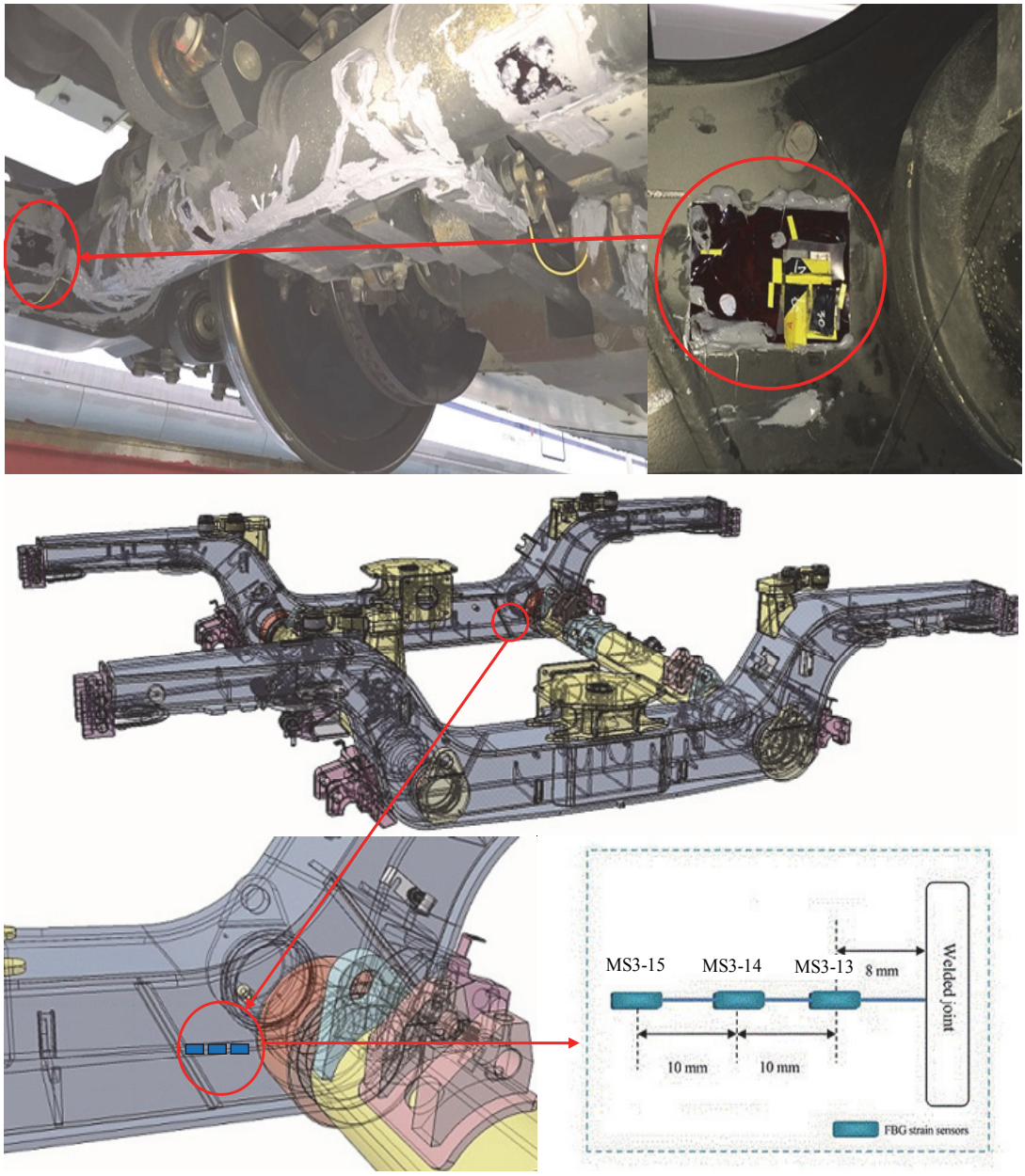


Figure 4.3 Location of 7th measurement point and three mounted strain sensors

It is understood that the stress concentration effect plays a crucial role in enlarging the stress at the weld toe, therefore the hot spot stress measured by sensors should be used for calculation of the fatigue life of welded joint (Hobbacher 2009). The hot spot stress responses of the sensors deployed near the welded joint can be calculated as:

$$\sigma_{hs} = 2.52\sigma_{0.4t} - 2.24\sigma_{0.9t} + 0.72\sigma_{1.4t} \tag{4.19}$$

where $\sigma_{0.4t}$, $\sigma_{0.9t}$, and $\sigma_{1.4t}$ are the stress data measured by the sensors near the weld toe, deployed at the distances of 0.4, 0.9, and 1.4 times of the thickness.

According to the guidance, the sensors in Figure 4.3, MS3-13, MS3-14, and MS3-15 are 8, 18 and 28 mm, respectively, from the weld toe, or 0.4, 0.9, and 1.4 times of the thickness of the steel plate. The data acquired from sensors MS3-13, MS3-14 and MS3-15 installed near the weld at the 7th measurement point is used in this study.

A comparison of the measured hot spot stress and the FEM-simulated hot spot stress at the welded joint of 7th measurement point (i.e. connection of a tubular crosspiece and frame) is provided as follows: According to the FEM analysis, the largest stress generated by a modeled high-speed train under operating load is 208 MPa, and the mean of stress ranges is 61.89 MPa. According to the monitoring result, the largest stress generated by an in-service train is on average 250 MPa, and the mean of stress ranges is on average 6.59 MPa.

This shows that the largest measured stress is 42 MPa larger than the FEM-simulated one, whereas the mean of stress ranges is 55.3 MPa smaller than the FEM-simulated

one. The reason for this phenomenon may be due to that the extreme condition experienced by an in-service high-speed train would generate relatively large stresses in comparison with the FEM-simulated condition; while there are more stress ranges of small amplitudes generated by the in-service train running on the high-speed railway, making the overall mean value of stress ranges become much lower than the FEM-simulated results which only consider several classical operating conditions with influencing loads.

4.3.2 Data preprocessing

The strain time histories from the FBG strain gauges MS3-13, MS3-14 and MS3-15 during three train trips at the monitoring period are selected for further data processing. The three trips took place on (a) January 3, (b) January 5, and (c) January 9, 2016, when high-speed train repeated running from one station to another, lasting around 1 hour. Below, the trips are denoted as "1st trip", "2nd trip" and "3rd trip".

These selected strain time histories are used to analyze the influence of mean stresses in the fatigue life assessment. The stress time histories of the sensors are obtained by multiplying measured strain values by the modulus of elasticity of steel SJ335G as the derived stress values are all lower than the yield strength of the material (355 MPa). Figure 4.4 shows a comparison among the stress histories of strain sensors MS3-13, MS3-14 and MS3-15 from the 1st train trip. The comparison reveals that the stress

responses of three sensors vibrate in the same manner with a similar pattern in terms of the stress amplitude.

The hot spot stress time histories from the data calculations from sensors MS3-13, MS3-14 and MS3-15 for the three train trips (i.e., 1st, 2nd and 3rd trips) are shown in Figure 4.5. It is observed from these hot spot stress time histories: (1) large and repeated stress vibrations are generated when the train runs at full-speed (around 200 km/h), as shown in Figure 4.6(a); (2) the low-amplitude stress variations with amplitudes lower than 4 MPa, are generated when the train is at railway stations in standby mode (0 km/h), as shown in Figure 4.6(b); (3) the stress responses with the pattern of Figure 4.6(c) are generated when the train runs along curves, there is a sudden increase in the level of stress amplitudes followed by a sudden decrease after approximately 80 seconds.

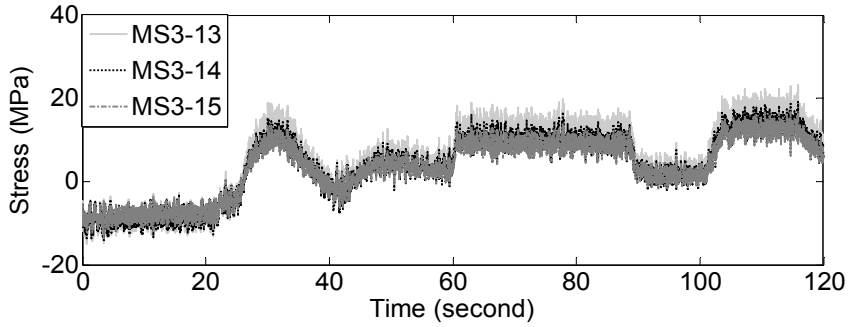
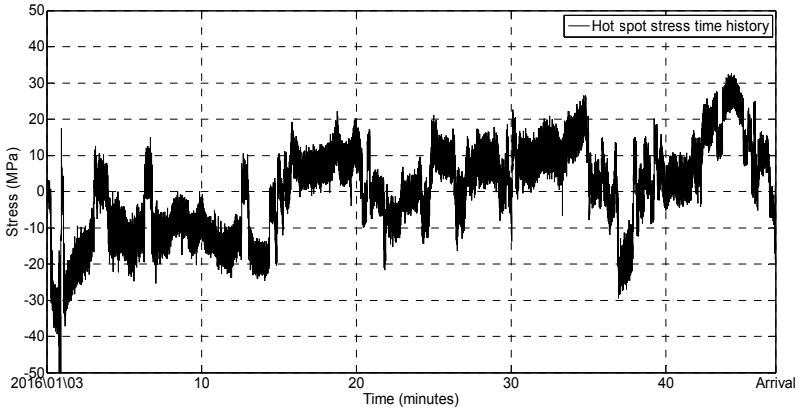
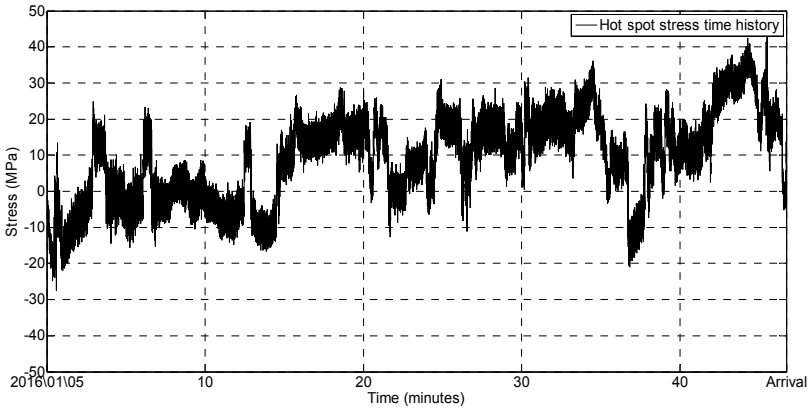


Figure 4.4 Stress time histories of three strain sensors at 7th measurement point

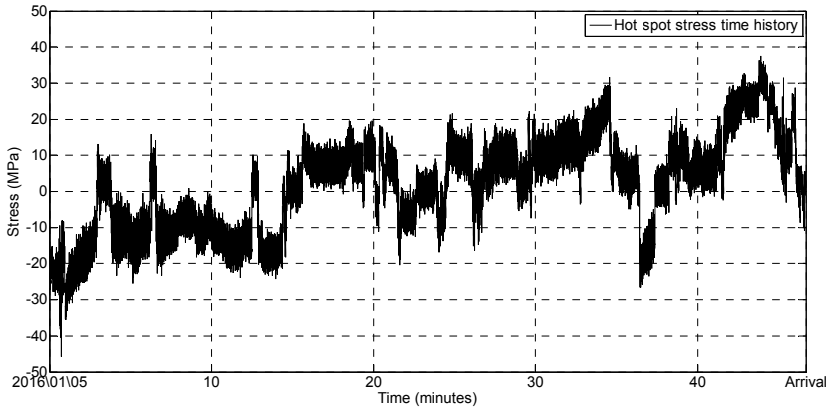
Figures 4.6(a) and (c) are some examples of the stress responses acquired during normal travel of the high-speed train bogie. These stress responses are the main causes of fatigue in bogie components, whereas the stress time histories recorded when the train sits in standby mode as shown in Figure 4.6(b), are of no significance. To assess the train fatigue life during normal high-speed railway service conditions, the stress time histories are selected for those periods when the train runs between stations, denoting as “normal service stress time history”. The normal service hot spot stress time histories obtained during the train trips are then used for further derivation of stress ranges, mean stresses, Q values, and exploration of the influence of mean stresses on fatigue.



(a)



(b)



(c)

Figure 4.5 Hot spot stress time histories at 7th measurement point during three train trips: (a) 1st trip, (b) 2nd trip, and (c) 3rd trip

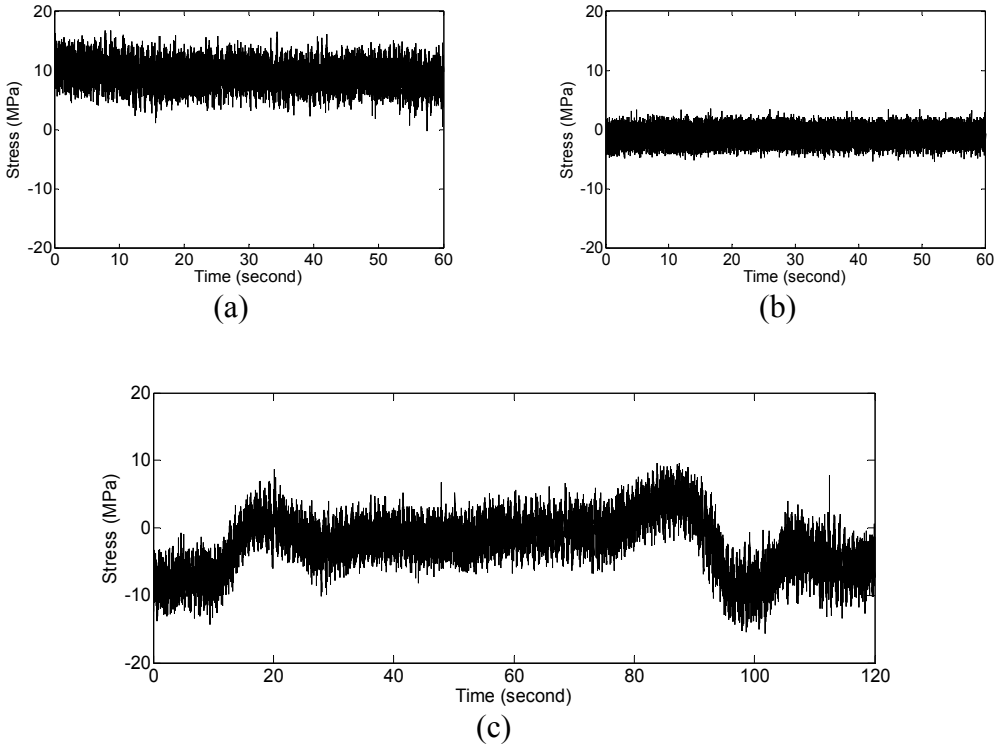


Figure 4.6 Hot spot stress time histories at 7th measurement point of train (a) during normal operation, (b) at station in standby mode (0 km/h), and (c) when runs in curves

4.3.3 PDF estimation of stress range and mean stress distributions

The hot spot stress time histories for the three trips are processed by rainflow counting technique, from which the stress ranges and mean stresses and the associated number of cycles are derived. The cycles of stress ranges with amplitudes lower than 4.0 MPa are filtered out because most of these stress cycles are caused by noise. To obtain the distribution of those mean stresses derived from rainflow counting technique, Equation (4.18) introduced in Section 4.2.2 to obtain Q values is used. The number of observations (n) of stress cycles, in the hot spot stress time history of three trips are

98184, 98784 and 95374. According to the Sturges rule (Fraley and Raftery 1998), the number of stress spectrum classes and the distribution of mean stresses (and associated Q) is k=18:

$$k = \lceil 1 + \log_2 n \rceil \tag{4.20}$$

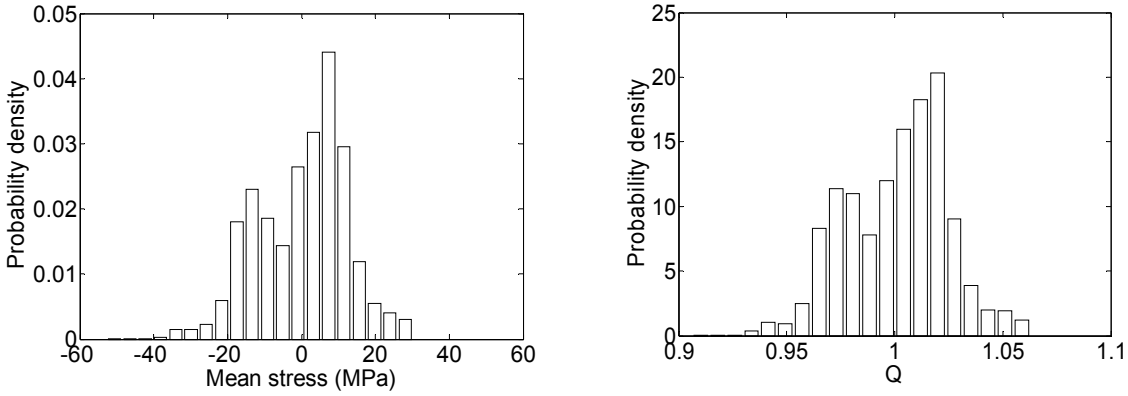
where $\lceil \rceil$ denotes the ceiling operator. The values of Q are obtained, and the distribution of in-service mean stresses and Q for the stress time histories of the three train trips are shown in Figure 4.7.

Figure 4.7 shows that values of mean stresses exist between -40 and 40 MPa, in both tensile (e.g., positive part) and compressive (negative part) forms. Distributions of the mean stresses of the hot spot stress time histories from the three train trips have a similar pattern. Figure 4.7 also reveals that the values of Q mainly accumulate within the range between 0.95 and 1.05 with a multi-modal distribution.

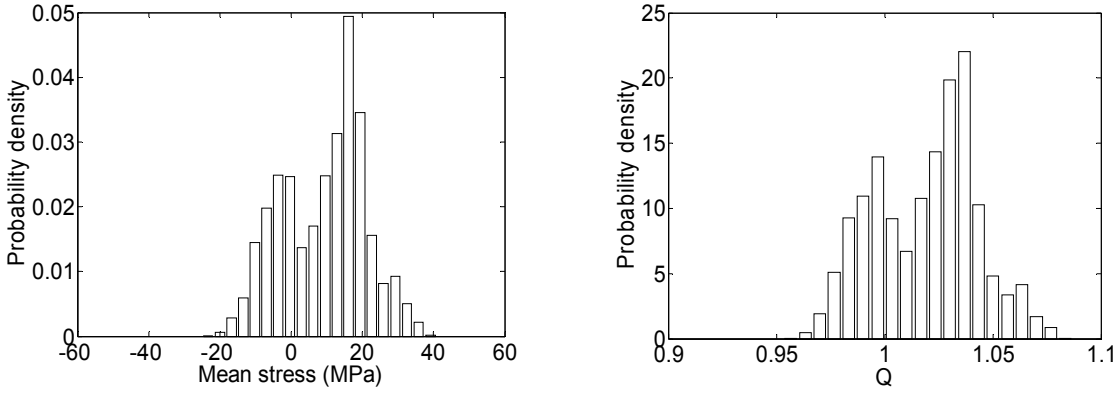
Multiple peaks can be found in the PDF of Q. This characteristic of Q makes it impossible to model the PDF of Q using a single distribution (e.g., Gaussian distribution). In view of this problem, an alternative solution is sought by using the finite mixture distributions technique.

To derive the best presentation of the PDF for Q, Gaussian, lognormal and Weibull mixture distributions expressed in Equations (3.7) to (3.9) are employed to model the distributions of Q. The EM algorithm is employed to identify the unknown parameters, and the AIC and BIC rules are in use to determine the most suitable number of

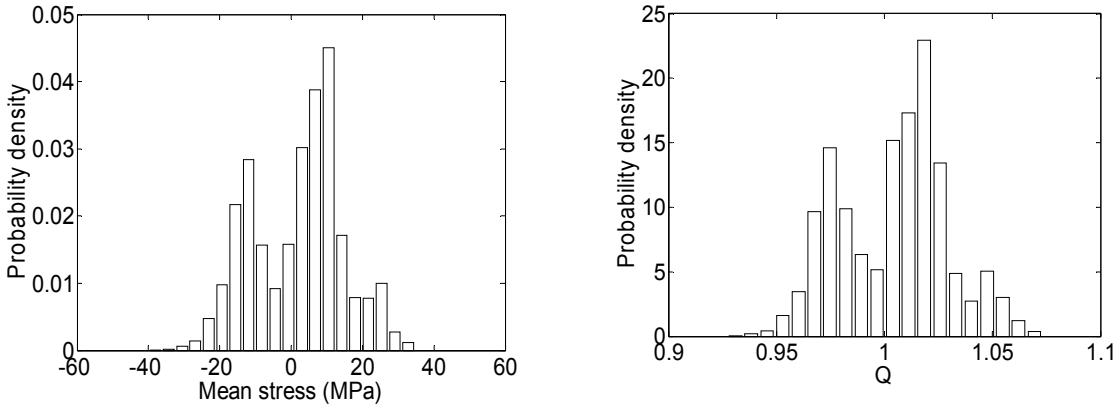
components for the mixture distributions model. In order to find the best distribution model for Q, the Q values derived from the hot spot stress time history of the 1st train trip are used for operating the modeling technique.



(a)



(b)



(c)

Figure 4.7 Distributions of mean stresses and associated *Q*: (a) 1st trip, (b) 2nd trip, and (c) 3rd trip

The three mixture distributions separately use different number of components from one to 10 for modeling the distribution of Q. Figures 4.8-10 illustrate the modeling results with two, five, seven and 10 components when using Gaussian, lognormal and Weibull mixture distributions respectively. It is seen that, with the increased number of components, the multimodality characteristic of the PDF of Q can be described more precisely by all three mixture distributions, although the PDF might be "overfitted" by the Gaussian and lognormal mixture distributions with excessive number of components.

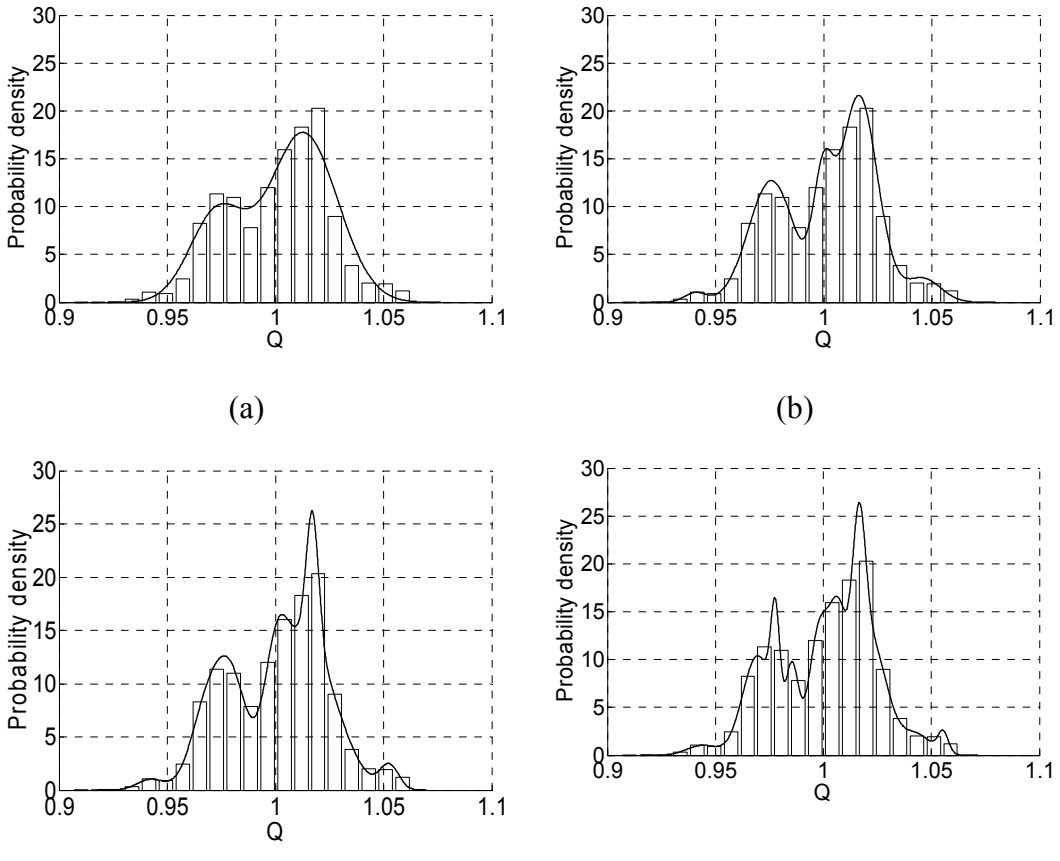


Figure 4.8 Modeling of PDF of Q using Gaussian mixture distribution with: (a) two components, (b) five components, (c) seven components, and (d) 10 components

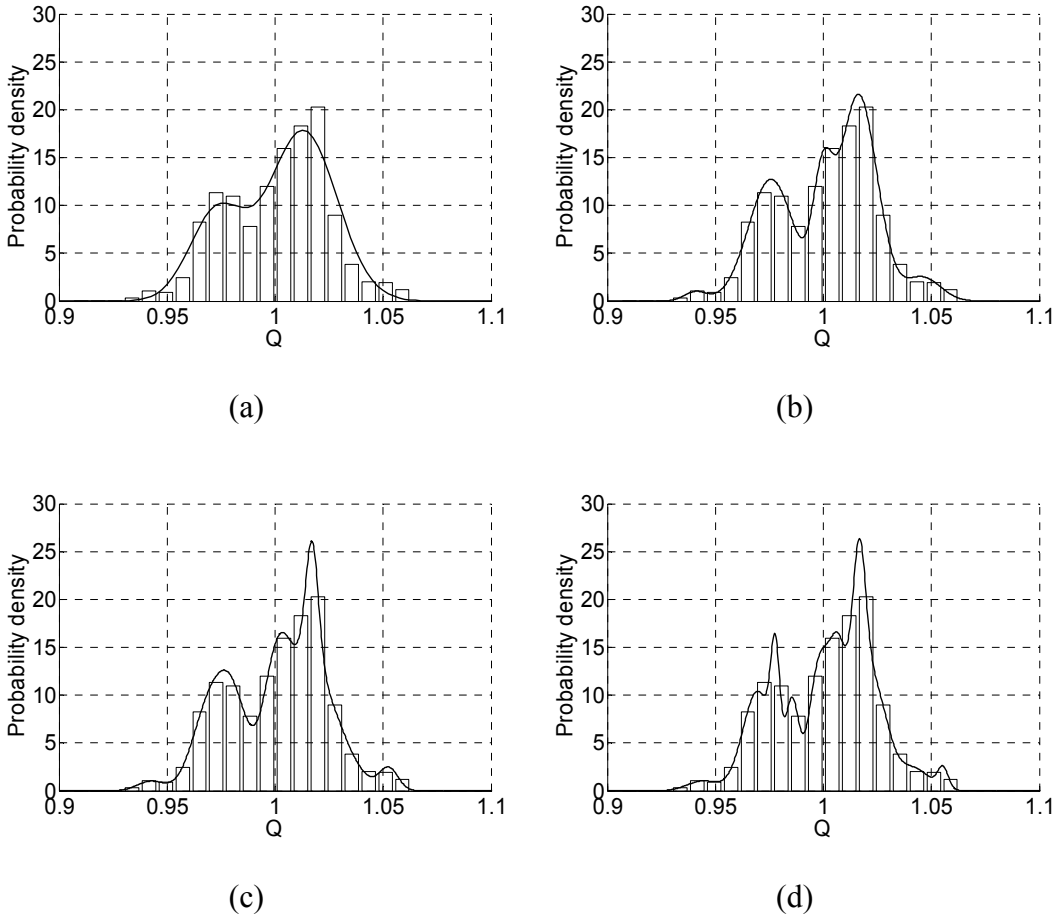
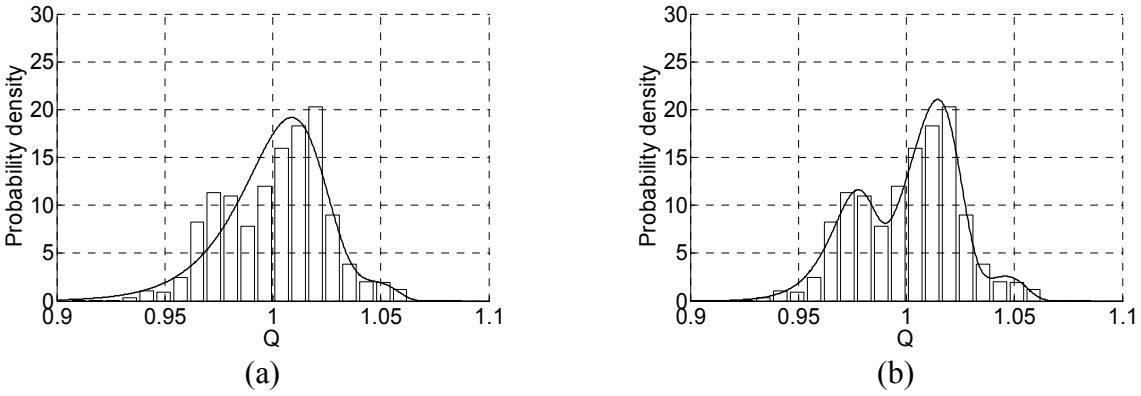


Figure 4.9 Modeling of PDF of Q using lognormal mixture distribution with: (a) two components, (b) five components, (c) seven components, and (d) 10 components



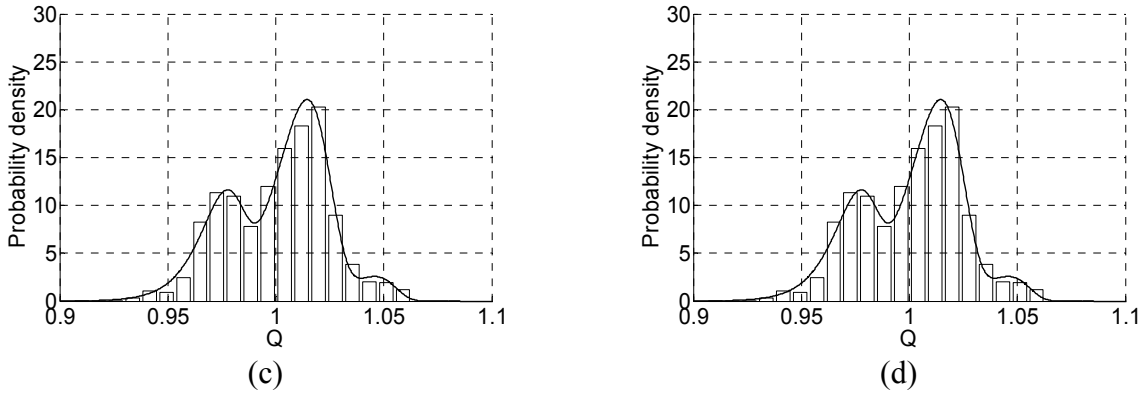


Figure 4.10 Modeling of PDF of Q using Weibull mixture distribution with: (a) two components, (b) five components, (c) seven components, and (d) 10 components

Figure 4.11 shows the results of AIC and BIC values when using different mixture distribution models with different number of parameters. It is seen that AIC and BIC values decrease and converge rapidly when the numbers of components used in the three mixture distribution models are increasing. Based on the selection rule for the best model, the lowest AIC or BIC values should apply, but from a practical perspective, Fraley and Raftery (1998) suggested selecting the number of components for best fitted PDF at the AIC or BIC values seem to converge. Thus, in this study the optimal number of components for the best model is selected at the AIC or BIC values not improving significantly.

Based on the above considerations, the Gaussian mixture distribution with six components is selected for modeling Q. Similarly, an investigation into the best modeling technique for the stress range distribution is already conducted and

concluded that the application of lognormal mixture distribution with two components is best for the stress spectra of the selected stress histories used in this chapter.

Figures 4.12 and 4.13 demonstrate the respective modeling results for both the PDF of Q and stress ranges for the hot spot stress time histories of the three trips. The calculated component parameters for the PDF of both stress spectrum and Q are shown in Tables 4.2 and 4.3, respectively. It is seen that the distribution of Q and the stress spectrum are adequately represented by mixture PDF. The information beyond the scope of the Q values can be extracted from the corresponding mixture PDF. The fatigue life assessment based on the PDFs for the Q values and for the stress ranges can be conducted.

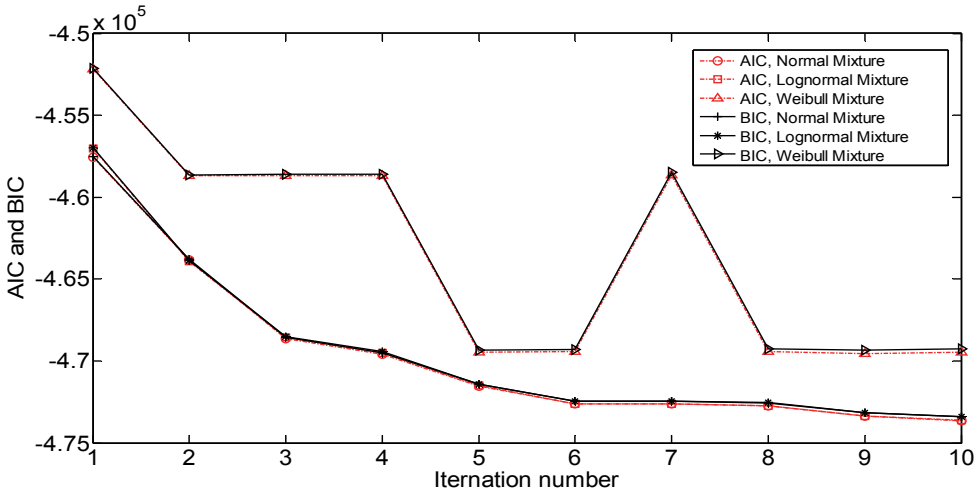


Figure 4.11 AIC and BIC values

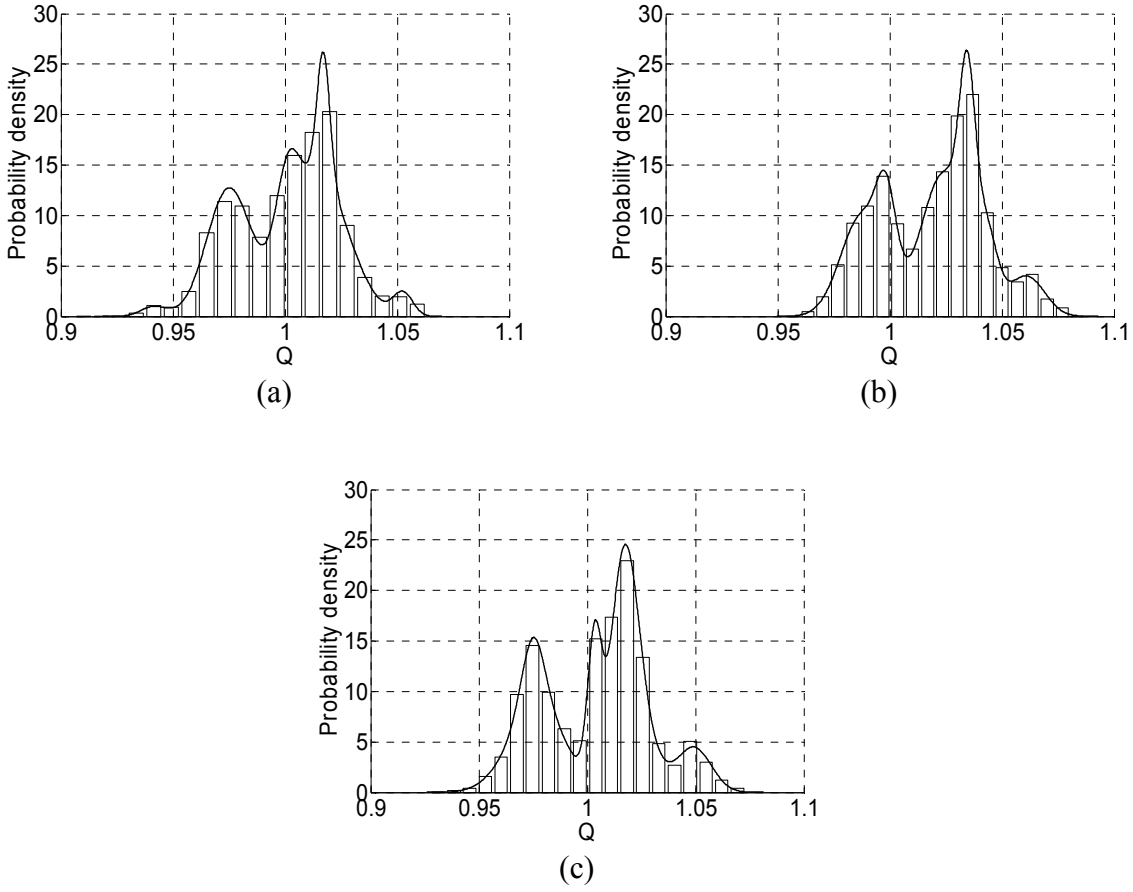
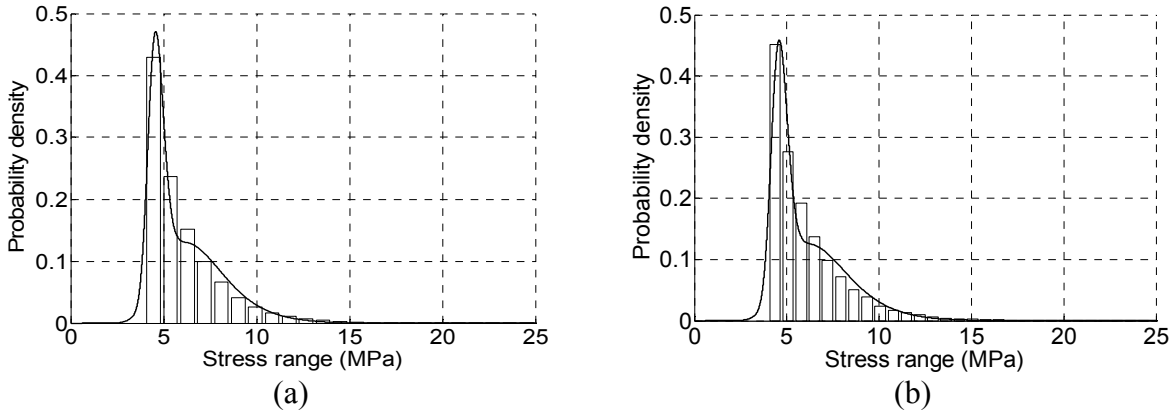
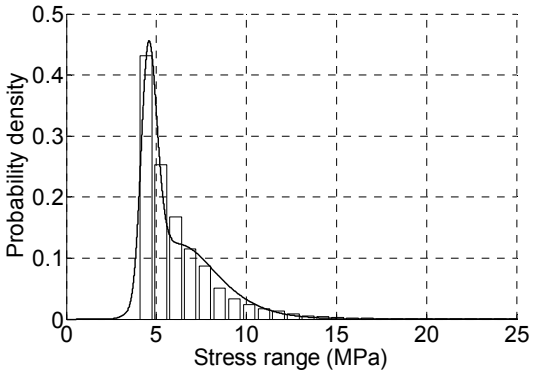


Figure 4.12 Modeling of PDF of Q using Gaussian mixture distribution with six components for (a) 1st trip, (b) 2nd trip, and (c) 3rd trip





(c)

Figure 4.13 Modeling of PDF of stress spectra using lognormal mixture distribution with two components for (a) 1st trip, (b) 2nd trip, and (c) 3rd trip

4.3.4 Fatigue life assessment results

The documentation of IIW-2008 has been widely used as the standard for the fatigue life evaluation of welded joints in train bogie components (Baek et al. 2008; Esderts et al. 2012). A study made by Li et al. (2007) concluded that the welded joint fatigue life assessment results for a rail bogie component using the principle stipulated in the IIW documentation were close to the results obtained by the fatigue tests of the same bogie component. The type of welded joint at the 7th measurement point is a fillet joint near the circular hollow section categorized as fatigue class FAT 90 (fatigue strength reference value in N/mm^2), which gives the respective values of the slopes of *S-N* curve above and below the knee point (number of cycles = 10^7) as $m_1 = 3.0$ and $m_2 = 5.0$.

Table 4.2 Estimated parameters of component distributions using Gaussian mixture distribution for modeling Q

Stress data	Values of parameters		
	Weight (w)	Mean value (μ)	Standard deviation (σ)
1st trip	0.0975	1.0010	0.0051
	0.3117	0.9749	0.0099
	0.4624	1.0147	0.0132
	0.0897	1.0170	0.0029
	0.0246	1.0526	0.0043
	0.0140	0.9415	0.0057
2nd trip	0.3639	1.0240	0.0102
	0.1423	1.0345	0.0035
	0.2370	0.9875	0.0095
	0.0801	1.0609	0.0080
	0.0836	1.0429	0.0049
	0.0930	0.9984	0.0044
3rd trip	0.0503	1.0287	0.0069
	0.0926	0.9750	0.0053
	0.0900	1.0492	0.0080
	0.3881	1.0174	0.0065
	0.1090	1.0033	0.0032
	0.2700	0.9768	0.0128

Table 4.3 Estimated parameters of component distributions using lognormal mixture distribution for stress spectra

Stress data	Values of parameters		
	Weight (w)	Mean value (μ)	Standard deviation (σ)
1st trip	0.5764	1.9064	0.2743
	0.4236	1.5237	0.0916
2nd trip	0.5685	1.9170	0.2792
	0.4315	1.5290	0.0953
3rd trip	0.4296	1.5278	0.0954
	0.5704	1.9219	0.2840

The fatigue ratio (D) for the normal service stress time history of each trip as determined by the PDFs of stress spectrum and Q, can then be obtained from Equation (4.13). The representative fatigue ratio (D_m) and fatigue life (F) for the bogie component during a train trip is calculated as the average value of the fatigue ratios obtained during the different trips, using Equation (4.16). As mentioned in Subsection 1.1.1, a typical duty of a high-speed train is 500,000 km/year. It is also known that the demanded operating period is 30 years (Grubisic and Fischer 2012). Thus, the design fatigue life in terms of mileage for a typical high-speed train component is $L_s = 1.5\times10^7$ km. The distance traveled for the normal service stress time history is H = 135 km. The fatigue ratio at failure, D_f, is taken as $D_f = 0.1$ for welded bogie components.

By applying the method proposed in this study, the fatigue life of the welded joint for the bogie is obtained as shown in Table 4.4. It is seen that the fatigue ratio predicted

for the whole service life (i.e., when the mileage reaches 1.5×10^7 km) does not reach the upper limit of the service life. The life prediction results using D_f =0.1 with and without taking into consideration mean stress effect are around 136 and 137 years, which are much longer than the demanded operating years (i.e., 30 years). This indicates the railway component is safe in terms of fatigue resistance. The results calculated with the method introduced in this chapter (with Q, the mean stress effect) and with the standard method (which does not consider the mean stress effect) are compared and shown in Table 4.4.

Table 4.4 Fatigue ratio and predicted fatigue life

Connection of tubular crosspiece and frame	Fatigue ratio at design traveling distance	Life prediction using D_f = 0.1	Fatigue life (number of operating years)
With mean stress	2.210E-03	6.787E+07 km	135.740
Without mean stress	2.193E-07	6.842E+07 km	136.840

4.3.5 Discussion

It is seen in Table 4.4 that the fatigue life for D_f=0.1 considering mean stress, is slightly lower than the case not considering mean stress, which means that although the overall influence of the mean stress is a negative factor to fatigue life, it is small and might be

ignored. The cause of the negligible influence of mean stress may be due to the proportions of the tensile and compressive mean stresses in mean stress distribution are almost equivalent. This equal proportions characteristic makes the modeled PDF of Q has almost equal areas in negative and positive parts. According to the constant life diagram in Equation (4.17) or Q in Equation (4.18), the compressive mean stress extends or is beneficial to fatigue life, while the tensile mean stress shortens or is detrimental to fatigue life. This finding is in accordance with the previous research findings (Rahman et al. 2008; Radaj and Vormwald 2013). Therefore, the contribution (in reducing fatigue life) of the positive part of the modeled PDF of Q is neutralized by the negative part of the modeled PDF of Q.

This also explains that for the hot spot stress time history of 2nd trip in which the tensile mean stress represents larger proportion of the overall mean stress distribution, the difference of the fatigue life considering mean stress (by counting PDF of Q) and without mean stress effect increases to 2% (6.787E+07 km versus 6.842E+07 km). In other words, when the mean stresses in the time histories are predominantly tensile, the fatigue life assessment approach gives more conservative results (i.e., difference of fatigue lives with and without considering mean stress becomes larger).

In addition, it is seen in Figure 4.11 that there are variations in the AIC and BIC values after the number of components is larger than seven. It would be because with the use of excessive number of components in mixture distributions to fit indicator Q, the fitted

results became unstable; it is therefore recommended to apply the mixture distribution with number of components less than seven to fit indicator Q.

4.4 Summary

In this chapter, a fatigue life assessment approach considering the variations in stress ranges and mean stresses is described for condition monitoring purpose. Firstly, a method for fatigue life assessment that is capable of taking mean stress effects into account is developed, by making use of a defined indicator Q and its associated PDF. The method is then used in relation to in-service monitoring stress time histories obtained using the sensors deployed on a high-speed train bogie for three train trips. This real application helps understand the effects of mean stresses. Distributions of mean stresses and associated probability mixture distribution models are derived respectively. Finally, the fatigue ratio and predicted fatigue life results for the measured high-speed train are calculated. The performance of the proposed fatigue life assessment method is demonstrated.

According to the research work accomplished, several points can be drawn as follows:

1. Distribution of the mean stresses exhibits the characteristic of multimodality;
2. PDF of the proposed indicator Q can be adequately modeled using a mixture distribution technique and its distribution conforms to Gaussian mixture distribution;

3. With the use of excessive number of components in mixture distributions to fit the indicator Q, the fitted results become unstable; it is recommended to apply the mixture distribution with number of components less than seven in fitting Q;
4. The influence of mean stresses on fatigue life is negligible for the case of normal train operating condition, and may significantly shorten fatigue life when tensile mean stresses are predominant in overall mean stresses, under the scope of Goodman relation for considering mean stress effect; and
5. A fatigue life assessment approach has been developed, which can extract and consider the impacts of varying stress ranges and mean stresses in the monitoring data on fatigue life, using frequentist modeling technique to consider the uncertainties in both stress ranges and mean stresses.

Overall, the proposed frequentist fatigue life assessment approach is capable of fatigue life assessment for high-speed trains instrumented with condition monitoring system.

CHAPTER 5

BAYESIAN MODELING AND UPDATING OF STRESS SPECTRUM USING PRIOR KNOWLEDGE AND CONTINUOUS MONITORING DATA

5.1 Introduction

As mentioned in Chapter 1, the decision-maker for maintenance prefers the fatigue life result to be assessed through integrating both the experience and observation. In short-term period of on-board vehicle monitoring, the priority is to have efficient solution to obtain monitoring-based fatigue assessment result. Due to the limitation of monitoring duration, only the monitoring data is the reliable source for quantifying the performance of measured train components, and therefore usually used to conduct fatigue analysis, as illustrated in Chapters 3 and 4. A frequentist approach can satisfy these needs to calculate results efficiently with monitoring data. However, in reality, train components are operated in the presence of external influence which is preferred to be included in the assessment. It is possible to obtain the information about external influence (i.e., prior information) when on-board monitoring periods become longer. In addition, continuous monitoring data is obtained from the on-board monitoring

system. When such information about the monitoring data is collected, it is worth to quantify and include this information in analysis results.

Hence, in long-term period of on-board vehicle monitoring, probabilistic fatigue life assessment should consider more complicated information: (1) the prior information should be integrated together with the monitoring data to jointly infer fatigue analysis results; and (2) it is also desired to predict the fatigue life result deriving from the continuous monitoring data.

However, classical frequentist approaches do not integrate prior information with current observations for assessment purposes, and the solution for prediction has not been well established. As mentioned in Subsection 1.2, in addition to frequentist inference, there also exists Bayesian inference to interpret probability. With recent advances in computing techniques, Bayesian methodology is increasingly applied for solving different types of complex problems in engineering field (Struss et al. 2008; Babuška 2016; Wan and Ren 2016; Wan and Ni 2017; Wang et al. 2017). In particular, Bayesian modeling technique is now a preferable way for establishing a probabilistic model for fatigue life assessment (Stone 2013; Chiachío et al. 2014). Bayesian methods take account of new external data as well as the monitoring observations in making the final inference (Freeman 1996). This characteristic ensures that Bayesian modeling and updating methods are able to predict results incorporating prior knowledge and continuous monitoring data.

Bayesian methods ensure the modeling and prediction results to be feasible by following Bayes' rule to conduct inductive learning. The associated result of combining the prior information and observed data in a Bayesian inference is the posterior distribution. It has an advantage that the posterior is proportional to the prior times likelihood (function for observed data). If the prior information is trivial, it has negligible weight in the posterior, which is then only dominated by the data information (Stone 2013). Similarly, as more monitoring data is collected, the influence of the prior information decreases, and again the inference results (i.e. posterior) effectively rely on the information in the data. Once the data amount is big enough to support an objective conclusion, there is essentially no room left for subjective opinion (i.e., prior information, O'Hagan 2004).

The research work in this Chapter addresses the issue of stress spectrum modeling through Bayesian modeling and updating methods for utilizing the characteristics of measurement data. The modeling and updating methods satisfy the needs for long-term period of on-board vehicle monitoring. Previously, a number of studies developed several Bayesian models for civil engineering structures. These models considered the influence of the uncertainties in fatigue resistance data for fatigue analysis (Madsen 1984; Amiri et al. 2013; Babuška et al. 2016). However, research studies on modeling the in-service stress spectrum are limited. Parameters in the estimated stress spectrum models are regarded as random variables. This assumption reflects the uncertainties in

varying observational characteristics. In addition, the Bayesian updating technique is also introduced to predict the distribution of stress spectrum.

In this chapter, the Bayesian modeling and updating solutions are verified using the in-service monitoring data obtained from a high-speed train. Stress ranges acquired from a measurement point during different train trips are used to demonstrate the performance of the proposed methods. PDFs are obtained from the modeling and prediction. The PDF predictions are updated continuously from the stress ranges of different train trips. The predicted results are useful for further exploration on constructing fatigue life assessment and prediction approaches.

5.2 Stress Spectrum Modeling Using Bayesian Inference

5.2.1 Theoretical basis

Assume the information of the stress ranges s derived according to the observations of monitoring system is shown as: $(s|\Theta)$, where Θ is comprised of the parameters which to determine the distribution of s. In addition, the information about stress ranges, other than that acquired from monitoring system, can be regarded as a "prior information"; in the algorithm, a prior for Θ is elicited and denoted by (Θ). Then the posterior density for Θ will be denoted by $(\Theta|s)$, which means the posterior density is conditional on the data. It has a relationship of:

$$P(\Theta \mid s) \propto P(\Theta) \times P(s \mid \Theta) \tag{5.1}$$

where $P(\Theta)$ is the prior knowledge about the parameter Θ, and $P(s|\Theta)$ denotes the likelihood of observation dataset. With the information, the posterior distributions of Θ given the observed stress ranges s are derived. In the following, this derivation principle is used to derive the posterior distributions of s.

5.2.2 Likelihood function

For the observations from monitoring system, the derived stress spectrum s can be described by the following equations:

$$s = \mu + \varepsilon \tag{5.2}$$

or

$$s_i = \mu_i + \varepsilon_i \tag{5.3}$$

where μ is true response of stress range and ε is noise. The observation of stress range s_i at ith moment is the results of μ_i and ε_i at ith moment. The only known information about the above equations is the distribution of noise, which follows Gaussian distribution:

$$\varepsilon = s - \mu \sim N(0, \sigma^2 I) \tag{5.4}$$

which is simplified as:

$$s \sim N(\mu, \sigma^2 I) \tag{5.5}$$

It is seen that the observation dataset of stress ranges s follows Gaussian distribution $\sim N(\mu, \sigma^2)$. Then the density form of s given μ, σ^2 is:

$$P(s \mid \mu, \sigma^2) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left[-\frac{1}{2\sigma^2} \left(s_i - \mu \right)^2 \right] \tag{5.6}$$

Assume $\bar{s}$ is the sufficient statistics of μ, then $\bar{s} = \frac{1}{n} \times \sum_1^n s_i$ and Equation (5.6) is simplified to:

$$P(s \mid \mu, \sigma^2) \propto \frac{1}{(\sigma^2)^{n/2}} \exp\left\{ -\frac{1}{2\sigma^2} \left[(n-1)g^2 + n(\bar{s} - \mu)^2 \right] \right\} \tag{5.7}$$

where $g^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(s_i - \bar{s} \right)^2$.

5.2.3 Prior information

The prior information in a Bayesian algorithm reflects the previous knowledge of the phenomenon before observing the data. Though there are numerous forms of priors, it is common to use the specific ones that are easy for calculation. Non-informative priors have advantages. Non-informative priors, combining with the likelihood, can yield a posterior that keeps the same type of distribution with the prior. More specifically, when no information before obtaining observation is provided for the parameters, a non-informative prior which follows Jeffrey's prior is used (Jeffrey 1946):

$$P(\mu, \sigma^2) \propto \frac{1}{\sigma^2} \tag{5.8}$$

5.2.4 Posterior distribution

5.2.4.1 Joint posterior distribution of μ and σ^2

According to the aforementioned analysis, the posterior distribution of the parameters can be given as:

$$P(\mu,\sigma^2 \mid s) \propto P(\mu,\sigma^2)\times P(s \mid \mu,\sigma^2)$$
$$\propto \frac{1}{\sigma^2}\frac{1}{(\sigma^2)^{n/2}}\exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)g^2+n(\bar{s}-\mu)^2\right]\right\} \tag{5.9}$$

5.2.4.2 Marginal posterior distribution of σ^2

The target of this study is to obtain the marginal posterior distribution of the parameters μ and σ^2. For σ^2, this can be done by integrating μ out of the above algorithm:

$$P(\sigma^2 \mid s) \propto \int_{-\infty}^{+\infty}\frac{1}{(\sigma^2)^{1+n/2}}\exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)g^2+n(\bar{s}-\mu)^2\right]\right\}d\mu$$
$$\propto \frac{1}{(\sigma^2)^{(n+1)/2}}\exp\left[-\frac{(n-1)g^2}{2\sigma^2}\right]\frac{\sqrt{2\pi}}{\sqrt{n}} \tag{5.10}$$
$$\times\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-(\frac{\mu}{\sqrt{n}}-\frac{\bar{s}}{\sqrt{n}})^2\frac{1}{2\sigma^2}\right]d\frac{\mu}{\sqrt{n}}$$

As is known that $F(x)=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx=1$, the above algorithm is simplified as:

$$P(\sigma^2 \mid s) \propto \frac{\left[\frac{(n-1)}{2} g^2\right]^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \frac{1}{(\sigma^2)^{\frac{n+1}{2}}} \exp\left\{\frac{1}{\sigma^2}\left[-\frac{(n-1)g^2}{2}\right]\right\} \tag{5.11}$$

The Posterior distribution of σ is obtained from dataset s, following Inverse-gamma distribution:

$$P(\sigma^2 \mid s) \propto IG(\frac{n-1}{2}, \frac{(n-1)\times g^2}{2}) \tag{5.12}$$

where $g^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(s_i - \bar{s}\right)^2$.

5.2.4.3 Marginal posterior distribution of μ

On the other side, the conditional posterior distribution of μ given σ^2 and s is:

$$P(\mu \mid \sigma^2, s) \sim N(\bar{s}, \frac{\sigma^2}{n}) \tag{5.13}$$

The joint distribution of μ and σ^2 can be expressed as:

$$\begin{aligned} P(\mu, \sigma^2 \mid s) &\propto P(\mu \mid \sigma^2, s) \times P(\sigma^2 \mid s) \\ &\propto N(\bar{s}, \frac{\sigma^2}{n}) \times IG(\frac{n-1}{2}, \frac{(n-1)g^2}{2}) \end{aligned} \tag{5.14}$$

5.2.4.4 Marginal posterior distribution of μ

For μ, it can be done by integrating σ^2 out of the above algorithm:

$$P(\mu \mid s) = \int_0^{+\infty} P(\mu, \sigma^2 \mid s) d\sigma^2 \tag{5.15}$$

$$= \frac{\left[\frac{(n-1)}{2} g^2\right]^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \frac{\sqrt{n}}{\sqrt{2\pi}} \times \int_0^{+\infty} \frac{1}{(\sigma^2)^{\frac{n+1}{2}}} \exp\left[-\frac{(n-1)g^2 + n(\mu - \bar{s})^2}{2\sigma^2}\right] d\sigma^2$$

It is noticed that a Gamma function is expressed as:

$$\frac{\Gamma(a)}{\lambda} = \int_0^{+\infty} x^{a-1} e^{-\frac{\lambda}{x}} dx \text{ and } \Gamma(a+1) = a \times \Gamma(a) \tag{5.16}$$

Therefore, the right part of Equation (5.15) can be simplified when assuming $x = \sigma^2$ and $\lambda = -\frac{(n-1)g^2 + n(\mu - \bar{s})^2}{2\sigma^2}$ in Gamma function. The whole equation can be integrated to:

$$\begin{aligned} P(\mu \mid s) &= \frac{\left[\frac{(n-1)}{2} g^2\right]^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \frac{\sqrt{n}}{\sqrt{2\pi}} \frac{\Gamma\left(\frac{n}{2}\right)}{\left[-\frac{(n-1)g^2 + n(\mu - \bar{s})^2}{2}\right]^{\frac{n}{2}}} \\ &= \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\sqrt{(n-1)\pi}\frac{g}{\sqrt{n}}\left[1 + \frac{1}{n-1}\left(\frac{\mu - \bar{s}}{g/\sqrt{n}}\right)^2\right]^{\frac{n}{2}}} \end{aligned} \tag{5.17}$$

It is found that the algorithm satisfies the structure of the non-standardized *Student-t* distribution, with parameters with μ_t as the mean: $\bar{s}$; σ_t as scale parameter: $(\frac{g}{\sqrt{n}})^2$ and v as degrees of freedom: $n-1$. The marginal posterior distribution of μ is obtained from dataset s, following *Student-t* distribution.

*5.2.4.5 Predicted distribution of future observation s**

The analytical expressions for the marginal posterior distributions of μ and σ^2 are obtained. Finally, the posterior density of future observations of stress ranges s^*, can be obtained by integrating with respect to μ *and* σ^2 (Webb 2006):

$$P(s^* \mid s_j) = \int\int P(s^* \mid \mu, \sigma^2) \times P_j(\mu, \sigma^2 \mid s_j) d\mu, \sigma^2 \tag{5.18}$$

which turns out to be a multivariate *Student-t* distribution expressed as:

$$P(s^* \mid s) = St_d(\bar{s}, \frac{n}{g^2(n+1)}, n-1) \tag{5.19}$$

where $g^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(s_i - \bar{s}\right)^2$.

5.3 Model Updating with Continuous Monitoring Data

5.3.1 Joint and marginal posterior distributions

When the continuous monitoring data is introduced to the model, the model established from Section 5.2 based on the initial monitoring data can be updated according to Bayesian updating algorithm. Below the analytical solution for the updating technique is introduced.

When new observations arrive, the prior information of the parameters relating to μ and σ^2 is the joint and marginal posterior distributions of μ and σ^2 obtained in Section

5.2. As shown in Equation (5.14), joint distribution of μ and σ^2 has a normal-inverse gamma prior distribution:

$$P(\mu,\sigma^2) = IG(a,\beta)\times N(\mu_0,\frac{1}{k}\sigma^2) \tag{5.20}$$

Therefore, the joint prior distribution of μ and σ^2 can now be generally expressed as:

$$P(\mu,\sigma^2) = \frac{\beta^a}{\Gamma(a)}\frac{1}{(\sigma^2)^{a+1}}\exp\left(-\frac{\beta}{\sigma^2}\right)\frac{\sqrt{k}}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{k(\mu-\mu_0)^2}{2\sigma^2}\right] \tag{5.21}$$

where $\alpha=\alpha_0=\frac{n-1}{2}$, $\beta=\beta_0=\frac{(n-1)g^2}{2}$, $g^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(s_i-\bar{s}\right)^2$, $\mu_0=\mu_t=\bar{s}$, and $k=k_0=n$.

Now consider a continuous monitoring dataset s_2 is observed and used for updating the model:

$$P(\mu,\sigma^2 \mid s_2) \propto P(\mu,\sigma^2 \mid s)\times P(s_2 \mid \mu,\sigma^2,s) \tag{5.22}$$

The updating joint posterior distribution of the parameters μ and σ^2 is expressed as:

$$\begin{aligned} P(\mu,\sigma^2 \mid s_2) = {} & \frac{\beta^a}{\Gamma(a)}\frac{1}{(\sigma^2)^{a+1}}\exp\left(-\frac{\beta}{\sigma^2}\right)\frac{\sqrt{k}}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{k(\mu-\mu_0)^2}{2\sigma^2}\right] \\ & \times\frac{1}{(\sigma^2)^{n_2/2}}\exp\left[-\frac{(n_2-1)g^2}{2\sigma^2}\right]\frac{\sqrt{n_2}}{\sigma^2}\exp\left[-\frac{n_2(\bar{s_2}-\mu)^2}{2\sigma^2}\right] \end{aligned} \tag{5.23}$$

where s is the initial monitoring data, s_2 is the new dataset; n_2 is the number of observations/stress ranges in the new dataset; $\bar{s}_2$ is the mean value of the dataset of s_2. The algorithm is simplified to:

$$P(\mu,\sigma^2 \mid s_2) \propto \frac{1}{(\sigma^2)^{a+n_2/2+1}} \exp\left\{-\frac{1}{\sigma^2}\left[\beta+\frac{(n_2-1)g^2}{2}\right]\right\}\frac{\sqrt{k}}{\sigma} \times \exp\left[-\frac{k\left(\mu-\mu_0\right)^2+n_2\left(\mu-\overline{s_2}\right)^2}{2\sigma^2}\right] \tag{5.24}$$

In the equation above, the last component can be calculated as:

$$k\left(\mu-\mu_0\right)^2+n_2\left(\mu-\overline{s_2}\right)^2=(k+n_2)\left(\mu-\frac{k\mu_0+n_2\overline{s_2}}{k+n_2}\right)^2+\frac{kn_2(\overline{s_2}-\mu_0)^2}{k+n_2} \tag{5.25}$$

With the converted expression in Equation (5.25), Equation (5.24) for $P(\mu,\sigma^2 \mid s_2)$ can be further expressed as:

$$P(\mu,\sigma^2 \mid s_2) \propto \frac{1}{(\sigma^2)^{a+n_2/2+1}} \exp\left\{-\frac{1}{\sigma^2}\left[\beta+\frac{(n_2-1)g^2}{2}\right]\right\}\frac{\sqrt{k}}{\sigma} \times \exp\left[-\frac{(k+n_2)\left(\mu-\frac{k\mu_0+n_2\overline{s_2}}{k+n_2}\right)^2}{2\sigma^2}+\frac{\frac{kn_2(\overline{s_2}-\mu_0)^2}{k+n_2}}{2\sigma^2}\right] \tag{5.26}$$

It is found that the distribution is a form of joint PDF for both Gaussian and inverse-Gamma distributions:

$$P(\mu,\sigma^2 \mid s_2) \sim IG\left((a+\frac{n_2}{2}),\beta+\frac{\left(n_2-1\right)s^2}{2}+\frac{kn(\overline{s_2}-\mu_0)}{2\left(k+n_2\right)}\right) \times N\left(\frac{k\mu_0+n_2\overline{y}}{k+n_2},\frac{\sigma^2}{k+n_2}\right) \tag{5.27}$$

which can be simplified as:

$$P(\mu,\sigma^2 \mid s_j) \sim IG\left(a_j,\beta_j\right)\times N\left(\mu_j,\frac{\sigma^2}{k_j}\right) \tag{5.28}$$

where $j=2$, $a_j=a_{j-1}+\frac{n_j}{2}$

$$\beta_j = \beta_{j-1} + \frac{(n_j - 1)g_j^2}{2} + \frac{k_{j-1} n_j (s_j - \mu_{j-1})^2}{2(k_{j-1} + n_j)}$$

$$k_j = k_{j-1} + n_j$$

$$\mu_j = \frac{k_{j-1}\mu_{j-1} + n_j \overline{s_j}}{k_{j-1} + n_j}$$

$$\overline{s_i} = \frac{1}{n_j} \sum_{i=1}^{n_j} s_{j,i}$$

$$g^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \left(s_{j,i} - \overline{s_j} \right)^2$$

With this joint PDF shown in Equation (5.28), the posterior distributions of the parameters μ (none-standardized student-t distribution) and σ^2 (Inverse-gamma distribution) can be obtained, which are:

$$P(\sigma_j^2) = IG(a_j, \beta_j) \tag{5.29}$$

and

$$P(\mu_j) = St_d(\mu_j | \mathrm{v}_j, \mu_{t,j}, \sigma_{t,j}) \tag{5.30}$$

where $\mu_{t,j} = \overline{y}_i$

$$\sigma_{t,j} = \frac{1}{k_{j-1} + n_j}$$

$$\mathrm{v}_j = \mathrm{v}_{j-1} + n_j$$

5.3.2 Predicted PDF

Based on the same principle introduced in the Section 5.3.1 for updating model for a stress spectrum, when M ($j=1, 2, 3, \ldots, W$) sets of dataset s_j are observed, the predicted final PDF of stress spectrum updated from Bayesian technique is expressed as:

$$P(s^* \mid s_j) = \int\int P(s^* \mid \mu, \sigma^2) \times P_j(\mu, \sigma^2 \mid s_j) d\mu, \sigma^2 \tag{5.31}$$

where

$$P(\sigma_j^2) = IG(a_j, \beta_j) \quad \text{and} \quad P(\mu_j) = \mathrm{St}_d(\mu_j | \mathrm{v}_j, \mu_{t,j}, \sigma_{t,j}) \tag{5.32}$$

where $a_1 = \frac{n_1 - 1}{2}$, $\beta_1 = \frac{(n_1 - 1)g_1^2}{2}$, $\mu_1 = \bar{s}_1$, $k_1 = n_1$, $a_j = a_{j-1} + \frac{n_j}{2}$,

$k_j = k_{j-1} + n_j$, $\beta_j = \beta_{j-1} + \frac{(n_j - 1)g_j^2}{2} + \frac{k_{j-1} n_j (s_j - \mu_{j-1})^2}{2(k_{j-1} + n_j)}$, $\mu_j = \frac{k_{j-1}\mu_{j-1} + n_j \bar{s}_j}{k_{j-1} + n_j}$,

$\sigma_{t,j} = \frac{1}{k_{j-1} + n_j}$, $\mu_{t,j} = \bar{s}_i$, and $\mathrm{v}_j = \mathrm{v}_{j-1} + n_j$.

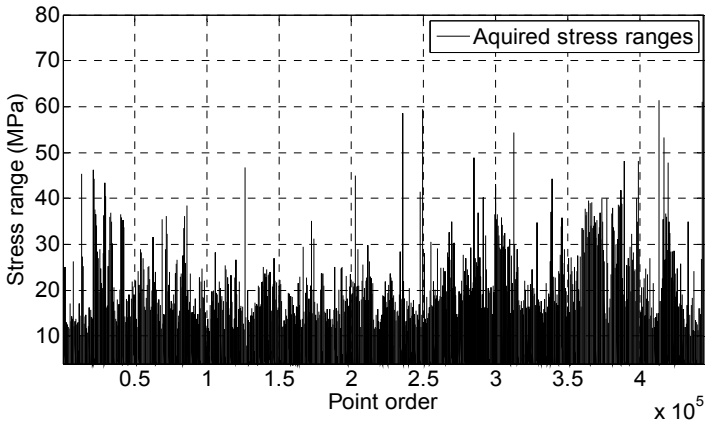
5.4 Demonstration Using In-service Monitoring Data

5.4.1 Stress ranges of in-service monitoring data

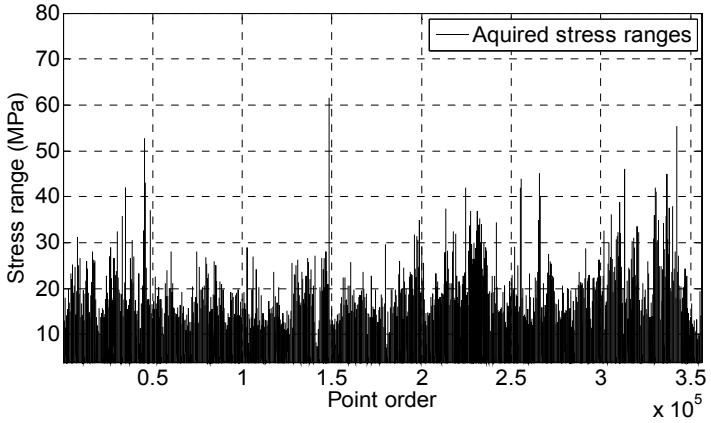
Bayesian modeling and updating of the stress spectra are performed for a welded joint on a train running on a Chinese high-speed railway using monitoring data. The stress ranges of six train trips are used to verify the established Bayesian model. In the stress ranges, the values mainly caused by noise are lower than 2 MPa. They have been eliminated from the original data.

Below, the stress ranges of six train trips (i.e., 1st, 2nd, 3rd, 4th, 5th and 6th train trips) are called "1st dataset", "2nd dataset", "3rd dataset", "4th dataset", "5th dataset" and "6th dataset". Examples of such stress ranges histories are shown in Figure 5.1. It is seen that the majority of stress ranges for the six train trips are below 40 MPa. The average amplitudes of 1st, 2nd, 3rd, 4th, 5th and 6th datasets are 3.84 MPa, 3.91 MPa, 3.88 MPa,

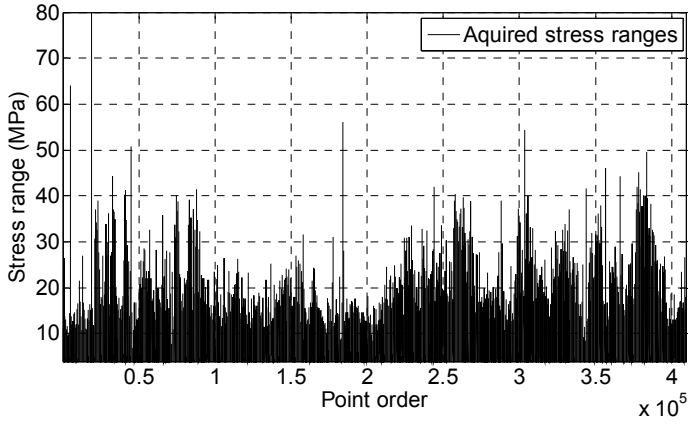
3.91 MPa, 3.88 MPa and 3.84 MPa, respectively. The largest stress ranges of 1st, 2nd, 3rd, 4th, 5th and 6th datasets are 79.27 MPa, 76.48 MPa, 77.41 MPa, 78.67 MPa, 77.26 MPa and 67.15 MPa, respectively. This shows that the range amplitudes obtained for these six train trips remain stable. They are suitable for verification of the proposed method on conducting modeling and continuous updating the fitted PDF.


Aquired stress ranges
Stress range (MPa)
Point order
x 10^5

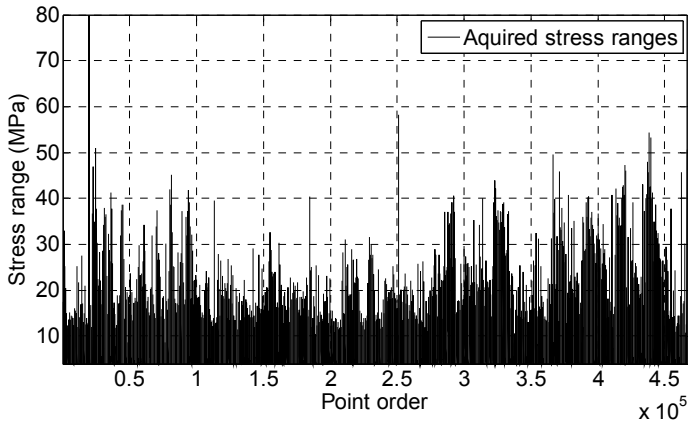
(a)


Aquired stress ranges
Stress range (MPa)
Point order
x 10^5

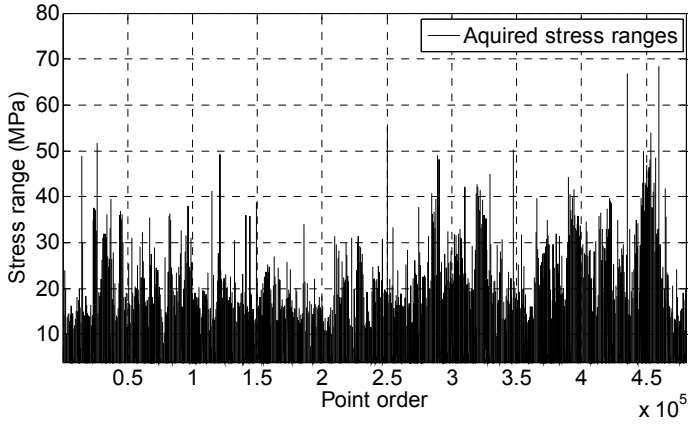
(b)


Aquired stress ranges
Stress range (MPa)
Point order
x 10^5

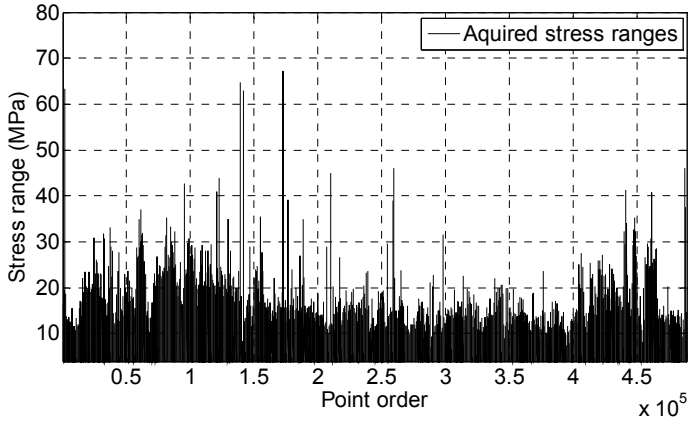
(c)



(d)



(e)



(f)

Figure 5.1 Acquired stress ranges from stress time histories of (a) 1st trip, (b) 2nd trip, (c) 3rd trip, (d) 4th trip, (e) 5th trip, and (e) 6th trip

The aforementioned six datasets are processed by the established modeling method. These stress spectra are used to investigate the performance of the Bayesian approach in estimating posterior distributions.

Firstly, to observe the performance of the modeled PDF derived by following the process introduced in Section 5.2, the Bayesian technique is performed separately on the stress ranges of each train trip.

Secondly, to verify the performance of the predicted PDF derived using the proposed Bayesian updating technique in Section 5.3, (1) the 1st dataset is processed initially by the Bayesian modeling method introduced in Sections 5.2, resulting in a derived PDF, and (2) this modeled result is then updated for prediction by Bayesian updating technique proposed in Section 5.3, which gradually adopts 2nd, 3rd, 4th and 5th, and 6th datasets to update the predicted results.

5.4.2 Modeling and updating results

The posterior distributions of the modeled stress spectra using the Bayesian modeling method are shown in Figure 5.1. The expression of the modeled stress spectrum s^* in relation to the posterior distributions of μ and σ can be simplified by integration:

$$f(s^*) = P(s^* \mid s) = St_d(\bar{s}, \frac{n}{g^2(n+1)}, n-1) \tag{5.33}$$

which can be written as:

$$f(s^*)=\frac{\tau\left(\frac{1}{2}n\right)}{\tau\left(\frac{n-1}{2}\right)(n-1)^{\frac{1}{2}}\pi^{\frac{1}{2}}}\left(\frac{n}{\frac{1}{n-1}\sum_{i=1}^{n}\left(s_i-\bar{s}\right)^2(n+1)}\right)^{\frac{1}{2}} \times\left[1+\frac{1}{n-1}\left(s^*-\bar{s}\right)^2\frac{n}{\frac{1}{n-1}\sum_{i=1}^{n}\left(s_i-\bar{s}\right)^2(n+1)}\right]^{-\frac{n}{2}} \tag{5.34}$$

The results of the posterior distributions of μ and σ as well as the statistical characteristics (e.g. mean, standard deviation, and 95% credibility intervals) of the predicted distribution of s^* are summarized in Table 5.1 and drawn in Figure 5.2. It is seen from Figure 5.2 that the modeled PDFs for the stress ranges of six train trips fit all stress spectra well. It can also be seen from Table 1 that the parameter distributions of both μ and σ for the six datasets are similar to each other. The 95% credibility intervals of parameter μ remain at [3.70, 3.86] for six datasets, and that of parameter σ remain at [1.57, 1.68] for six datasets. The 95% credibility intervals of the modeled stress range s^* using each of 6 datasets are also similar.

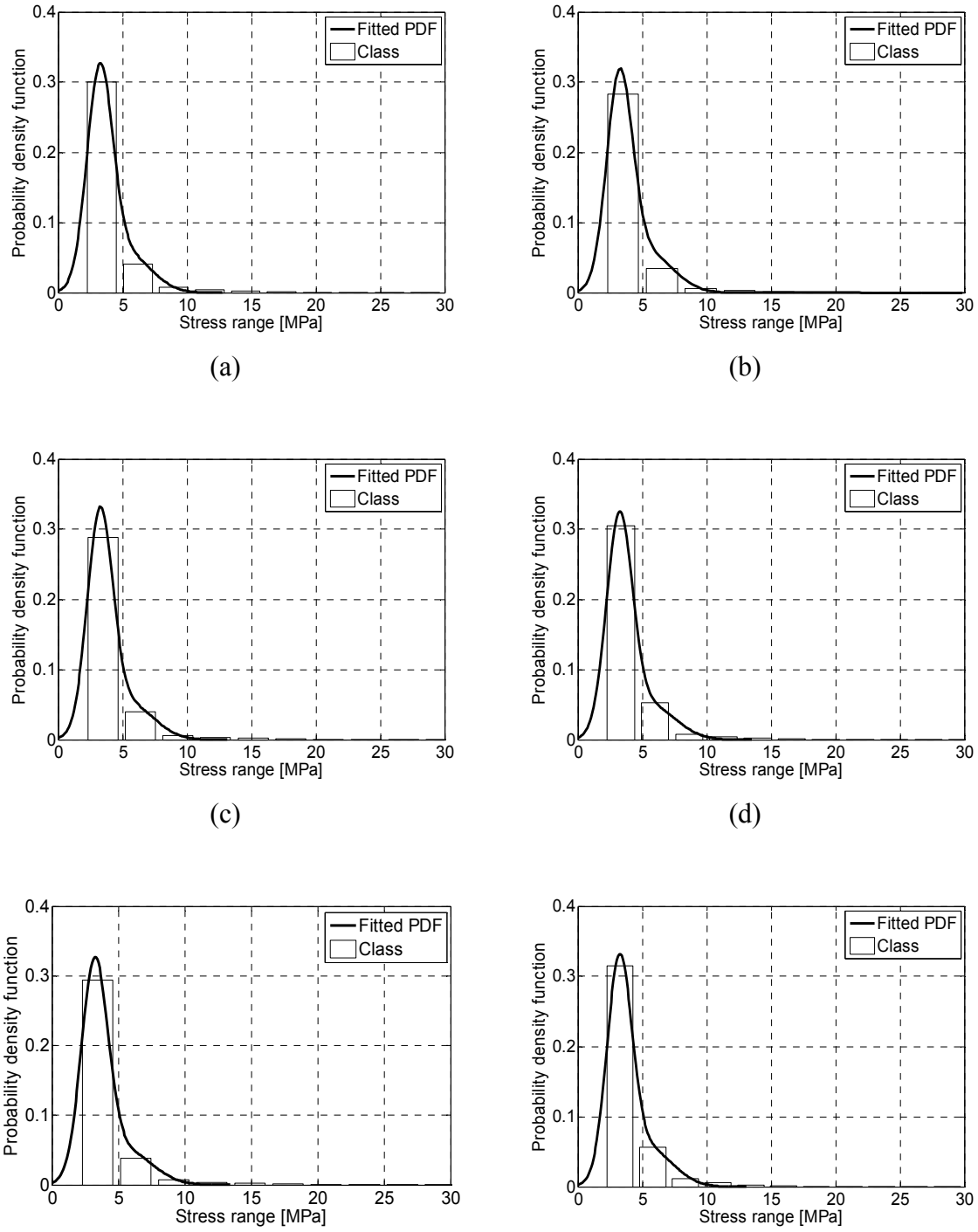


Figure 5.2 Fitted PDF for stress spectra of (a) 1st trip, (b) 2nd trip, (c) 3rd trip, (d) 4th trip, (e) 5th trip, and (f) 6th trip

Table 5.1 Calculation of posterior distributions of parameters and modeled results using stress ranges of single train trip

Dataset	Posterior distribution				Predicted distribution of stress ranges s^*		
	μ		σ				
	Mean	95% credibility interval	Mean	95% credibility interval	μ	σ	95% credibility interval
1st	3.75	[3.72, 3.78]	1.593	[1.57, 1.62]	3.75	1.59	[0.65, 6.88]
2nd	3.83	[3.79, 3.86]	1.65	[1.63, 1.68]	3.83	1.65	[0.61, 7.03]
3rd	3.81	[3.78, 3.84]	1.59	[1.57, 1.61]	3.81	1.59	[0.64, 6.93]
4th	3.78	[3.75, 3.81]	1.65	[1.63, 1.67]	3.78	1.65	[0.55, 7.05]
5th	3.73	[3.70, 3.76]	1.62	[1.60, 1.64]	3.73	1.62	[0.55, 7.01]
6th	3.77	[3.74, 3.80]	1.58	[1.56, 1.60]	3.77	1.58	[0.66, 6.87]

These findings prove that the stress spectra of six repeated train trips can be modeled by the proposed Bayesian approach. In addition, the modeling results are similar for different datasets, which verify the stable performance of the method in PDF modeling.

Figure 5.3 illustrates that the continuous prediction results of PDFs for the stress ranges of six train trips, showing stable performance of the modeling technique for predicting

stress spectra. Comparing the predicted results and the stress ranges to be used for updating, it is shown that the predicted results are close to the observations to be input in the next step (e.g. predicted results using 1st and 2nd datasets are close to the distribution shape of 3rd dataset). This performance shows the feasibility of the proposed Bayesian prediction method for modeling and predicting stress ranges.

5.4.3 Comparison of proposed frequentist and Bayesian modeling methods

According to the literature about frequentist and Bayesian modeling techniques and the verified results about the proposed frequentist and Bayesian modeling methods in Chapter 3 and 5, the pros and cons of two proposed approaches can be summarized.

The pros and cons of the proposed frequentist approach include:

- The modeling method is very efficient, for that once the best fitted PDF is defined through preliminary study, the developed method for parameters estimation in the PDF can be rapidly implemented;
- The modeling method is very flexible to apply for modeling distributions with different types of shapes (i.e. either single or mixture distribution). By using the modeling method, different models with various number of components can be used to model a distribution;
- But the developed modeling method cannot separately consider the uncertainty of the occurrence of each set of the continuous monitoring data. Specifically, the monitoring data integrates the upcoming monitoring data by averaging the

values of previous and new monitoring data, during which the probability of the occurrence of each set of the monitoring data is viewed as equal to the other sets. However, in reality, the occurrence probabilities of different sets of data may be different.

The pros and cons of the proposed Bayesian approach include:

- In the Bayesian modeling method, the parameters used to describe the model are viewed as random variables with distributions, reflecting the uncertainty in our belief about the modeling result;
- The Bayesian modeling method is preferable in updating and making prediction. Each set of the monitoring data is viewed as a specific evidence for future prediction, and will be integrated with the previous modeling result based on the formulated Bayesian updating scheme, through which the uncertainty involved in each set of the monitoring data can be separately considered;
- The Bayesian modeling method, however, is computationally intensive, consuming more time to obtain the fatigue life estimation results in comparison with the frequentist method.

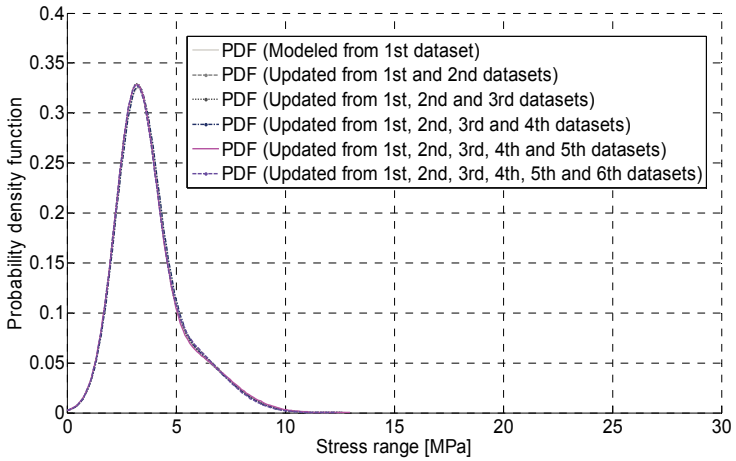


Figure 5.3 Predicted PDF of stress spectrum by Bayesian updating technique

Table 5.2 Modeled and predicted results

Data	Posterior distribution				Predicted distribution of stress ranges s*		
	μ		σ				
	Mean	95% credibility interval	Mean	95% credibility interval	μ	σ	95% credibility interval
Modeling	3.76	[3.72, 3.79]	1.59	[1.57, 1.62]	3.76	1.59	[0.65, 6.88]
1st Update	3.79	[3.76, 3.82]	1.62	[1.60 1.65]	3.79	1.59	[0.63, 6.96]
2nd Update	3.79	[3.76, 3.82]	1.61	[1.59, 1.63]	3.79	1.61	[0.63, 6.95]
3rd Update	3.79	[3.75, 3.82]	1.62	[1.59, 1.64]	3.79	1.62	[0.61, 6.98]
4th Update	3.79	[3.76, 3.83]	1.63	[1.61, 1.66]	3.79	1.63	[0.59, 6.99]
5th Update	3.78	[3.75, 3.81]	1.62	[1.59, 1.64]	3.78	1.62	[0.61, 6.97]

5.5 Summary

In this chapter, the Bayesian method is introduced for stress spectrum modeling and updating for on-board vehicle condition monitoring purpose. The performances of analytical solutions by the Bayesian method have been verified. This method is feasible in modeling and updating the PDF of stress spectra for high-speed train components. In addition, in-service monitoring stress time histories of a train running on a high-speed railway in China are used to establish the PDF of stress spectra. Specifically, the stress ranges acquired from a measurement point of the instrumented high-speed train are used in association with the Bayesian updating process. The modeling and prediction performances are analyzed.

In line with the results and findings, the following points can be drawn:

- A procedure using Bayesian inference has been established for the stress spectrum modeling, through integrating prior information and continuous monitoring data;
- The proposed stress spectrum model is sufficient to represent the variation in the in-service monitoring dataset; and
- Predicted PDF of the stress spectra acquired from continuous monitoring train trips can maintain good and stable performance when introducing more trips of data into the updating process. The modeling results can then be used for further formulation of fatigue life assessment and prediction method for long-term condition monitoring purpose, which will be shown in the next chapter.

CHAPTER 6

PROBABILISTIC FATIGUE LIFE ASSESSMENT OF HIGH-SPEED TRAINS USING IN-SERVICE MONITORING DATA: BAYESIAN APPROACH

6.1 Introduction

Recent advances in computing techniques have been in favor of the Bayesian model for probabilistic condition assessment (Stone 2013; Chiachío et al. 2014). In the last chapter, a Bayesian model has been developed for the stress spectrum. This chapter is to develop a Bayesian fatigue life prediction method for high-speed trains integrating the model developed in Chapter 5. There have been a number of studies in which Bayesian models were developed for investigating the uncertainty influence of the fatigue resistance data on fatigue analysis of civil engineering structures (Babuška et al. 2016). However, the work done on developing Bayesian approaches for fatigue life assessment using in-service monitoring data is very limited. In the Bayesian approach, the parameters of the estimated models are regarded as random variables, aiming at reflecting the uncertainties and varying observations characteristics. In addition, the predicted distributions of parameters can be obtained through updating process.

In the Bayesian inference, two solutions are usually applied to derive the posterior distribution of parameters. One is analytical (Stone 2013), and the other is simulation-based, also called the Markov chain Monte Carlo (MCMC) simulation approach (Babuška et al. 2016). The analytical solution by definition, seeks to derive an algebraic expression for the posterior distributions of parameters and predicted results. The analytical solution is able to express the result precisely and explicitly. However, it is difficult and even impossible to obtain the analytical expression for a very complex model with complicated expression. The MCMC solution reduces the calculation effort in obtaining posterior distributions of parameters through the generation of sample parameter values, and effectively achieves a simulation-based distribution for practical purposes. Although the advantage of MCMC approach to fatigue life assessment is substantial, it is still of practical value to explore the similarity between the performances of analytical and MCMC approaches in condition assessment.

Considering the above issues, in Section 6.2 a Bayesian fatigue life assessment method is developed using in-service monitoring data of high-speed train. Subsequently, the in-service monitoring data obtained from a high-speed train is used to predict the stress spectrum within the fatigue life assessment model. The time histories of both nominal and hot spot stresses are selected for verification of the effect of stress concentration on fatigue life. The prediction results of the two types of stress spectra are compared. Section 6.4 details the fatigue life prediction results using the proposed method, which integrates the estimated posterior distributions of material parameter values and the

established posterior distribution of in-service stress spectra to calculate fatigue life using the MCMC technique. The predicted distributions of fatigue life are illustrated for hot spot stresses.

6.2 Proposed Method for Fatigue Life Prediction

In the following subsections, firstly a fatigue life assessment approach is proposed utilizing the posterior distributions of the material parameters m, C derived from fatigue resistance data, and the stress spectrum $f(s)$ derived from in-service stresses data. Subsequently, the Bayesian updating technique is implemented on measured fatigue resistance data for m, C as shown in Section 6.2.2. In addition, the Bayesian modeling and updating method for formulating $f(s)$ has been demonstrated in Section 5.2.

6.2.1 Presentation of assessment procedure integrating stress spectrum and fatigue resistance data

The classical method for predicting fatigue life is based on S-N curve, in which the relationship shows the number of cycles to failure N of a material, under a constant stress level S:

$$NS^m = C \tag{6.1}$$

where C and m are material parameters of the component. To establish a relationship for the various stress spectra causing fatigue failure, the Miner's rule is used as follows:

$$D_f = \sum_i^k \frac{n_{s,i}}{N_{s,i}} \tag{6.2}$$

In a real world application, the material is manufactured to be the component of a train, experiencing continuously varying stresses imposed on the component during the service period of the train component, usually classified as k of classes of one stress spectrum, S_1, S_2,…, S_k .

The continuously varying stresses or stress spectrum occurring during each train trip show a similar pattern before the occurrence of fatigue failure. Affected by the varying stress ranges imposed on the component, the welded joint accumulates inherent fatigue damage due to the imposed stress ranges within each successive train trip, before the accumulated and irreversible fatigue damage (after many train trips) leads to component failure/cracking after reaching the fatigue limit. In order to quantify the accumulated fatigue damage impact on the component during a train trip, a fatigue ratio D is defined. The fatigue ratio D for one train trip is calculated based on Miner's rule:

$$D = \sum_i^k \frac{n_i}{N_i} \tag{6.3}$$

where n_i is the number of stress cycles that the component experienced of the ith stress class (S_i); and N_i is the number of cycles that would cause the component to fail at the ith stress class (S_i) amplitude and can be calculated based on Equation (6.1).

When the distribution of stress ranges is regarded as a continuous function, the fatigue ratio in Equation (6.3) can be rewritten as:

$$D = \int_0^P \frac{n_{tot} f(s)}{N_s} ds \tag{6.4}$$

where N_s is the number of cycles to failure when the stress range is s. P is the maximum stress range obtained during the train trip. As stress range is always positive, the lower bound of the integration is zero. The conventional assumption is that S-N curve terminates at a fatigue limit, below which failure will not occur, or one in which the S-N relationship becomes a horizontal line. For rail structures, however, the stress ranges below the knee point (also referred to as "constant amplitude fatigue limit", below which the slope of the S-N relationship declines) are included in calculating D_m by introducing a reduction factor λ_i for n_{tot}:

$$n_{tot} = \sum_{i=1}^{k} \lambda_i n_i \tag{6.5}$$

and

$$\lambda_i = \begin{cases} (S_i / \Delta\sigma)^{m-1} & \text{if } S_i \le \Delta\sigma \\ 1 & \text{if } S_i > \Delta\sigma \end{cases} \tag{6.6}$$

where n_i is the number of stress cycles as defined in Equation (6.2), m is the material parameter as defined in Equation (6.1), $\Delta\sigma$ is the knee point value that can be obtained from the documentation published by the International Institute of Welding (IIW), and the value of $\Delta\sigma$ is calculated and catalogued according to a FAT class.

By introducing Equation (6.1), Equation (6.4) is expressed as:

$$D = \int_0^P \frac{n_{tot} f(s)}{N_s} ds = \int_0^P \frac{n_{tot} f(s)}{C s^{-m}} ds \tag{6.7}$$

Therefore, if the fatigue ratio for one typical train trip is known, the predicted fatigue life F of the component is calculated as:

$$F = \frac{D_f H}{D} \tag{6.8}$$

where D_f is the fatigue life threshold of the component, with different values selected for different components. For a welded joint, the fatigue standard IIW normally sets a constant value 0.5 for D_f; H is the service life per unit, namely, the railway length in kilometers for a train trip. Combining Equations (6.7) and (6.8), the predicted fatigue life F gives:

$$F = \frac{0.5H}{\int_0^P \frac{f(s)}{N_s} ds} = \frac{0.5H}{\frac{n_{tot}}{C} \int_0^P f(s) s^m ds} \tag{6.9}$$

where the logarithm of the equation is shown as:

$$\log(F) = \log(0.5) + \log(H) - \log(n_{tot}) - \log\left(\int_0^P f(s) s^m ds \right) + \log(C) \tag{6.10}$$

6.2.2 Estimation of material parameters *m*, *C*

In Equation (6.10), two parameters C and m govern the fatigue property of a material. Their relationship with the fatigue property of a material is expressed as:

$$\log(N) = \log(C) - m\log(S) + \varepsilon \tag{6.11}$$

where N is the number of cycles to failure of a material, under a constant stress level S. It is derived as:

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i \quad \text{or} \quad y = X\beta + \varepsilon \tag{6.12}$$

where $y_i = \log(N_i)$, $x_i = \log(S_i) - \frac{1}{n}\sum_1^n \log(S_i)$, $\beta_1 = \log(C) - \frac{m}{n}\sum_1^n \log(S_i)$, $\beta_2 = -m$, $\beta = [\beta_1, \beta_2]^T$ and ε is the error.

In Equation (6.11), the error ε is assumed to follow the Gaussian distribution:

$$\varepsilon \sim N(0, \sigma^2 I) \tag{6.13}$$

$$\varepsilon = y - X\beta \sim N(0, \sigma^2 I) \tag{6.14}$$

which is expressed as:

$$f(y - X\beta \mid \beta, \sigma^2) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left(-\frac{(y - X\beta)^T (y - X\beta)}{2\sigma^2}\right) \tag{6.15}$$

It is known that:

$$f(\beta, \sigma^2) = f(\beta \mid \sigma^2) f(\sigma^2) \tag{6.16}$$

A normal-inverse Gamma prior is then used for representing knowledge on the distributions of these two parameters:

$$f(\beta \mid \sigma^2) = N(\beta_0, \sigma^2 \Sigma_0) = \frac{1}{(2\pi\sigma^2)^{2/2} \mid \Sigma_0 \mid^{1/2}} \exp\left[-\frac{(\beta - \beta_0)^T \Sigma_0^{-1} (\beta - \beta_0)}{2\sigma^2}\right] \tag{6.17}$$

$$f(\sigma^2) = IG\left(\frac{a}{2}, \frac{b}{2}\right) = \frac{\left(\frac{b}{2}\right)^{a/2}}{\Gamma\left(\frac{a}{2}\right)} (\sigma^2)^{-a/2-1} \exp\left(-\frac{b}{2\sigma^2}\right) \tag{6.18}$$

Accordingly, the posterior distribution of these two parameters can be finalized to:

$$f(\beta,\sigma^2 \mid y,X) = f(y \mid X,\beta,\sigma^2) \times f(\beta,\sigma^2) \tag{6.19}$$

It is calculated as:

$$\begin{aligned} f(\beta,\sigma^2 \mid y,X) \propto (\sigma^2)^{-2/2} \exp\left[-\frac{(\beta-\beta^*)^T(\Sigma^*)^{-1}(\beta-\beta^*)}{2\sigma^2} \right] \\ \times (\sigma^2)^{-a^*/2-1} \exp(-\frac{b^*}{2\sigma^2}) \end{aligned} \tag{6.20}$$

where

$$\Sigma^* = (X^T X + \Sigma_0^{-1})^{-1}$$

$$\beta^* = (X^T X + \Sigma_0^{-1})^{-1}(X^T y + \Sigma_0^{-1}\beta_0)$$

$$a^* = a + n$$

$$b^* = b + y^T y + \beta_0^T \Sigma_0^{-1} \beta_0 - \beta^* (\Sigma^*)^{-1} \beta^*$$

For the conjugate distribution of parameters β and σ^2, the marginal posterior distributions follow:

$$f(\beta \mid \sigma^2, y, X) \sim N(\beta^*, \sigma^2 \Sigma^*) \tag{6.21}$$

and

$$f(\sigma^2 \mid y) \sim IG(\frac{a^*}{2}, \frac{b^*}{2}) \tag{6.22}$$

6.2.2.1 Marginal posterior distribution of β

The marginal posterior distribution of β can be calculated by integrating σ^2 out from the joint posterior distribution:

$$f(\beta \mid y, X) = \int_0^\infty f(\beta,\sigma^2 \mid y, X) d\sigma^2 \tag{6.23}$$

It is expressed as:

$$f(\beta \mid y, X) \propto \int_0^{\infty} (\sigma^2)^{-(2+a^*)/2-1} \exp\left[-\frac{(\beta-\beta^*)^T (\Sigma^*)^{-1} (\beta-\beta^*) + b^*}{2\sigma^2} \right] d\sigma^2 \qquad (6.24)$$

If $L = (\beta-\beta^*)^T (\Sigma^*)^{-1} (\beta-\beta^*) + b^*$, then the above equation is simplified as:

$$f(\beta \mid y, X) = \int_0^{\infty} (\sigma^2)^{-(2+a^*)/2-1} \exp\left(-\frac{L}{2\sigma^2} \right) d\sigma^2 \qquad (6.25)$$

Assume $u = \frac{L}{2\sigma^2}$, then $\sigma^2 = \frac{L}{2u}$ and $\frac{d\sigma^2}{du} = -\frac{L}{2} u^{-2}$. Then the algorithm is simplified as:

$$f(\beta \mid y, X) = \left(\frac{L}{2u}\right)^{-(2+a^*)/2-1} \int_0^{\infty} u^{\left(k+a^*\right)/2-1} \exp\left(-u\right) du \qquad (6.26)$$

It is found that the component of Equation (6.26) satisfies the structure of Gamma function, based on the assumption that: $r=(2+\alpha^*)/2$ and $t=u$. Therefore, the marginal posterior distribution of β can be simplified as:

$$f(\beta \mid y, X) \propto \left[1 + \frac{1}{a^*} (\beta-\beta^*)^T \left(\frac{b^*}{a^*} \Sigma^*\right)^{-1} (\beta-\beta^*) \right]^{-(2+a^*)/2} \qquad (6.27)$$

Thus, the marginal posterior distribution of β follows *Student-t* distribution with mean: β^* and covariance: $\frac{b^*}{a^*-2}\Sigma^*$.

6.2.2.2 Predicted distribution of future observations $\hat{\boldsymbol{y}}$

The predictive distribution of new observations $\hat{\boldsymbol{y}} = \left(y_{n+1}, y_{n+2}, \ldots, y_{n+q}\right)$ satisfies the relationship:

$$\hat{y} = \hat{X}\beta + \hat{\varepsilon} \tag{6.28}$$

where $\hat{\boldsymbol{X}}$ is a $q \times 2$ matrix of the know elements; $\hat{\boldsymbol{\varepsilon}}$ is a $q \times 1$ vector of errors. Then the predicted distribution function of $\hat{\boldsymbol{y}}$ is given by:

$$f(\hat{y} \mid \hat{X}) = f(\hat{y} \mid \hat{X}, \beta, \sigma^2) f(\beta, \sigma^2 \mid y, X) d\beta d\sigma^2 \tag{6.29}$$

which turns out to be a multivariate *Student-t* distribution:

$$f(\hat{y} \mid \hat{X}) = \frac{\Gamma(a^* + q)}{\Gamma(\frac{a^*}{2})(\pi a^*)^{q/2}} \left|\hat{\Sigma}\right|^{-1/2} \left[1 + \frac{1}{a^*}\left(\hat{y} - \hat{\mu}\right)^T \hat{\Sigma}^{-1} \left(\hat{y} - \hat{\mu}\right)\right]^{-(a^* + q)/2} \tag{6.30}$$

where

$$\hat{\mu} = \hat{X}\beta^*$$

$$\hat{\Sigma} = \frac{b^*}{a^*}\left(I + \hat{X}\Sigma^{*-1}\hat{X}^T\right)$$

Therefore, the predictive mean and variance matrices of $\hat{\boldsymbol{y}}$ given $\hat{\boldsymbol{X}}$ are $\hat{X}\beta^*$ and $\hat{\Sigma} = \frac{b^*}{a^* - 2}\left(I + \hat{X}\Sigma^{*-1}\hat{X}^T\right)$, respectively.

6.3 Simulation-based Inference for Parameter Estimation

It is shown that for the posterior distributions of β and σ^2 based on S and N data in Section 6.2.2 and the posterior distributions of μ and σ_t^2 based on s data, the analytical

expressions can be obtained. However, for other complex models (e.g., linear regression model with different prior distributions, more sets of parameters need to be estimated with more sophisticated expressions for posterior distribution), it is difficult or too complicated to obtain the algebraic expression of the posterior distributions of parameters. Sampling-based approaches were developed by previous scholars to tackle this problem, such as the Gibbs sampler and/or the Metropolis-Hastings (MH) algorithm. These approaches reduce the calculation effort. Particularly, the Gibbs sampler is widely used and regarded as a special case of the more general MH algorithm. It applies a Markovian updating scheme to generate posterior samples by sweeping through each variable to sample from its conditional distributions with the remaining variables fixed at current values. The method can be used when the conditional distributions for all parameters are available for sampling, whereby samples of parameters can be generated straight-forwardly and efficiently by giving specified values of the conditions variables.

In this study, Gibbs sampling technique is employed to obtain samples of the C and m values by drawing samples of β and σ^2 based on S and N data from the conjugate distributions shown in Equations (6.17-19). The full conditional distributions for Gibbs sampling are formed as:

$$\left\{\begin{array}{l}\pi(\beta^{(t)} \mid (\sigma^{t-1})^2) = \pi(\beta^t) L(\beta^{t-1}, (\sigma^{t-1})^2 \mid S, N) \\ \pi((\sigma^{(t)})^2 \mid \beta^{(t)}) = \pi(\sigma^2) L(\beta^{(t)}, (\sigma^{t-1})^2 \mid S, N)\end{array}\right\} \tag{6.31}$$

Thus, for iterations T large enough it is regarded the subset of last M runs in Step 2 as the simulated observations from the corresponding posterior distributions (e.g., [β],

[σ^2], [μ], [σ_t^2]). The posterior distributions of the parameters are obtained using Gibbs samplers.

The predicted density of fatigue life can also be estimated through the Gibbs sampler output based on observations of *S, N*, and *s*:

$$\begin{aligned} P(F \mid s,S,N) &= \iiint P(F \mid s^*,c,m) \times P(s^* \mid s) \times P(c,m \mid S,N) ds^*,c,m \\ &\approx \frac{1}{M} \sum_{K=T-M+1}^{T} P(F \mid \mu,\sigma_t^2,c,m) \end{aligned} \tag{6.32}$$

The process for calculating *C* and *m* can be described as follows:

(1) Gibbs sampler for parameters β and σ^2

Step 0: Initialize the parameters $\beta^{(0)}$ and $(\sigma^2)^0$ and the number of runs *T*;

Step 1: While t< *T*, draw $\beta(t)$ from the conditional distribution $f(\beta|S, N, (\sigma^2)^{(t-1)})$, and draw $(\sigma^2)^{(t)}$ from the conditional distribution $f(\sigma^2|S, N, \beta^{(t)})$;

Step 2: Observe the samples of $\{\beta^{(1)}...,\beta^{(T)}, (\sigma^2)^{(1)}...,(\sigma^2)^{(T)}\}$, pick up the samplers from the last *M* runs (e.g., $\{\beta^{(T-M+1)}\ldots,\beta^{(T)}, (\sigma^2)^{(T-M+1)}\ldots,(\sigma^2)^{(T)}\}$) showing the stationary part of the generated realization as a subset; and

Step 3: Use the subset to calculate the samples of the parameters *C* and *m* (e.g., $\{C^{(T-M+1)}\ldots,C^{(T)}, (m)^{(T-M+1)}\ldots,(m)^{(T)}\}$), which are then applied to estimate the posterior distributions of *C* and *m*.

(2) Gibbs sampler for parameters μ and σ_t^2

When the analytical solution cannot be obtained for the posterior distribution of the parameters for a stress spectrum in Session 5.2, Gibbs sampler technique can be used to obtain the posterior distributions of μ and σ_t^2 based on s data. Description of the process is given below:

Step 0: Initialize the parameters $\mu^{(0)}$ and $(\sigma_t^2)^0$ and the number of runs T; and initialize a dataset r, a unit matrix of q dimension, where q is the length of dataset s;

Step 1: While $t< T$, draw $\mu^{(t)}$ from the conditional distribution $f(\mu|s, r, (\sigma_t^2)^{(t-1)})$; and draw $(\sigma_t^2)^{(t)}$ from the conditional distribution $f(\sigma^2|s, r, N, \mu^{(t)})$;

Step 2: Observe the samples of $\{\mu^{(1)}..., \mu^{(T)}, (\sigma_t^2)(1)...,(\sigma_t^2)^{(T)}\}$, pick up the samplers from the last M runs (e.g., $\{\mu^{(T-M+1)}..., \mu^{(T)}, (\sigma_t^2)^{(T-M+1)}...,(\sigma_t^2)^{(T)}\}$) showing the stationary part of the generated realization as a subset; and

Step 3: Use the subset of $\{\mu^{(T-M+1)}..., \mu^{(T)}, (\sigma_t^2)^{(T-M+1)}...,(\sigma_t^2)^{(T)}\}$ to estimate the posterior distributions of μ and σ_t.

After T runs of such iteration shown in Steps 2 and 3, the stationary samples arrive at $\{\mu^{(T-M+1)}\ldots, \mu^{(T)}, (\sigma_t^2)^{(T-M+1)}\ldots,(\sigma_t^2)^{(T)}\}$. It was discovered by previous research that as the number of Gibbs iterations increases to infinity, the draws from the conditional

distributions converge to the joint distribution, that are $f(\beta, \sigma^2|S, N)$ and $f(\mu, \sigma_t^2|s, r)$ in this study.

(3) Gibbs sampler for prediction of fatigue life F

Step 0: Draw $T\text{-}M+1$ runs of samples for β and σ^2 as shown in Clause (1) of this session;

Step 1: For each run of the sample $\{\beta^{(T-M+1)}\ldots,\beta^{(T)}, (\sigma^2)^{(T-M+1)}\ldots,(\sigma^2)^{(T)}\}$, calculate the associated fatigue life value F according to Equation (6.32) to obtain the dataset of samples of fatigue lives $\{F^{(T-M+1)},\ldots, F^{(T)}\}$; and

Step 2: Use the dataset of $\{F^{(T-M+1)},\ldots, F^{(T)}\}$ to estimate the posterior distribution of F.

6.4 Application to Predict Fatigue Life of High-speed Train

6.4.1 Comparison of analytical and simulation estimation performances

A set of typical fatigue resistance data (Yildirim et al. 2012) is obtained from the literature. It is used for comparing the modeling results derived from the Gibbs sampler and from the analytical solutions. As shown in Figure 6.1, this dataset consists of 47 sets of experimental fatigue test data, on the specimens that had the same welded joint

geometry. The *S-N* relationship can be classified as FAT 260 class, according to the traditional approach provided in the IIW standard.

The dataset is separated into training (30) and validation (17) datasets, as shown in Figure 6.1. The training dataset is used to estimate the posterior distributions of the parameters in the *S-N* model using the Bayesian approach, and then the obtained N^* (e.g., predicted cycles to failure) results (e.g., mean values and creditable intervals) is predicted according to different S^* (e.g., stress ranges) values. The validation dataset is used to verify the performance of the estimated parameter results, by constructing 95% credibility interval which is expected to cover the major area of the validation dataset. Both analytical and Gibbs sampler solutions are used for Bayesian modeling. Once the predicted results are obtained using the training dataset, they are verified through the validation dataset.

Figure 6.2 demonstrates the T=10000 samples drawn by the Gibbs sampler technique for the parameters β_1, β_2, and σ^2. It is seen that the ranges of the values for the three parameters stabilize at [12.5, 13.5], [-7, -2], and [0.5, 3], respectively, showing the samples for the parameters stably converge. Based on this, M is set as T in Section 6.3 and all samples are reserved for the construction of posterior distributions of the parameters. The modeled PDFs are shown in Figure 6.2.

As β_1 and β_2 in Equation (6.12) are the crucial factors determining the posterior distributions of C and m, an important focus is to compare Gibbs sampler and

analytical solutions to ensure the accuracy and applicability of β_1 and β_2. Table 6.1 lists the results of the posterior distributions of the three parameters obtained by analytical and Gibbs sampler solutions for comparison. The analytical solution and Gibbs sampler processes are introduced in Sections 6.2 and 6.3. The posterior mean and variance of β_1 estimated by Gibbs sampler are 12.84 and 0.17, and estimated by analytical solution are 12.84 and 0.17. Posterior mean and variance of β_2 estimated by Gibbs sampler are -4.82 and 0.65, and estimated by analytical solution are -4.82 and 0.66. This proves that both analytical and Gibbs sampler solutions obtain similar results (identical results in this particular case) for the parameters of the *S*-*N* model. Apart from this information, the 95% credibility intervals for the three parameters can also be estimated by the analytical and Gibbs sampler approaches. The results from the Gibbs sampler are [12.50, 12.84], [-3.53, -6.12] and [0.69, 1.97].

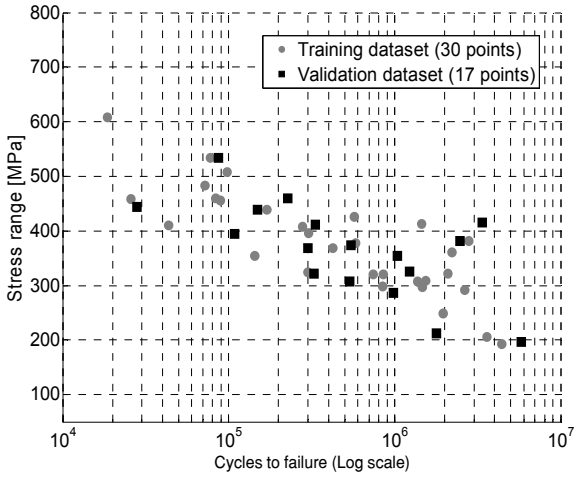


Figure 6.1 Fatigue resistance data: original data separates into training and validation datasets

Table 6.1 Summary of parameters by analytical and MCMC solutions on fatigue model by training dataset

Parameters	Mean		Variance		MC error	Median	95% credibility interval (Gibbs)
	Gibbs	Analytical	Gibbs	Analytical			
β_1	12.84	12.84	0.17	0.17	0.0018	12.84	[12.50, 12.84]
β_2	-4.82	-4.82	0.65	0.66	0.0078	4.81	[-3.53, -6.12]
σ^2	1.25		0.33		0.0034	1.22	[0.69, 1.97]

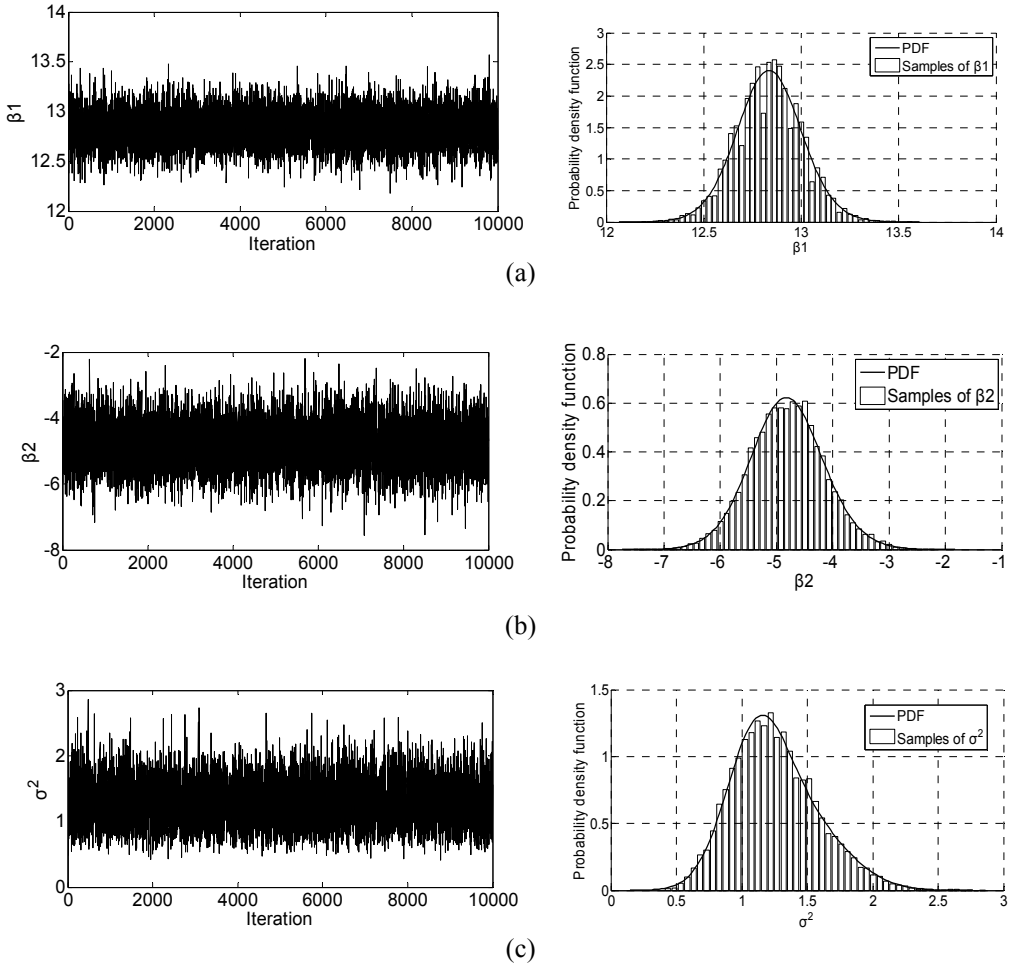


Figure 6.2 Sample paths and estimated posterior densities for parameters: (a) β_1, (b) β_2, and (c) σ^2

Once the posterior distributions of parameters are obtained, the *S*-*N* model can be established and the prediction of values of N^* corresponding to different S^* can be obtained. Table 6.2 lists the variations of N^* estimated by Gibbs sampler and analytical solutions for values of S^* from 50 to 800 MPa. It is observed from the results by the Gibbs sampler that mean values of N^* decrease with the rise of S^*, while the variance of N^* remains stable. Among the predicted results for S^*=50, …, 800 MPa, the largest mean value and variance of N^* are at S^*=800 MPa and 400 MPa, respectively. The

trend found in the results by Gibbs sampler can also be seen in the results by analytical solutions. Overall, the predicted results by the Gibbs sampler and analytical solutions are almost identical.

Comparisons of the Gibbs sampler and analytical solutions on estimation of the posterior distributions of parameters (Table 6.1) and the predicted values of N^* (Table 6.2) indicate that the two solutions have similar performance as for the linear fatigue *S-N* model applied in Section 6.2.2. Bayesian inference results obtained by Gibbs sampler and analytical solutions are very close, with respect to the posterior distributions of parameters and the predicted values of N^*.

Based on this, the Gibbs sampler approach is applied to calculate and to demonstrate the 95% credibility interval of N^* for the stress ranges from 50 to 800 MPa as shown in Table 6.2. Figure 6.3 shows a data plot of the region of this 95% credibility interval between a 2.5% lower bound and a 97.5% upper bound of predicted distributions of N^*, together with the median of the predicted results, for the training dataset (30 points) and validation dataset (17 points). It is observed that the median curve of N^* crosses the mean area of the validation dataset and that the 95% credibility interval region covers the majority of points of the validation dataset (15 included).

Table 6.2 Comparison of predicted results (i.e. number of cycles) obtained by Gibbs sampler and analytical solutions

Stress range (MPa)	Mean		Variance		95% credibility interval (Gibbs)
	Gibbs	Analytical	Gibbs	Analytical	
50	$10^{9.79}$	$10^{9.80}$	$10^{1.12}$	$10^{1.13}$	$[10^{8.39}, 10^{11.14}]$
100	$10^{8.34}$	$10^{8.35}$	$10^{0.69}$	$10^{0.70}$	$[10^{7.26}, 10^{9.41}]$
200	$10^{6.89}$	$10^{6.90}$	$10^{0.45}$	$10^{0.46}$	$[10^{6.04}, 10^{7.77}]$
300	$10^{6.05}$	$10^{6.05}$	$10^{0.40}$	$10^{0.40}$	$[10^{5.22}, 10^{6.85}]$
400	$10^{5.45}$	$10^{5.45}$	$10^{0.38}$	$10^{0.39}$	$[10^{4.63}, 10^{6.27}]$
500	$10^{4.97}$	$10^{4.98}$	$10^{0.41}$	$10^{0.41}$	$[10^{4.14}, 10^{5.81}]$
600	$10^{4.59}$	$10^{4.60}$	$10^{0.44}$	$10^{0.44}$	$[10^{3.73}, 10^{5.46}]$
700	$10^{4.28}$	$10^{4.27}$	$10^{0.48}$	$10^{0.47}$	$[10^{3.37}, 10^{5.18}]$
800	$10^{4.00}$	$10^{3.99}$	$10^{0.50}$	$10^{0.50}$	$[10^{3.09}, 10^{4.94}]$

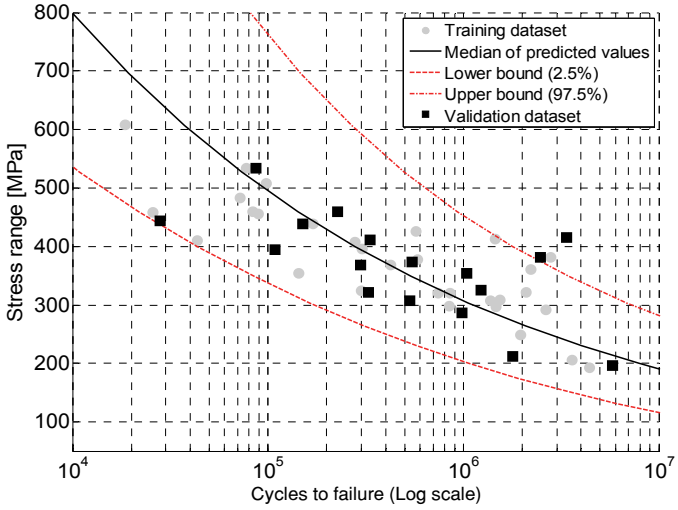


Figure 6.3 Modeled credibility intervals by Bayesian-inferenced model with validation dataset

This comparison results show that the predicted distributions of N^* can well represent the *S-N* experimental data relationship, indicating the applicability and accuracy of the Bayesian approach, either using the Gibbs sampler solution or using the analytical solution, in constructing the fatigue *S-N* model. This Bayesian approach is then applied to construct the fatigue *S-N* model and to estimate the posterior distributions of the parameters C and m for the high-speed train component material.

6.4.2 Material parameters estimation for high-speed train component

According to IIW 2008, the fatigue class FAT100 is recommended for welded joints in steel structures (Fricke and Kahl 2005). The *S-N* relationship established from the data of FAT100 is applicable for fatigue life assessment approaches based on different principles, such as nominal stress, hot spot stress, structural stress according to Dong (2005) and structural stress according to Xiao and Yamada approaches (Fricke and Kahl 2005). Therefore, fatigue resistance data supporting the FAT100 class is extracted from previous literature (Maddox 2001) for further establishment of the *S-N* relationship and the associated parameters of the welded joints of a high-speed train. The FAT100 fatigue resistance data is shown in Figure 6.4. The data is then used to construct the *S-N* relationship for welded joints of a high-speed train. Firstly, the means and variances of the posterior distributions of β_1, β_2 and σ^2 in Equation (6.12) are deduced according to the processes of Section 6.2.2, as shown in Table 6.3. Based on the analysis results of β_1 and β_2 in Table 6.3, the posterior distributions of parameters

C and *m* are constructed and plotted in Figure 6.5. As mentioned in Section 6.2.2, the posterior distributions of log(*C)* and *m* follow *Student-t* distributions. The mean and variance of *C* are $10^{12.31}$ and $10^{0.06}$ and those of *m* are -2.81 and 0.01. The constructed distribution is used below for drawing samples in order to calculate fatigue life.

Table 6.3 Summary of parameters by Gibbs sampler and analytical solutions on fatigue model by FAT 100 fatigue resistance data

Parameters	Mean	Variance	Median	95% credibility interval
	Gibbs	Gibbs		
β_1	13.65	0.05	13.65	[13.55, 13.65]
β_2	-2.81	0.11	2.81	[-2.60, -3.02]
σ^2	4.22	0.63	4.18	[3.09, 4.18]

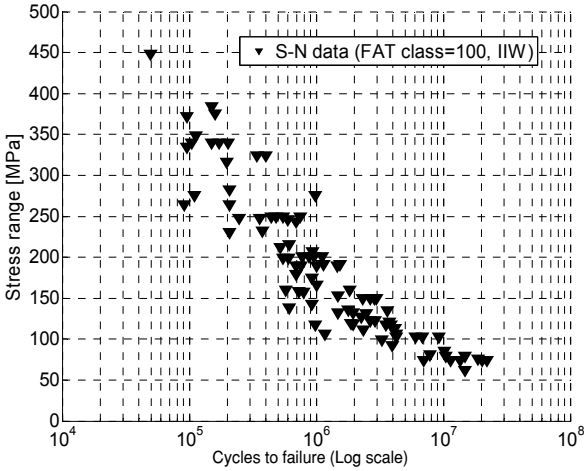
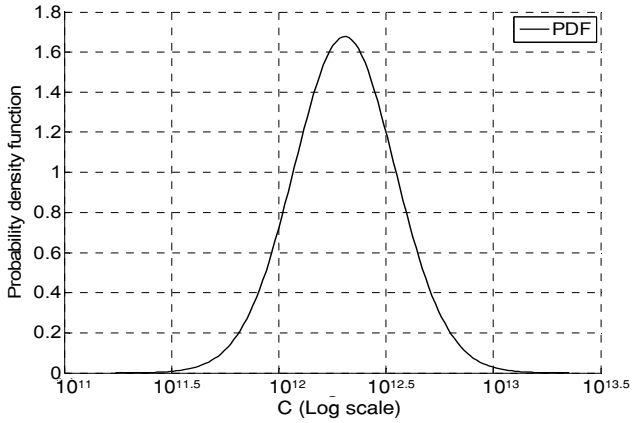
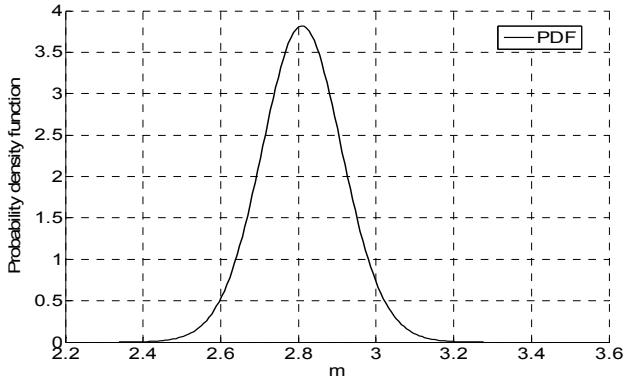


Figure 6.4 FAT100 data (Yildirim and Marquis 2012) to build up *S-N* relationship of welded joint of high-speed trains

The 95% credibility interval of the N^* for different stress ranges S^* from 50 to 500 are plotted in Figure 6.6(a). Again it is seen that the major points lie within the region of the constructed N^*. Based on the predicted results, the distributions N^* for different S^* are depicted in Figure 6.6(b). It is observed that the mean values of N^* decrease with the increase of S^*, while the variances of N^* remain almost unchanged. The information on the established N^* in the figure can be used to draw samples for the calculation of fatigue life.

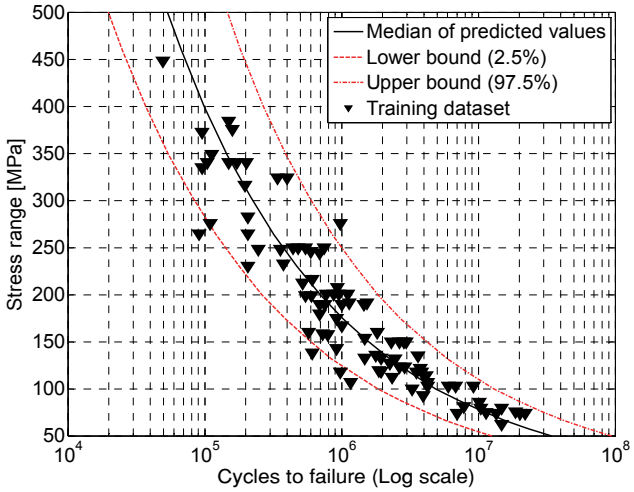


(a)

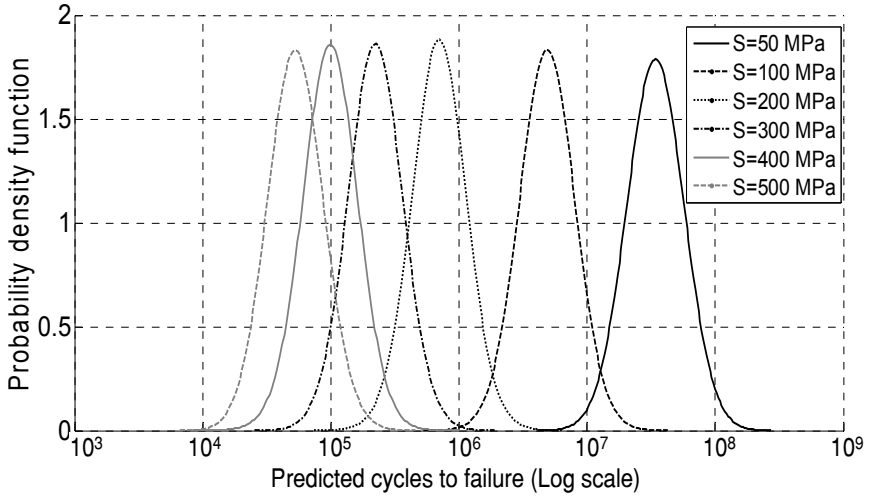


(b)

Figure 6.5 Posterior distributions estimated from Bayesian modeling: (a) C, and (b) m



(a)



(b)

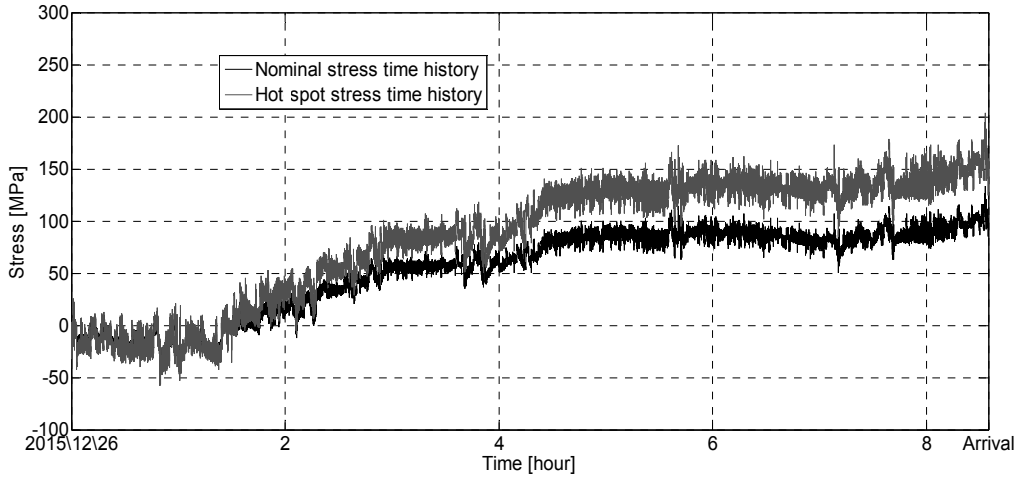
Figure 6.6 Establishment of predicted *S-N* relationship for FAT 100 data at different stress range levels: (a) predicted 95% credibility intervals, and (b) predicted number of cycles to failure

After obtaining the *S-N* relationship for the welded joint of a high-speed train, the next step is to establish the predicted distributions of stress spectrum by the Bayesian modeling method using in-service monitoring data.

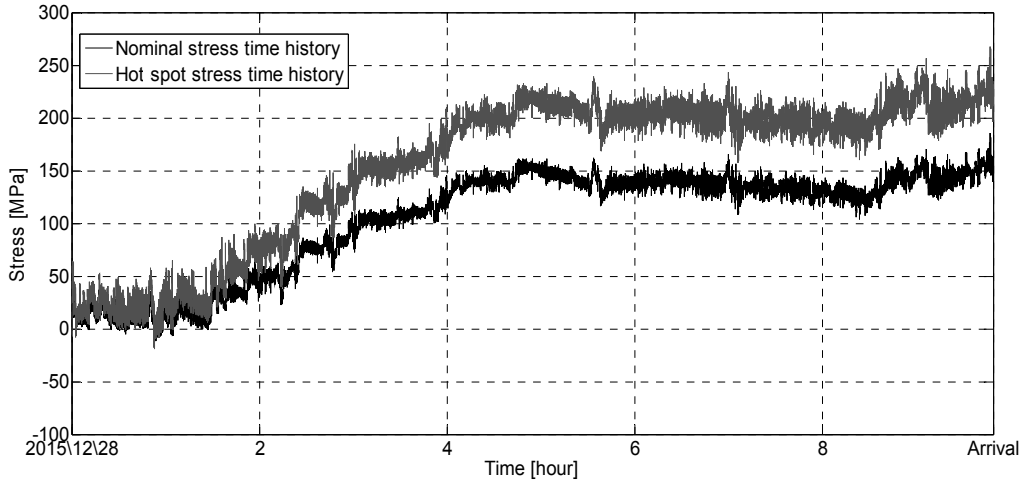
6.4.3 Constructed PDF of stress spectrum using in-service monitoring data

Bayesian updating of the stress spectrum is performed for a welded joint on a train running on a high-speed railway in China using the monitoring data introduced in Chapter 4. The welded joint is selected as the 7th measurement point shown in Table 4.1 and Figure 4.3. Detail description about the on-board monitoring can be found in Section 4.3. The strain time histories of the whole train trips measured at the 7th measurement point are used for analysis of stress spectrum modeling and fatigue life assessment in the following sections. To investigate the influence of hot spot stress on fatigue life, the stress spectra for nominal and hot spot stresses, on the same monitored welded joint are compared.

Specifically, two trips of data from the strain sensors are selected for analysis. The data was acquired from the monitoring system and was recorded on 2015/12/26 and 2015/12/28. Below, this work names the 2015/12/26 trip as "1st trip" and the 2015/12/28 trip as "2nd trip". By multiplying strain data with the elasticity modulus E of steel, the stress time histories are obtained. Examples of such strain-time histories for both the nominal and hot spot stress regions of the welded joint are shown in Figure 6.7.



(a)



(b)

Figure 6.7 Nominal and hot spot stresses time histories obtained by strain sensors of welded joint during train running on high-speed train at (a) 1st trip, and (b) 2nd trip

Once the stress time history is obtained, it is processed by the following steps, to determine the spectrum of stress ranges:

- The stress time history is composed of a series of repeated stress cycles. Rainflow counting method is used to extract the stress ranges for each stress cycle. From each individual stress cycle, the stress range S_a is obtained by

calculating the difference between the maximum stress (σ_{max}) and minimum stress (σ_{min});

- Filtering out those stress ranges with amplitudes smaller than 1.3 MPa for nominal stresses and 4 MPa for hot spot stresses (the majority of such can be considered as noise);

- Calculating all of the stress ranges in the stress time history;

- Constructing the stress spectrum of the stress ranges and illustrating the relationship between the stress range levels and the number of cycles at each stress range level.

Figures 6.8 and 6.9 demonstrate the stress ranges and the stress spectrum for the nominal and hot spot stress time histories. It is seen that the majority of stress ranges for both nominal and hot spot stresses are below 40 MPa, and the mean of the hot spot stress ranges are larger than that of the nominal stresses. The largest nominal and hot spot stress ranges for the 1st trip are 65.06 and 91.42 MPa, respectively; and for the 2nd trip are 62.90 and 90.64 MPa, respectively. This shows that the range amplitudes obtained for these two trips remain stable for both nominal and hot spot stresses. The distribution profiles of the stress spectra for the two different trips are similar to each other. The distribution profiles are also similar for both nominal and hot spot stresses. These stress spectra are then used to investigate the performance of the Bayesian approach in estimating posterior distributions.

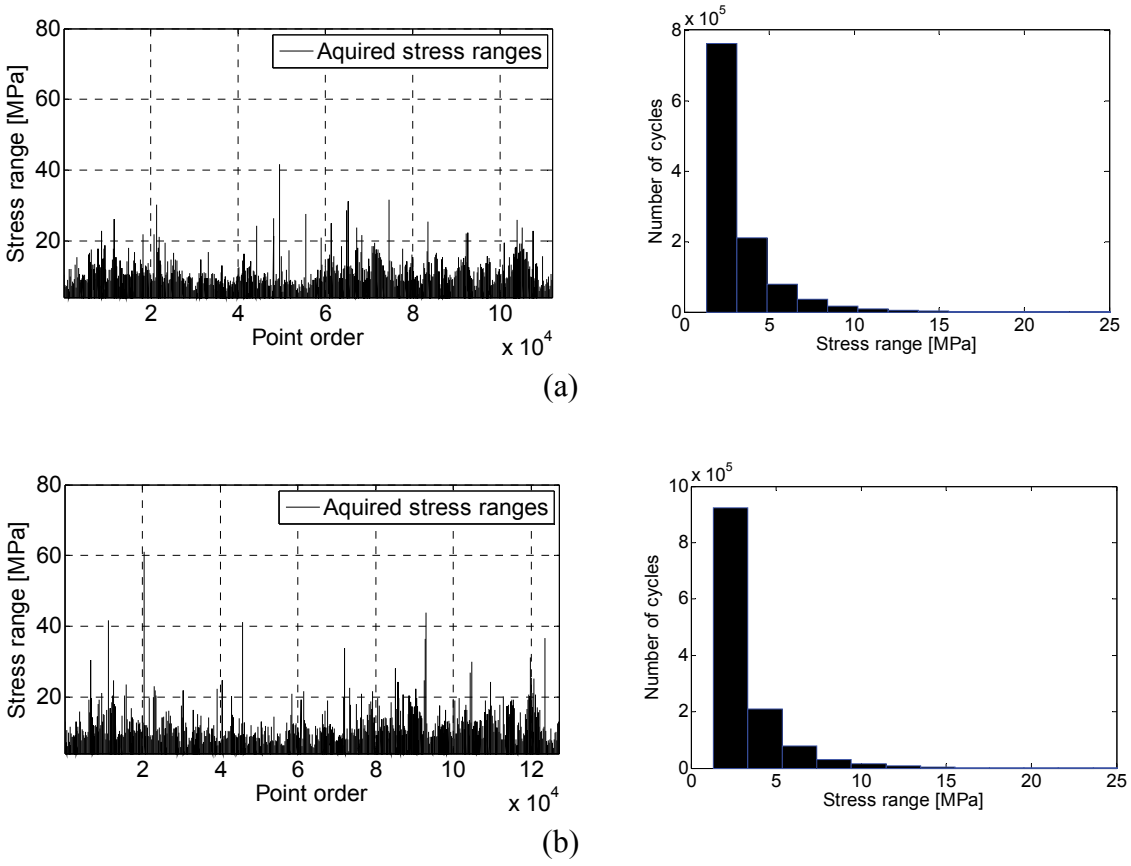


Figure 6.8 Stress ranges and corresponding stress spectrum of nominal stress time histories of (a) 1st trip, and (b) 2nd trip

The results of the posterior distributions of μ and σ as well as the information relating to the predicted distribution of s^* are summarized in Table 6.4 and drawn in Figure 6.10, for the stress spectra of both nominal and hot spot stresses relating to the 1st and 2nd trips. For the nominal stress, the posterior distribution of μ and σ and the predicted stress spectra PDFs are similar for the 1st and 2nd trips, as shown in Figure 6.10. The phenomenon also exists for in the cases of hot spot stress.

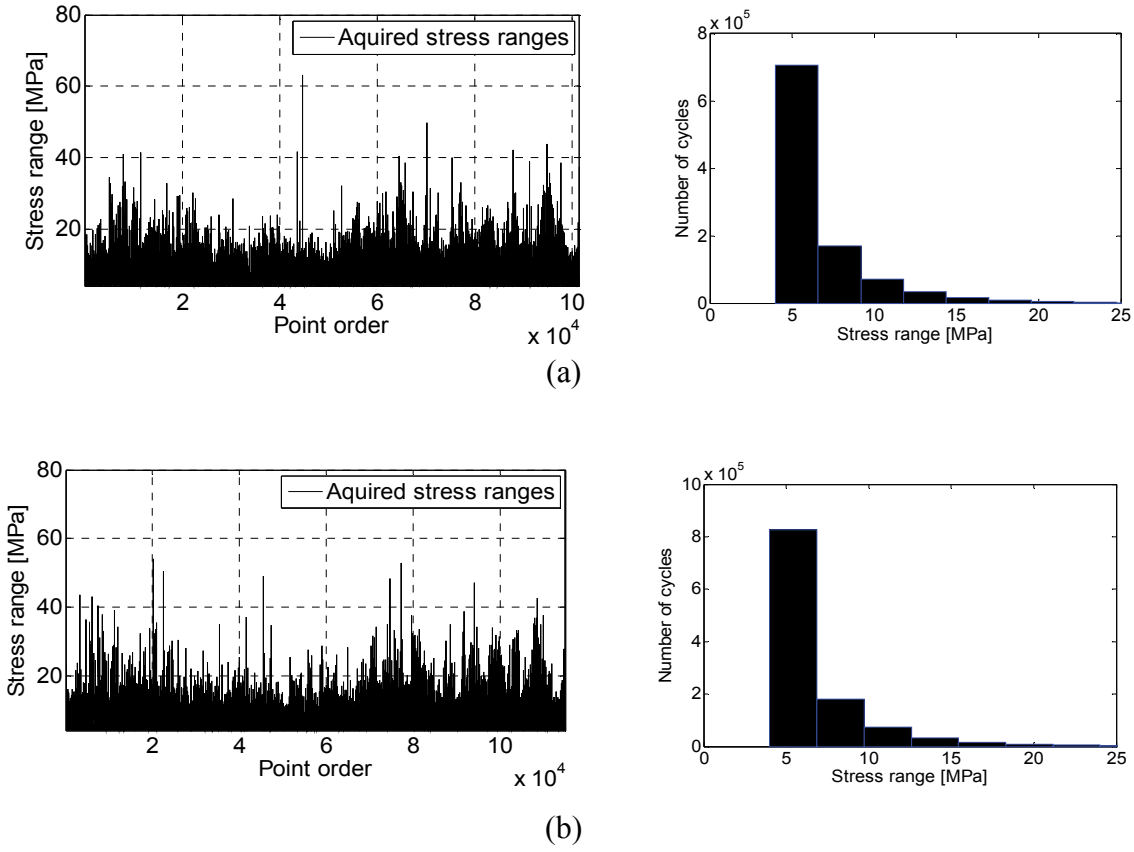
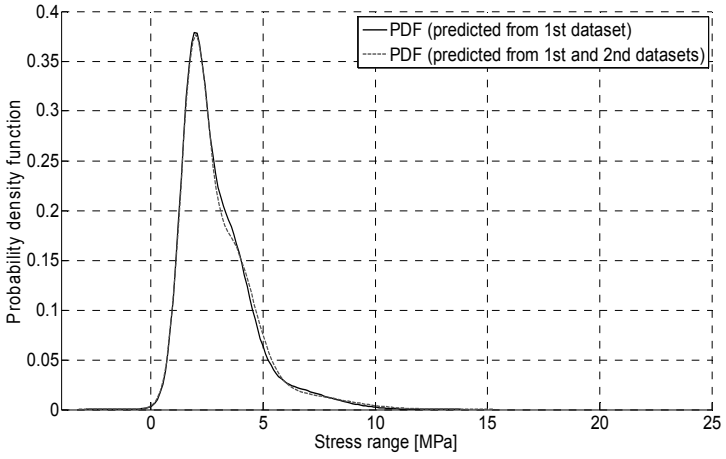
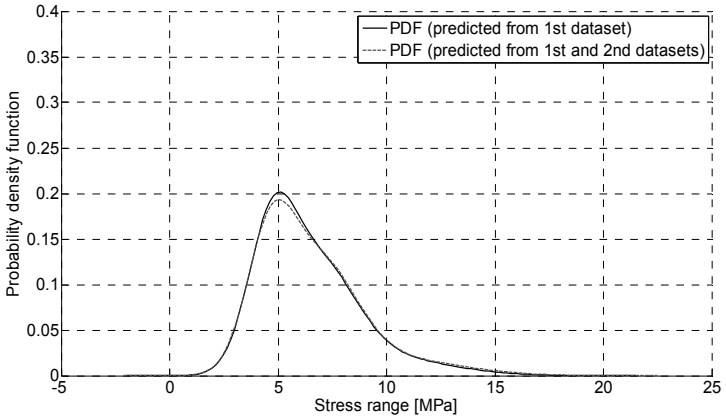


Figure 6.9 Stress ranges and corresponding stress spectrum of hot spot stress time histories of (a) 1st trip, and (b) 2nd trip

In comparison with the nominal and hot spot stresses obtained during the same trips (either first or second trips), it is also seen that the mean values of μ in the hot spot stress cases are larger than in the cases of nominal stress, whereas the mean values of σ in both two cases are similar. The 95% credibility interval of s^* in hot spot stress case is more scattered than in the nominal stress case. The stress spectra of both nominal and hot spot stresses are modeled and described by the proposed Bayesian approach.



(a)



(b)

Figure 6.10 Predicted PDF of stress spectrum for stress ranges by Bayesian updating technique: (a) nominal stress spectrum, and (b) hot spot stress spectrum

Table 6.4 Calculation of posterior distributions of parameters and predicted stress ranges

Data	Stress type	Posterior distribution				Predicted distribution of stress ranges s^*		
		μ		σ				
		Mean	95% credibility interval	Mean	95% credibility interval	Mean	Std	95% credibility interval
1st trip	Nominal	2.931	[2.927, 2.935]	2.590	[2.576, 2.604]	2.931	2.590	[0, 6.795]
	Hot spot	6.452	[6.446, 6.458]	2.465	[2.459, 2.471]	6.452	2.465	[1.621, 11.283]
2nd trip	Nominal	2.956	[2.953, 2.959]	2.703	[2.693, 2.713]	2.956	2.703	[0, 7.293]
	Hot spot	6.494	[6.489, 6.498]	2.527	[2.522, 2.531]	6.494	2.527	[1.527, 11.447]

6.4.4 Fatigue life prediction results

Based on the results obtained for parameters C and m in Session 6.4.2 and the stress spectra in Section 6.4.3, the corresponding algorithm to predict fatigue life F is expressed as follows:

$$\log(F) = \log(0.5) + \log(H) - \log(n_{tot}) - \log\left(\int_0^P f(s^*)(s^*)^m ds^*\right) + \log(C) \qquad (6.33)$$

where

$$f(s^*) = \frac{\pi^2 \tau\left(\frac{n}{2}\right)(n-1)^2}{\tau\left(\frac{n-1}{2}\right)} \left(\frac{n}{\frac{1}{n-1}\sum_{i=1}^{n}\left(s_i - \bar{s}\right)^2 (n+1)} \right)^{\frac{1}{2}} \left[1 + \frac{n\left(s^* - \bar{s}\right)^2}{\sum_{i=1}^{n}\left(s_i - \bar{s}\right)^2 (n+1)} \right]^{-\frac{n}{2}} \quad (6.34)$$

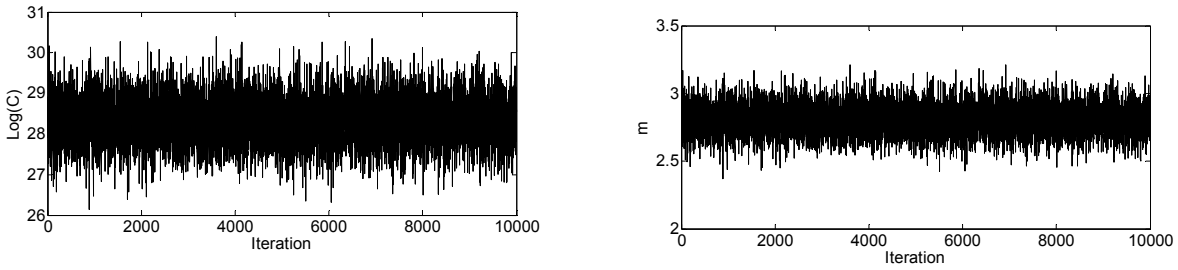
Equation (6.34) gives the posterior distribution of fatigue life *P(F| c, m, s)*. It is then used to calculate the predicted fatigue life using the Gibbs sampler technique. The sample values of *log(C)* and *m* used in Equation (6.34) can be drawn from the posterior distributions of parameters shown in Figure 6.5. In this study, 10000 samples of *log(C)* and *m* are generated to calculate the samples of fatigue life *F*, as shown in Figure 6.11.

The samples of predicted fatigue life *F* for nominal stresses for the 1st and 2nd trips and the PDF constructed are shown in Figure 6.12, and the samples of predicted fatigue life *F* for hot spot of both trips and the constructed PDF are shown in Figure 6.13. Information such as mean, standard deviation, PDF and 95% credibility interval is also provided in the figures. It is observed that the values are stable. For the nominal stress case, the fatigue life obtained based on the 1st trip is close to that based on the 2nd trip in terms of mean, standard deviation as well as the 95% credibility interval. This is also true for the hot spot stress case.

By considering the duty of a high-speed train being 500,000 km/year and the demanded operating period being 30 years, the design fatigue life in terms of mileage has been obtained as $L_s = 1.5 \times 10^7$ km. It can be seen that the lower bound of the 95%

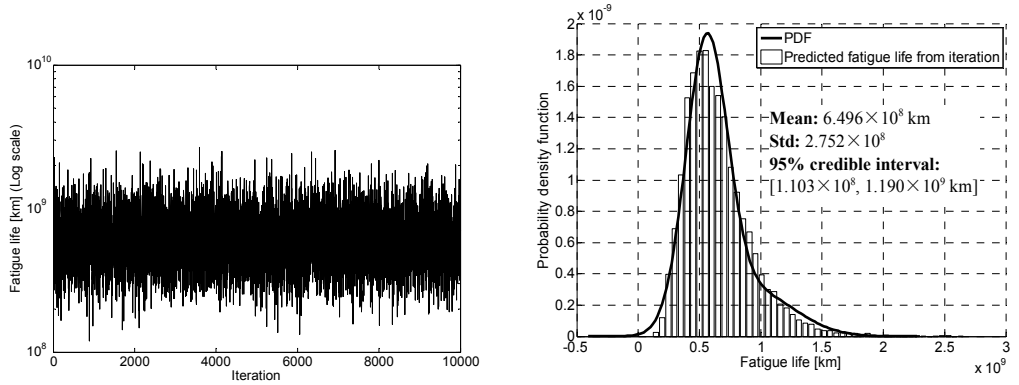
credibility interval of the fatigue life for the hot spot stress case is around 2.7×10^7 km, being higher but approaching the design fatigue life. In addition, the mean of the estimated fatigue life can be expressed as "number of operating years" by taking 500,000 km as yearly traveling distance: for the hot spot stress cases, the mean values of the evaluated fatigue lives are around 177 years from the 1st dataset, and around 174 years from the 1st and 2nd datasets. The above findings verify the stability of the proposed fatigue life prediction technique.

The mean values of fatigue life predictions in the nominal stress case are larger than that in the hot spot stress case, which is as expected according to the results of Section 4.2. For the welded joint of the tubular crosspiece, the average ratio of the mean of the predicted fatigue life from the nominal stresses to the mean from the hot-spot stresses is about 7.3:1. This value indicates that the estimated fatigue life results from the two types of stresses are significantly different, and only the predicted fatigue life result from the hot-spot stresses is appropriate when using the fatigue life assessment approaches developed in this study. Therefore, it is recommended to use hot spot stresses for less risk fatigue life predictions when utilizing in-service monitoring data.

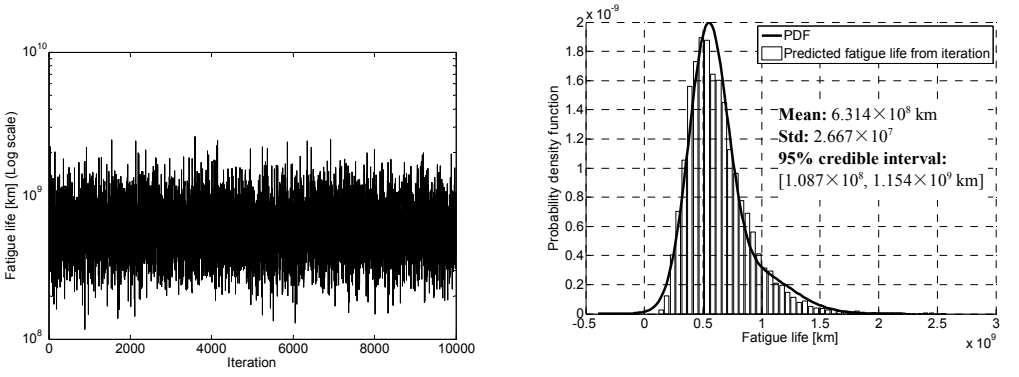


(a) (b)

Figure 6.11 Samples drawn from predicted conjugate parameters distributions using MCMC technique: (a) *log(C)*, and (b) *m*

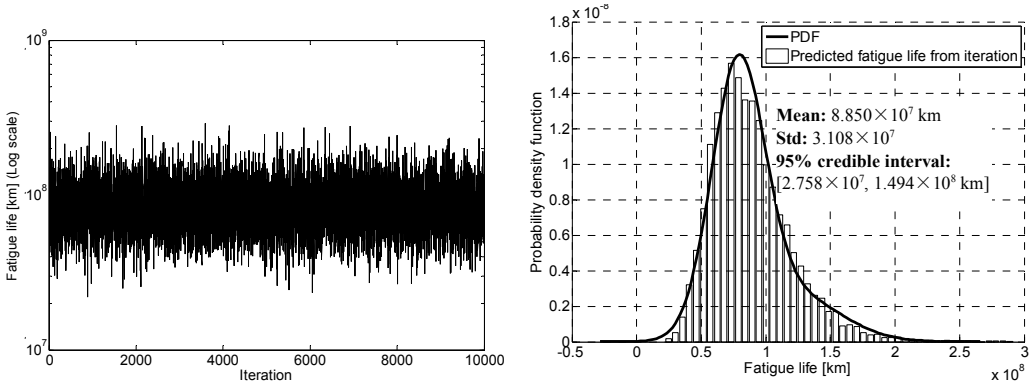


(a)

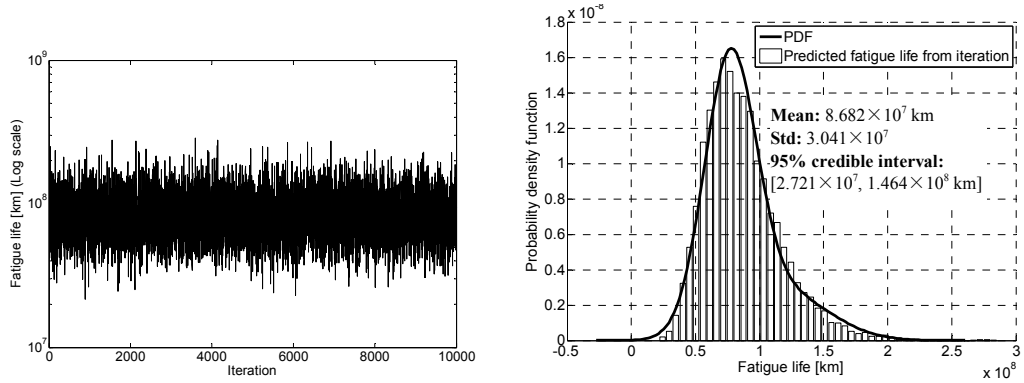


(b)

Figure 6.12 Samples of predicted fatigue lives and estimated PDF by Gibbs sampler technique using predicted nominal stress spectrum s^* obtained from (a) 1st dataset, and (b) 1st and 2nd datasets



(a)



(b)

Figure 6.13 Samples of predicted fatigue lives and estimated PDF by Gibbs sampler technique using predicted hot spot stress spectrum s^* obtained from (a) 1st dataset, and (b) 1st and 2nd datasets

6.4.5 Comparison of proposed frequentist and Bayesian fatigue life assessment approaches

The fatigue life assessment approach presented in Chapter 6 is fully based on Bayesian framework. The approach employs the PDF of the stress spectrum obtained by the Bayesian method described in Chapter 5, which utilizes the prior information and monitoring data. In addition, as the distributions of material parameters in *S-N*

relationship are obtained using Bayesian inference, the final fatigue life estimation results consider the uncertainties in both material parameters and monitoring data.

In contrast, the frequentist fatigue life assessment approach presented in Chapter 4 considers only one kind of uncertainty that arises in the stress spectrum, which is derived from the monitoring data. Different from the frequentist fatigue life assessment approach, the Bayesian fatigue life assessment approach takes into account two kinds of uncertainty that arise in the stress spectrum and in the material parameters of *S*-*N* relationship. In addition, the Bayesian fatigue life assessment approach can incorporate the prior information about the random parameters involved. Another difference of the two approaches is that the fatigue life prediction result from the frequentist approach is a fixed value, while the fatigue life prediction result from the Bayesian approach is a distribution from which both mean and variation of the estimation can be obtained.

A comparison is made between the fatigue life assessment results given in Chapters 4 and 6 for the same type of welded joint. The predicted fatigue life using the hot spot stresses is 6.842×10^7 km by the proposed frequentist approach; whereas the proposed Bayesian approach yields the prediction result in terms of a distribution with its mean 8.682×10^7 km and its standard deviation 3.041×10^7 km.

The fatigue life prediction results for the same type of welded joint are different by the two methods. It is reasonable since the two methods are based on different modeling

assumptions (including model shape, model selection rule and parameter estimation rules). Specifically, in the Bayesian method, the estimated result of fatigue life is a distribution rather than a fixed value. This is partially because the uncertainty of material parameters is considered, marking the parameters C and m in the fatigue calculation algorithm be variables. In this way, the predicted fatigue life result is obtained as a probability distribution.

As the average yearly running distance (mileage) of a typical high-speed train is 500,000 km, then in the frequentist approach, the fatigue life of high-speed train is around 137 years; and in the Bayesian approach, the mean of the predicted fatigue life of high-speed train is around 174 years. They are much longer than the design life of a typical high-speed train (30 years). This indicates the train component is safe in terms of fatigue resistance.

6.5 Summary

In this chapter, a probabilistic fatigue life assessment approach using Bayesian modeling technique has been developed for on-board condition monitoring of high-speed trains components. It makes use of fatigue resistance data and in-service monitoring data. The material parameters and stress spectrum relating to train components are estimated through Bayesian updating techniques, integrating both analytical and MCMC solutions.

The performances of both analytical and MCMC solutions (e.g., Gibbs sampler) for Bayesian updating have been compared using the fatigue resistance data. The performances of both approaches are similar, confirming the accuracy and feasibility of the Bayesian approach in estimating material parameters. Based on these, this Bayesian approach using the Gibbs sampler is applied to construct a fatigue *S*-*N* model for the high-speed train component material and to estimate material parameter values relating to fatigue life assessment.

In addition, in-service monitoring stress time histories for a train running on a high-speed railway in China are used to establish the predicted stress spectrum with a Bayesian approach as an essential part of the proposed fatigue life assessment method. The time histories of nominal and hot spot stresses at the same measurement point are used in association with the Bayesian updating process. The prediction results of the two stress spectra are compared. Finally, the estimated posterior distributions of material parameters and established posterior distribution of in-service stress spectra are integrated to calculate fatigue life using the Gibbs sampler technique. The predicted distributions of fatigue life are obtained.

In line with the results and findings, the following points can be drawn:

1. In the Bayesian model for fatigue resistance data, the 95% credibility interval is sufficient to describe the statistical scatter in the data;
2. The uncertainties in material properties and in-service stress histories can be taken into account by the proposed Bayesian fatigue life prediction model;

3. Similar performance is observed in the MCMC solution (e.g., Gibbs sampler) and the analytical solution on PDF modeling of material parameters; and
4. The predicted fatigue life of high-speed trains can be estimated by using the proposed probabilistic approach with in-service monitoring data. It would be feasible to implement the proposed Bayesian fatigue life assessment and prediction approach for high-speed trains with instrumentation system for long-term monitoring.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

This study has been devoted to connecting on-board vehicle condition monitoring technique with fatigue life assessment implementations on high-speed train components. The main focus is to develop probabilistic methods for integrating the stress monitoring data into fatigue life assessment for high-speed trains instrumented with on-board monitoring system. Both frequentist inference and Bayesian inference techniques are used to develop probabilistic fatigue life assessment approaches. In addition to high speed trains, the developed modeling methods and fatigue life assessment approaches can be applied for monitoring-based fatigue life assessment of a variety of welded structures such as aircraft, automobile, bridges, buildings, wind turbine, pressure vessels, and pipelines.

Through recent advanced on-board vehicle condition monitoring technique, monitoring data can be acquired to reflect in-service train status, which is expected to help develop accurate and implementable condition assessment methods for train life prediction. Taking in-service high-speed trains as research objects, this study explores the use of the monitoring data for probabilistic fatigue life assessment. To achieve this

goal, the influence of the in-service stress spectrum with inherent uncertainties in fatigue life assessment is particularly considered using the proposed processes based on frequentist inference and Bayesian inference. Specifically, this thesis has addressed the issues relating to formulating *S*-*N*-based fatigue life assessment method in consideration of uncertainties in the monitoring data, which is done by modeling the stress spectrum with appropriate PDF, and integrating the formulated PDF into fatigue life assessment approaches.

In summary, the major contributions of the study are as follows: the study develops a probabilistic fatigue life assessment methodology for high-speed trains instrumented with on-board vehicle condition monitoring system. The methodology has established two branches of fatigue life assessment approaches for in-service high-speed trains. One branch is based on frequentist framework, and the other branch is based on Bayesian framework.

In the frequentist inference for fatigue life assessment, the proposed approach offers entirely quick inference via the constructed PDF. The specific findings and conclusions are as follows:

1. Similar patterns can be found for the stress ranges distributions of repeated train trips. Therefore, it is appropriate to model the standard PDF to extract general features and describe the tail behaviors of stress spectra;
2. The proposed frequentist modeling technique is a feasible solution for describing the stress spectra derived from in-service stress time histories;

3. The proposed frequentist modeling technique is particularly applicable for constructing stress spectrum with the use of monitoring data;
4. PDF of the proposed indicator Q can be adequately modeled using a mixture distribution technique and the Q distribution conforms to Gaussian mixture distribution;
5. The influence of mean stresses on fatigue life is negligible for the case of normal train operating condition, and may significantly shorten fatigue life when tensile mean stresses are predominant in overall mean stresses, under the scope of Goodman relation in reflecting mean stress effect;
6. A fatigue life assessment approach has been developed, which can extract and consider the impacts of varying stress ranges and mean stresses in the monitoring data on fatigue life, using frequentist modeling technique to consider the uncertainties in both stress ranges and mean stresses; and
7. The proposed frequentist fatigue life assessment approach can effectively extract information from monitoring data, and is relatively reliable for implementation on high-speed trains with instrumentation system for short-term monitoring.

In the Bayesian inference for fatigue analysis, a stress spectrum modeling process is presented and fatigue life assessment approach is further developed from a fully Bayesian perspective. The approach can predict the PDF of modeled stress spectrum and the results about fatigue life. The specific findings and conclusions in this part are as follows:

1. A procedure using Bayesian inference has been established for the stress spectrum modeling, through integrating prior information and continuous monitoring data;
2. The proposed stress spectrum model is sufficient to represent the variation in the in-service monitoring dataset;
3. Robust performance is observed in obtaining the predicted PDF results by Bayesian updating process for the stress spectra acquired from continuous monitoring train trips;
4. It is suitable by using the proposed Bayesian modeling and updating technique to derive stress spectrum for further assessment of fatigue life;
5. The uncertainties in material properties and in-service stress histories can be taken into account by the proposed Bayesian fatigue life prediction model;
6. Similar performance is observed in the MCMC solution (e.g., Gibbs sampler) and the analytical solution on PDF modeling of material parameters; and
7. It would be feasible to implement the proposed Bayesian fatigue life assessment and prediction approach for high-speed trains with instrumentation system for long-term monitoring.

7.2 Future Work

The established methodology for probabilistic fatigue life assessment of high-speed trains using in-service monitoring data needs further development to facilitate the railway authorities to make decisions more timely and reliably. The specific recommendations are as follows.

(1) Establishment of a time-varying fatigue life assessment by using continuous monitoring data and Bayesian Dynamic Model (BDM)

In the current study, the modeling technique is employed to understand the distribution behavior of stress ranges, by regarding the stress ranges as the frequency-domain information converted from the original monitored stress time history. This conversion causes the information loss about the relationship between the time moment and stress value, which may result in less precise assessment outcomes. Therefore, a better way for probabilistic fatigue life assessment can be achieved through BDM to treat the time series as the output of a dynamic system perturbed by random disturbances. Previous studies indicated that the BDM is suitable for modeling univariate and multivariate time series, also for representing the structural changes, non-stationarity and irregular patterns (Yang 2013). In the future study, current research work can be extended to apply BDM for probabilistic fatigue life assessment and prediction.

(2) Development of Bayesian fatigue life assessment approach using mixture distribution model

In the current study, the frequentist approach applies mixture PDFs to model the stress spectrum and mean stresses, achieve more precise description of the tail behavior and shape of the original distributions. This fitted results are integrated into further fatigue life assessment. But in the current Bayesian approach, only the single Gaussian distribution is applied to model the PDF for stress spectrum and does not model the effect of mean stresses. As mean stresses naturally exhibit multimodality, the distribution of mean stresses is unable to be modeled by single PDF. Currently, Bayesian modeling using mixture PDF is still an ongoing research topic in the field of machine learning (Bishop 2006). There have been some research progresses in using Gaussian mixture distribution to conduct modeling task. But its applicability in real monitoring data has not yet been confirmed. Recognizing this, the future work is recommended to investigate the application of Gaussian mixture PDF modeling for the multimodal mean stress representation and the Bayesian fatigue life assessment approach considering mean stress effect.

(3) Development of fatigue life assessment approach for general high-speed train components with *S*-*N*-based and fracture mechanics theories

As mentioned in Section 2.1.1, the fatigue life for welded joints is mainly consumed during the fatigue initiation period. There are many locations on train components

other than welded joints also experiencing fatigue problems and worthy of conducting fatigue life assessment. Their fatigue lives may comprise both crack initiation and crack propagation periods. Some investigators have investigated the fatigue crack propagation phenomenon of train components interpreted by crack propagation theory, which also provides a good explanation of fatigue crack during crack propagation period (Smith 1984, Zerbst et al. 2005). In addition, there are hybrid algorithms developed by utilizing both *S*-*N* and fracture mechanics theories which can consider the fatigue life in crack initiation and propagation periods (Cui 2002). The future focus of the research is to investigate the development of fatigue life assessment approach using the aforementioned hybrid algorithms for high-speed trains. By considering the two-state fatigue life characteristics, it is targeted to address the fatigue life consumption according to different train components. This approach can be further implementable to the on-board vehicle condition monitoring system.

APPENDIX I

PROCESS OF LABORATORY VERIFICATION OF DEVELOPED METHODS

The fatigue life assessment approach proposed in this thesis can be verified using laboratory experiments. The procedures for experimental verification are as follows:

Firstly, construct a stress/load spectrum which is generated from the modeled PDF of in-service stress spectra. According to the modeled PDF, the distribution of the stress ranges at different amplitudes and the total number of stress cycles per train trip are obtained, which will be used to reconstruct the load spectrum to be imposed on a test specimen in the laboratory;

Secondly, use the stress spectrum to establish a pattern of repeated loadings which will be imposed on the test specimen;

Thirdly, impose the repeating loadings on the specimen through the fatigue test rig and begin to record the total number of stress cycles imposed on the specimen;

Fourthly, continuously impose the repeating loads on the test specimen until fatigue failure occurs. And then record the number of stress cycles applied;

Fifthly, convert the recorded number of stress cycles to failure, to the corresponding fatigue life of the specimen; meanwhile, based on the same stress spectrum mentioned in the first step, use the proposed probabilistic fatigue life assessment method to estimate the fatigue life;

Finally, the authentic fatigue life obtained from the laboratory experiment is compared with the fatigue life estimated by the proposed method. The result of this comparison can verify the proposed method.

APPENDIX II

FREQUENTIST VERSUS BAYESIAN STATISTICS

Probability density function (PDF) is in general used to describe the probability that an investigated random variable would fall within a particular range of values. The shape of the PDF is defined by its (model) parameters (e.g. mean, standard deviation).

In frequentist statistics, each of the model parameters is a fixed value.

In Bayesian statistics, the shape of PDF is also defined by the model parameters. However, those parameters are not fixed values, but rather they are random variables which are also described by distributions.

The different treatments on the parameters of the PDF in frequentist and Bayesian statistics reveal the difference in how probability is used and also result in the difference in how inference or estimation result is described. Especially, Bayesian statistics enables the incorporation of prior knowledge and the quantification of uncertainty in inference/estimation.

The frequentist defines probability as the long-run frequency of a certain measurement or observation. In frequentist statistics, it is believed that there is a single true PDF

about the investigated object. But it is unable to know the true PDF. Feasible solution is to obtain samples from the investigated object. The distribution of the samples can be regarded as one paradigm generated by the true distribution. By modeling a PDF based on the samples, the modeled PDF is a representative of the true distribution. And when the number of samples becomes larger, the modeled PDF is believed to have more accurate representation about the true distribution.

Taking an analysis of a particular experiment where 20 coin-tosses were performed as an example. Suppose the result got 13 heads and 7 tails. The frequentist sees the particular outcome as an instance of an infinity of possible outcomes, given a particular truth. Therefore, the frequentist would then calculate probabilities of the particular outcome, given the truth, based on the following equation*:* $p(13h7t|truth)$. The probability of possible future occurrence is calculated accordingly. In all, the frequentist calculates a probability of an event given a truth and tries to provide any statement based on using this calculated result.

On the other hand, the Bayesian defines probability as the plausibility of a hypothesis given incomplete knowledge, and uses probability more widely to model both sampling and uncertainty (e.g. model uncertainty). In Bayesian statistics, it is believed that other than the samples of investigated object, there always exist other kinds of information (e.g. previous results, past experience) to be integrated for obtaining the final decision about the PDF of investigated object. Those other types of information are quantified as the prior information in obtaining the PDF. With the information from

the samples and the prior information as well, the final PDF can be obtained. In the PDF, the parameters used to describe the PDF also have distributions, reflecting the uncertainty of the model. And when more samples are available, the probability about the modeled PDF is updated.

Again taking an analysis of a particular experiment where 20 coin-tosses were performed as the example, the result was still 13 heads and 7 tails. Then in Bayesian perspective, it does not only consider the probability of "13 heads and 7 tails" given the truth *p(13h7t|truth)*, but it also views the knowledge about the truth as a hypothesis, which is quantified as a belief about (or probability of) the truth *p(truth)*. Then probability relating to the extent of belief about the truth given the occurrence of "13 heads and 7 tails" can be obtained by:

$$p(truth \mid 13h7t) = \frac{p(13h7t \mid truth)\,p(truth)}{p(13h7t)} \tag{II.1}$$

This resulting probability is used to predict the probability of next event. In all, the Bayesian uses the samples to test and update the hypothesis, and uses the updated hypothesis to predict possible future outcomes.

REFERENCES

Akaike, H. (1977). "On entropy maximisation principle", *Proceedings of the Symposium on Applications of Statistics*. North Holland, Netherland, 27-47.

Al-Rubaie, K.S. (2008). "A general model for stress-life fatigue prediction", *Materialwissenschaft und Werkstofftechnik*, 39(6): 400-406.

Ambaum, M.H. (2012). "Frequentist vs Bayesian statistics: a non-statisticians view", *arXiv preprint arXiv*,1208: 2141.

An, D., Choi, J.H., Kim, N.H. and Pattabhiraman, S. (2011). "Fatigue life prediction based on Bayesian approach to incorporate field data into probability model", *Structural Engineering and Mechanics*, 37(4): 427-442.

Babuška, I., Sawlan, Z., Scavino, M., Szabó, B. and Tempone, R. (2016). "Bayesian inference and model comparison for metallic fatigue data", *Computer Methods in Applied Mechanics and Engineering*, 304: 171-196.

Baek, S.H., Cho, S.S. and Joo, W.S. (2008). "Fatigue life prediction based on the rainflow cycle counting method for the end beam of a freight car bogie", *International Journal of Fatigue*, 9(1): 95-101.

Balageas, D. (2006). *Introduction to Structural Health Monitoring*. ISTE Ltd., London, UK.

Balageas, D., Fritzen, C.P. and Güemes, A. (2010). *Structural Health Monitoring*. John Wiley and Sons, New Jersey, USA.

Barke, D. and Chiu, W.K. (2005). "Structural health monitoring in the railway industry: a review", *Structural Health Monitoring*, 4(1): 81-93.

Barron, I. (2016). "Coming back from a trip on high-speed trains in the 2040s", *Engineering*, 2(3): 292-293.

Bathias, C. (1999). "There is no infinite fatigue life in metallic materials", *Fatigue and Fracture of Engineering Materials and Structures*, 22(7): 559-565.

Bathias, C., Drouillac, L. and Le Francois, P. (2001). "How and why the fatigue *S-N* curve does not approach a horizontal asymptote", *International Journal of Fatigue*, 23: 143-151.

Bayraktar, M., Tahrali, N. and Guclu, R. (2010). "Reliability and fatigue life evaluation of railway axles", *Journal of Mechanical Science and Technology*, 24(3): 671-679.

Beer, F.P., Johnston, R., DeWolf, J. and Mazurek, D. (2011). "Stress and strain: axial loading", *Mechanics of Materials*, 57-69.

Beganovic, N. and Dirk, S. (2017). "Remaining lifetime modeling using State-of-Health estimation", *Mechanical Systems and Signal Processing*, 92: 107-123.

Benedettini, O., Baines, T.S., Lightfoot, H.W. and Greenough, R.M. (2009). "State-of-the-art in integrated vehicle health management", *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 223(2): 157-170.

Bentley, D.E. and Hatch, C.T. (2003). "Fundamentals of rotating machinery diagnostics", *Mechanical Engineering-CIME*, 125(10): 61-62.

Beretta, S., Carboni, M. and Cervello, S. (2011). "Design review of a freight railway axle: fatigue damage versus damage tolerance", *Materialwissenschaft Und Werkstofftechnik*, 42(12): 1099-1104.

Beretta, S. and Regazzi, D. (2016). "Probabilistic fatigue assessment for railway axles and derivation of a simple format for damage calculations", *International Journal of Fatigue*, 86: 13-23.

Bishop, C.M. (1995). *Neural Networks for Pattern Recognition.* Oxford University Press, Oxford, UK.

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning*. Springer, New York, USA.

Bladon, K., Rennison, D., Izbinsky, G., Tracy, R. and Bladon, T. (2004). "Predictive condition monitoring of railway rolling stock", *Proceedings of the Conference on Railway Engineering*. Darwin, UK.

Boiler, A.S.M.E. and Code, P.V. (2007). *Rules for Construction of Pressure Vessels. Section VIII, Div, 1*.

Boller, C. and Buderath, M. (2007). “Fatigue in aerostructures: where structural health monitoring can contribute to a complex subject”, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1851): 561-587.

Bosselmann, T. (2005). “Innovative applications of fiber-optic sensors in energy and transportation”, *Proceedings of the 17th International Conference on Optical Fibre Sensors.* Bellingham, USA*,* 5855: 188-193.

Bozdogan, H. (1987). “Model selection and Akaike’s information criterion (AIC): the general theory and its analytical extensions”, *Psychometrika*, 52(3): 345-370.

Braithwaite, F. (1854). “On the fatigue and consequent fracture of metals”, *Institution of Civil Engineers, Minutes of Proceedings.* London, UK.

Branger J. (1969). “Full-scale test and service experience in the Swiss Venom case”, *Proceedings of the ICAF Technical Session*. National Aeronautical Research Institute, Sweden.

BS7608 (2014). *Code of Practice for Fatigue Design and Assessment of Steel Structures*. British Standards Institution, UK.

BS7910 (2005). *Guidance on Methods for Assessing the Acceptability of Flaws in Metallic Structures*. British Standards Institution, UK.

Buch, A. (1979). “Comparison of various gust loading test results for notched specimens with the aid of the relative miner rule”, *Materialprüfung*, 21:

193-198.

Buch, A. (1988). *Fatigue Strength Calculation*. Trans Tech Publications, Switzerland.

Campos, J. and de Rus, G. (2009). "Some stylized facts about high-speed rail: a review of HSR experiences around the world", *Transport Policy*, 16(1): 19-28.

Castillo, E. and Fernandez-Canteli, A. (2009). *A Unified Statistical Methodology for Modeling Fatigue Damage.* Springer Science and Business Media, Germany.

Castillo, E., Fernández-Canteli, A., Koller, R., Ruiz-Ripoll, M.L. and García, A. (2009). "A statistical fatigue model covering the tension and compression Wöhler fields", *Probabilistic Engineering Mechanics*, 24(2): 199-209.

Cera, A., Mancini, G., Leonardi, V. and Bertini, L. (2008). "Analysis of methodologies for fatigue calculation for railway bogie frames", *Proceedings of the 8th World Congress on Railway Research*. Seoul, Korea.

Chang, F.K. and Tinoco Jr, I. (1996). "1997. In structural health monitoring, current status and perspectives", *Proceedings of the International Workshop on Structural Health Monitoring*. Stanford, USA, 18-20.

Charles, G., Goodall, R. and Dixon, R. (2006). "Wheel-rail profile estimation", *Proceedings of International Conference on Railway Condition*

Monitoring. Birmingham, UK, 32-37.

Chattopadhyay, A., Glinka, G., El-Zein, M., Qian, J. and Formas, R. (2011). "Stress analysis and fatigue of welded structures", *Welding in the World*, 55(7-8): 2-21.

Chiachío, M., Juan C., Guillermo, R. and James L.B. (2014). "Predicting fatigue damage in composites: A Bayesian framework", *Structural Safety,* 51: 57-68.

Chong, S.Y., Lee, J.R. and Shin, H.J. (2010). "A review of health and operation monitoring technologies for trains", *Smart Structures and Systems*, 6(9): 1079-1105.

Chookah, M., Nuhi, M. and Modarres, M. (2011). "A probabilistic physics-of-failure model for prognostic health management of structures subject to pitting and corrosion-fatigue", *Reliability Engineering and System Safety*, 96(12): 1601-1610.

Cui, W. (2002). "A state-of-the-art review on fatigue life prediction methods for metal structures", *Journal of Marine Science and Technology*, 7(1): 43-56.

Dietz, S., Netter, H. and Sachau, D. (1998). "Fatigue life prediction of a railway bogie under dynamic loads through simulation", *Vehicle System Dynamics,* 29(6): 385-402.

Donahue, R.J., Clark, H.M., Atanmo, P., Kumble, R. and McEvily, A.J. (1972).

"Crack opening displacement and the rate of fatigue crack growth", *International Journal of Fracture*, 8(2): 209-219.

Donelson III, J., Edwards, M.C., Filkins, M.H., Punwani, S.K., Stewart, M.F., Toth, D.G. and Zavis, W.M. (2005). "Performance of an on-board monitoring system in a revenue service demonstration", *Proceedings of the 2005 ASME/IEEE Joint Rail Conference*. San Diego, USA.

Dong, P., Zhang, J. and Hong, J.K. (2005). *U.S. Patent No. 6,901,809*. U.S. Patent and Trademark Office, Washington, USA.

Downing, S.D. and Socie, D.F. (1982). "Simple rainflow counting algorithms", *International Journal of Fatigue*, 4(1): 31-40.

EN13445 (2009). *Unfired Pressure Vessels-Part 3*. BSI British Standards, UK.

EN13749 (2005). *Railway Applications-Methods of Specifying Structural Requirements of Bogie Frames*. European Standard, Brussels.

Esderts, A., Willen, J. and Kassner, M. (2012). "Fatigue strength analysis of welded joints in closed steel sections in rail vehicles", *International Journal of Fatigue*, 34(1): 112-121.

Ewing, J.A. and Humfrey, J.C.W. (1903). "The fracture of metals under repeated alternations of stress", *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 200: 241-250.

European Council (2011). *High Speed Rail, III*. Retrieved from http://www.publi

cations.parliament.uk/pa/cm201012/cmselect/cmtran/writev/rail/m53.htm, UK.

Farrar, C.R. and Worden, K. (2007). "An introduction to structural health monitoring", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851): 303-315.

Fatemi, A. and Yang, L. (1998). "Cumulative fatigue damage and life prediction theories: a survey of the state of the art for homogeneous materials", *International Journal of Fatigue*, 20(1): 9-34.

Federal Railroad Administration Office of Safety Analysis (2008). *Accident Causes* http://safetydata.fra.dot.gov/officeofsafety/publicsite/Query/inccaus.aspx, USA.

Forsyth, P.J.E. (1969). *The Physical Basis of Metal Fatigue*. Blackie and Son Ltd., London, UK.

Fraley, C. and Raftery, A.E. (1998). "How many clusters? Which clustering method? Answers via model-based cluster analysis", *The Computer Journal*, 41(8): 578-588.

Freeman, L. (1996). "Bayesian statistical methods-a natural way to access clinical evidence", *British Medicine Journal*, 313: 569-570.

Frangopol, D.M. (2011). "Life-cycle performance, management, and optimisation of structural systems under uncertainty: accomplishments and challenges", *Structure and Infrastructure Engineering*, 7(6): 389-413.

Freudenthal, A.M. and Gumbel, E.J. (1956). “Physical and statistical aspects of fatigue”, *Advances in Applied Mechanics*, 4: 117-158.

Fricke, W. (2002). “Evaluation of hot spot stresses in complex welded structures”, *Proceedings of the IIW Fatigue Seminar.* Tokyo, Japan.

Fricke, W. and Kahl A. (2005). “Comparison of different structural stress approaches for fatigue assessment of welded ship structures”, *Marine Structures,* 18(7): 473-488.

Frýba, L. (1996). *Dynamics of Railway Bridges*. Thomas Telford, London, UK.

Fuchs, H.O. and Stephens, I. (1980). *Metal Fatigue in Engineering*. John Wiley and Sons, New York, USA.

Gedeon, M. (2014). “Goodman (Haigh) and other related diagrams”, *Technical Tidbits*, Materion Brush Inc., USA, 64.

Gelfand, A.E., Hills, S.E., Racine-Poon, A. and Smith, A.F. (1990). “Illustration of Bayesian inference in normal data models using Gibbs sampling”, *Journal of the American Statistical Association*, 85(412): 972-985.

Givoni, M. (2006). “Development and impact of the modern high-speed train: a review”, *Transport Reviews*, 26(5): 593-611.

Grassie, S.L. and Elkins, J.A. (2006). “Traction and curving behaviour of a railway bogie”, *Vehicle System Dynamics*, 44(1): 883-891.

Griffiths, J., Mogford, I. and Richards, C. (1971). “Influence of mean stress on fatigue-crack propagation in a ferritic weld metal”, *Metal Science*

Journal, 5(1): 150-154.

Grubisic, V. and Fischer, G. (2006). "Procedure for reliable durability validation of train axles", *Materialwissenschaft Und Werkstofftechnik*, 37(12): 973-982.

Grubisic, V. and Fischer, G. (2012). "Railway axle failures and durability validation", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 226(5): 518-529.

Guida, M. and Penta, F. (2010). "A Bayesian analysis of fatigue data", *Structural Safety*, 32(1): 64-76.

Gurney, T.R. (1979). *Fatigue of Welded Structures*. Cambridge University Press, London, UK.

Haagensen, P.J. (1997). "IIW's round robin and design recommendations for improvement methods", *Proceedings of the IIW Conference on Performance of Dynamically Loaded Welded Structures*. San Francisco, USA, 305.

Henrysson, H.F. (2002) "Effects of mean stress and crack closure on fatigue life of spot welds", *Fatigue and Fracture of Engineering Materials and Structures*, 25(12): 1175-1185.

Hirakawa, K., Toyama, K. and Kubota, M. (1998). "The analysis and prevention of failure in railway axles", *International Journal of Fatigue*, 20(2): 135-144.

Ho, T.K., Liu, S.Y., Ho, Y.T., Ho, K.H., Wong, K.K., Lee, K.Y., Tarn, H.Y. and Ho, S.L. (2008). "Signature analysis on wheel-rail interaction for rail defect detection", *Proceedings of the 4th IET International Conference on Railway Condition Monitoring*. Derby, UK, 36.

Hobbacher, A. (2009). *Recommendations for Fatigue Design of Welded Joints and Components*. Welding Research Council, New York, USA.

Hobbacher, A. (2010). "New developments at the recent update of the IIW recommendations for fatigue of welded joints and components", *Steel Construction*, 3(4): 231-242.

Hong, M., Wang, Q., Su, Z. and Cheng, L. (2014). "In situ health monitoring for bogie systems of CRH380 train on Beijing-Shanghai high-speed railway", *Mechanical Systems and Signal Processing*, 45(2): 378-395.

Hotchkiss, E. and Larney, C. (2011). *California High-speed Rail Authority Strategic Energy Plan*. National Renewable Energy Laboratory, Golden, Colorado, USA.

Hunt, D.L., Weiss, S.P., West, W.M., Dunlap, T.A. and Freesmeyer, S. (1990). *Development and Implementation of a Shuttle Modal Inspection System*. NASA Technical Reports, USA.

Hu, W., Liu, Z., Liu, D. and Hai, X. (2017). "Fatigue failure analysis of high-speed train gearbox housings", *Engineering Failure Analysis*, 73: 57-71.

Ifeng (2011). Retrieved from http://news.ifeng.com/mainland/special/jinghu

gaotie/content-2/detail_2011_08/22/8573483_0.shtmlChina.

INFRAS/IWW (2000). “External costs of transport, report commissioned by UIC”, *UIC Publications*, Paris.

Iwnicki, S. (2006). *Handbook of Railway Vehicle Dynamics.* CRC press, Florida, USA.

Iwnicki, S., Stow, J. and Andersson, E. (2006). “Field testing and instrumentation of railway vehicles”, *Handbook of Railway Vehicle Dynamics*, 445-452.

Jackson, C. (2015). “World railways in an era of change: yesterday, today and tomorrow”, *Proceedings of the 13th International Conference on Railway Engineering*. Edinburgh, UK.

Jeffreys, H. (1946). “An invariant form for the prior probability in estimation problems”, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 453-461.

Jin, X.S. and Shen, Z.Y. (2001). “Rolling contact fatigue of wheel / rail and its advanced research progress”, *Journal of The China Railway Society*, 23(2): 92-108.

JIS E4207 (2004). *Truck Frames for Railway Rolling Stock: General Rules for Design*. Japanese Industrial Standard, Japan.

JSSC (1995). *Fatigue Design Recommendations for Steel Structures*. Japanese Society of Steel Construction Technical Report No. 32, Tokyo, Japan.

Kassner, M. (2012) “Fatigue strength analysis of a welded railway vehicle

structure by different methods", *International Journal of Fatigue*, 34(1): 103-111.

Kihl, D.P. and Sarkani, S. (1999). "Mean strain effects in fatigue of welded steel joints", *Probabilistic Engineering Mechanics*, 14(1): 97-104.

Kim, J.S. (2006). "Fatigue assessment of tilting bogie frame for Korean tilting train: analysis and static tests", *Engineering Failure Analysis*, 13(8): 1326-1337.

Klesnil, M. and Lukáš, P. (1992). *Fatigue of Metallic Materials*. Elsevier, Netherlands.

Ko, J.M. and Ni, Y.Q. (2005). "Technology developments in structural health monitoring of large-scale bridges", *Engineering Structures*, 27(12): 1715-1725.

Kwon, K. and Frangopol, D.M. (2010). "Bridge fatigue reliability assessment using probability density functions of equivalent stress range based on field monitoring data", *International Journal of Fatigue*, 32(8): 1221-1232.

Lagnebäck, R. (2007). *Evaluation of Wayside Condition Monitoring Technologies for Condition-based Maintenance of Railway Vehicles*. PhD Thesis, Luleå Tekniska Universitet, Sweden.

Lagoda, T. and Sonsino, C.M. (2004). "Comparison of different methods for presenting variable amplitude loading fatigue results",

Materialwissenschaft Und Werkstofftechnik, 35(1): 13-20.

Leander, J., Andersson, A. and Karoumi, R. (2010). "Monitoring and enhanced fatigue evaluation of a steel railway bridge", *Engineering Structures*, 32(3): 854-863.

Leboeuf, M. (2016). "High-speed rail: opportunities and threats", *Engineering*, 2(4): 402-408.

Lee, C.W. and Lee, D.H. (2015). "Design life analysis of KTX running gear systems on various operating conditions", *Key Engineering Materials*, 625: 674-677.

Lee, J.S., Choi, S., Kim, S.S., Park, C. and Kim, Y.G. (2012). "A mixed filtering approach for track condition monitoring using accelerometers on the axle box and bogie", *IEEE Transactions on Instrumentation and Measurement*, 61(3): 749-758.

Lee, Y.L., Pan, J., Hathaway, R.B. and Barkey, M.E. (2005). *Fatigue Testing and Analysis: Theory and Practice*. Elsevier Butterworth-Heinemann, New Jersey, USA.

Leung, C.M., Or, S.W., Ho, S.L. and Lee, K.Y. (2014). "Wireless condition monitoring of train traction systems using magnetoelectric passive current sensors", *IEEE Sensors Journal*, 14(12): 4305-4314.

Li, X.F., Fang, J., Liang, S.L., Ma, J.J., Li, M.G., Ma, S.Q. and Zhao, W.Z. (2015) "Structural fatigue capacity assessment on gearbox in the high-speed

train based on dynamic monitoring data and fatigue evaluation system", *Proceedings of the 10th International Workshop on Structural Health Monitoring*. Stanford, USA.

Li, X.F., Xie, S., Shi, H. and Zhao, W.Z. (2007). "Study on the analysis method for fatigue life prediction of vehicle welded structure", *China Railway Science*, 28(3): 74-78.

Li, Z., Molodova, M., Nunez, A. and Dollevoet, R. (2015). "Improvements in axle box acceleration measurements for the detection of light squats in railway infrastructure", *IEEE Transactions on Industrial Electronics*, 62(7): 4385-4397.

Liu, X.Z. and Ni, Y.Q. (2017). "Wheel tread defect detection for high-speed trains using wheel impact load detector", *Proceedings of the 2017 International Conference on Computational Technologies in Concrete Structures*. Ilsan, Korea.

Lonsdale, C. and Stone, D. (2004). "North American axle failure experience", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 293-298.

Lou, J., Gui, A., Moh, A., Zhu, K., He, L. and Ching, C. (2011). *China High-speed Rail on the Economic Fast Track*. Morgan Stanley Blue Paper, USA.

Luo, R.K., Gabbitas, B.L., Brickle, B.V. and Wu, W.X. (1998). "Fatigue damage evaluation for a railway vehicle bogie using appropriate sampling

frequencies", *Vehicle System Dynamics*, 29: 404-415.

Maddox, S.J. (2001). "Recommended hot spot stress design *S-N* curves for fatigue assessment of FPSOs", *Proceedings of the 11th International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers, Stavanger, Norway.

Madshus, C., Bessason, B. and Hårvik, L. (1996). "Prediction model for low frequency vibration from high-speed railways on soft ground", *Journal of Sound and Vibration*, 193(1): 195-203.

Mancini, G. and Cera, A. (2008). "Design of railway bogies in compliance with new EN 13749 European standard", *Proceedings of the 7th World Congress on Railway Research*. Montreal, Canada.

Mann, J.Y. (1970, 1978, 1983, 1990). *Bibliography on the Fatigue of Materials, Components and Structures, Volumes 1 to 4*. Pergamon Press, Oxford, UK.

Manson, S.S. and Hirschberg, M.H. (1964). *Fatigue: An Interdisciplinary Approach*. Syracuse University Press, Syracuse, USA, 133.

Matsumoto, M., Masai, K. and Wajima, T. (1999). "New technologies for railway trains", *Hitachi Review*, 48(3): 134-138.

Maurin, L., Boussoir, J., Rougeault, S., Bugaud, M., Ferdinand, P., Landrot, A.G. and Chauvin, T. (2002). "FBG-based smart composite bogies for railway applications", *Proceedings of 15th Optical Fiber Sensors Conference*

Technical Digest. Portland, USA, 91-94.

McLachlan, G. and Krishnan, T. (2007). *The EM Algorithm and Extensions*. John Wiley and Sons, New Jersey, USA.

Miles J.W. (1954). "On structural fatigue under random loading", *Journal of the Aeronautical Sciences*, 753-762.

Miller, K.J. (1999). "A historical perspective of the important parameters of metal fatigue and problems for the next century", *Proceedings of the 7th International Fatigue Congress*. Higher Education Press, Beijing, China, 15-39.

Miller, K.J. and O'donnell, W.J. (1999). "The fatigue limit and its elimination", *Fatigue and Fracture of Engineering Materials and Structures*, 22(7): 545-557.

Milne, I., Ritchie, R.O. and Karihaloo, B.L. (2003). *Comprehensive Structural Integrity: Cyclic Loading and Fatigue*. Elsevier, Netherlands.

Miner, M.A. (1945). "Cumulative damage in fatigue", *Journal of Applied Mechanics*, 12(3): 159-164.

Mita, A. and Yokoi, I. (2000). "Fiber Bragg grating accelerometer for structural health monitoring", *Proceedings of the 5th International Conference on Motion and Vibration Control*. Sydney, Australia.

Mohammadi, J., Garg, V. K. and Subei, N. (1986). "Probabilistic fatigue failure analysis of a railroad freight car", *Vehicle System Dynamics*, 15(2): 93-

104.

Montgomery, D.C. (1997). *Introduction to Statistical Quality Control*. John Wiley and Sons Inc., New York, USA.

Moore, P.L. and Booth, G. (2014). *The Welding Engineer's Guide to Fracture and Fatigue*. Elsevier, Netherlands.

Morgado, T.L.M., Branco, C.M. and Infante, V. (2008). "A failure study of housing of the gearboxes of series 2600 locomotives of the Portuguese Railway Company", *Engineering Failure Analysis*, 15(1-2): 154-164.

NASA (2007). *System Failure Case Studies: Derailed*. NASA Office of Safety and Mission Assurance, USA.

Neale, B.S. (2001). "Forensic engineering: the investigation of failures", *Proceedings of the 2nd International Conference on Forensic Engineering.* The Institution of Civil Engineers. Thomas Telford, UK.

Ni, Y.Q. (2015). "SHM-enriched high speed rail systems", *Proceedings of the 10th International Workshop on Structural Health Monitoring*. Stanford, USA.

Ni, Y.Q., Wang, X. and Ye, X.W. (2015). "Probabilistic fatigue assessment for high-speed train bogies using in-service monitoring data", *Proceedings of the 13th International Conference on Railway Engineering*. Edinburgh, UK.

Ni, Y.Q., Xia, H.W., Wong, K.Y. and Ko. J.M. (2011). "In-service condition assessment of bridge deck using long-term monitoring data of strain

response", *Journal of Bridge Engineering*, 17(6): 876-885.

Ni, Y.Q., Xia, Y., Lin, W., Chen, W.H. and Ko, J.M. (2012). "SHM benchmark for high-rise structures: a reduced-order finite element model and field measurement data", *Smart Structures and Systems*, 10(4-5): 411-426.

Ni, Y.Q., Ye, X.W. and Ko, J.M. (2011). "Modeling of stress spectrum using long-term monitoring data and finite mixture distributions", *Journal of Engineering Mechanics*, 138(2): 175-183.

Nicholas, T. and Zuiker, J.R. (1989). "On the use of the Goodman diagram for high cycle fatigue design", *International Journal of Fracture*, 80(2-3): 219-235.

Niedźwiecki, M., Michal, M. and Przemysław, P. (2016). "System identification based approach to dynamic weighing revisited", *Mechanical Systems and Signal Processing*, 80: 582-599.

Niemi, E. and Marquis, G. (2002). *Introduction to the Structural Stress Approach to Fatigue Analysis of Welded Components*, *IIw-Doc.XIII-1819-00.* International Institute of Welding, France.

NRAC. (2014). *The High-speed Railway in China*. National Railway Administration of China, Retrieved from http://www.nra.gov.cn/zggstlzt (in Chinese), China.

NRAC. (2016). *Mid-to-Long Term Railway Development Plan*. National Railway Administration of China, Retrieved from http://www.ndrc.gov.cn/zcfb

/zcfbghwb/201607/t20160720_813863.html (in Chinese), China.

O'Hagan, A. (2004). "Bayesian statistics: principles and benefits", *Frontis*, 3: 31-45.

Orowan, E. (1939). "Theory of the fatigue of metals", *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 79-106.

Pan, W.F., Hung, C.Y. and Chen, L.L. (1999). "Fatigue life estimation under multiaxial loadings", *International Journal of Fatigue*, 21(1): 3-10.

Paris, P.C., Gomez, M.P. and Anderson, W.E. (1961). "A comprehensive analytic theory of fatigue", *The Trend in Engineering*, 13(1): 9-14.

Park, B.H., Kim, N.P., Kim, J.S. and Lee, K.Y. (2006). "Optimum design of tilting bogie frame in consideration of fatigue strength and weight", *Vehicle System Dynamics*, 44(12): 887-901.

Petershagen, H, Fricke, W. and Massel, T. (1991). *Application of the Local Approach to the Fatigue Strength Assessment of Welded Structures in Ships*. *IIW Doc. XIII-1409-91*. International Institute of Welding, France.

Peterson, R.E. (1952). "Brittle fracture and fatigue in machinery", *Fatigue and Fracture of Metals*, 74.

Platt, L., Steger, M., Angeles, L.O.S. and The, D. (2014). "U.S. high-speed-rail effort gets a boost from China", *Executive Intelligence Review*, 48-49.

Pook, L. (2007). *Why Metal Fatigue Matters*. Springer, Netherlands.

Radaj, D. and Vormwald, M. (2013). *Advanced Methods of Fatigue Assessment*. Springer, Netherlands.

Rahman, M.M., Ariffin, A.K., Abdullah, S., Noor, M.M., Bakar, R.A. and Maleque, M.A. (2008). "Finite element based fatigue life prediction of cylinder head for two-stroke linear engine using stress-life approach", *Journal of Applied Sciences,* 8(19): 3316-3327.

Rankine W.J.M. (1842). "On the causes of the unexpected breakage of the journals of railway axles, and on the means of preventing such accidents by observing the law of continuity in their construction", *Institution of Civil Engineers, Minutes of Proceedings.* London, UK.

Reuters (2010). Retrieved from http://www.reuters.com/article/idustre61619q20100207s, London, UK.

Ritchie, R.O. (1988). "Mechanisms of fatigue crack propagation in metals, ceramics and composites: role of crack tip shielding", *Materials Science and Engineering: A*, 103(1): 15-28.

Roux, C., Lorang, X., Maitournam, H. and Nguyen-Tajan, M.L. (2014). "Fatigue design of railway wheels: a probabilistic approach", *Fatigue and Fracture of Engineering Materials and Structures*, 37: 1136-1145.

Rudlin, J., Muhammed, A. and Schneider, C. (2006). "Inspection reliability and periodicity for rail axle inspection", *Insight*, 48(6): 348-351.

Schijve, J. (1967). "Significance of fatigue cracks in micro-range and macro-

range", *ASTM Selected Technical Papers*, 415: 415-459.

Schijve, J. (2001). *Fatigue of Structures and Materials*. Springer, Netherlands.

Schijve, J. (2003). "Fatigue of structures and materials in the 20th century and the state of the art", *International Journal of Fatigue*, 25(8): 679-702.

Schutz, W. (1996). "A history of fatigue", *Engineering Fracture Mechanics*, 54(2): 263-300.

Schwarz, G. (1978). "Estimating the dimension of a model", *The Annals of Statistics*, 6(2): 461-464.

Scott, D.W. (2009). *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley and Sons, New Jersey, USA.

Seewald, F. (1932). "Messungen mit dem glasritzdehnungsschreiber der deutschen versuchsanstalt fiir luftfahrt", *Maschinenbau*, 10: 725-727.

Seki, K. (2012). "A study on intelligent trains to improve safety and reliability of operation", *Railway Technology Avalanche*, 39: 233.

Sendeckyj, G.P. (2001). "Constant life diagrams-a historical review", *Journal of Sound and Vibration,* 23(4): 347-353.

Shang, D.G., Yao, W.X. and Wang, D.J. (1998). "A new approach to the determination of fatigue crack initiation size", *International Journal of Fatigue*, 20(9): 683-687.

Shull, P.J. (2002). *Nondestructive Evaluation Theory, Techniques, and Applications*. Marcel Dekker, Inc., New York, USA.

Smith, R.A. (1984). “Thirty years of fatigue crack growth-an historical review”, *Fatigue Crack Growth: 30 Years of Progress*, 1-16.

Smith, R.A. (1990). “The Versailles railway accident of 1842 and the first research into metal fatigue”, *Fatigue,* 90: 2033-2041.

Smith, R.A. (2005). “Railway fatigue failures: An overview of a long standing problem”, *Materialwissenschaft Und Werkstofftechnik*, 36(11): 697-705.

Smith, R.A. and Zhou, J. (2014). “Background of recent developments of passenger railways in China, the UK and other European countries”, *Journal of Zhejiang University Science A*, 15(12): 925-935.

Sneed, W.H. and Smith, R.L. (1998). “On-board real-time railroad bearing defect detection and monitoring”, *Proceedings of the 1998 ASME/IEEE Joint Railroad Conference*. Philadelphia, USA.

Song, Y., Wu, P. and Jia, L. (2016). “Study of the fatigue testing of a car body underframe for a high-speed train”, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 230(6): 1614-1625.

Sonsino, C.M. (2007). “Fatigue testing under variable amplitude loading”, *International Journal of Fatigue*, 29(6): 1080-1089.

Stephens, R.I., Fatemi, A., Stephens, R.R. and Fuchs, H.O. (2000). *Metal Fatigue in Engineering*. John Wiley and Sons, New Jersey, USA.

Stone, J.V. (2013). *Bayes' Rule: A Tutorial Introduction to Bayesian Analysis.*

Sebtel Press, UK.

Stone, D.H., Kalay, S.F. and Tajaddini, A. (1992). "Statistical behaviour of wheel impact load detectors to various wheel defects", *Proceedings of the 10th International Wheelset Congress*. Sydney, Australia, 9-13.

Stratman, B., Liu, Y. and Mahadevan, S. (2007). "Structural health monitoring of railroad wheels using wheel impact load detectors", *Journal of Failure Analysis and Prevention*, 7(3): 218-225.

Strauss, A., Frangopol, D.M. and Kim, S. (2008). "Use of monitoring extreme data for the performance prediction of structures: Bayesian updating", *Engineering Structures*, 30(12): 3654-3666.

Sudret, B., Guédé, Z., Hornet, P., Stéphan, J.M. and Lemaire, M. (2003). "Probabilistic assessment of fatigue life including statistical uncertainties in the SN curve", *Proceedings of the 17th International Conference on Structural Mechanics in Reactor Technology*. Prague, Czech Republic.

Susmel, L. and Tovo, R. (2011). "Estimating fatigue damage under variable amplitude multiaxial fatigue loading", *Fatigue and Fracture of Engineering Materials and Structures*, 34(12): 1053-1077.

Susmel, L., Tovo, R. and Lazzarin, P. (2005). "The mean stress effect on the high-cycle fatigue strength from a multiaxial fatigue point of view", *International Journal of Fatigue,* 27(8): 928-943.

Tam, H.Y., Lee, T., Ho, S. and Haber, T. (2007). "Utilization of fiber optic Bragg

Grating sensing systems for health monitoring in railway applications", *Proceedings of the 6th International Workshop on Structural Health Monitoring,* Stanford University, Stanford, USA.

Tam, H.Y., Liu, S.Y., Guan, B.O., Chung, W.H., Chan, T.H. and Cheng, L.K. (2005). "Fiber Bragg grating sensors for structural and railway applications", *Advanced Sensor Systems and Applications, International Society for Optics and Photonics*, 2: 5634.

Tam, H.Y., Liu, S.Y., Ho, S.L. and Ho, T.K. (2011). "Fiber Bragg grating sensors for railway systems", *Fiber Bragg Grating Sensors: Recent Advancements, Industrial Applications and Market Exploitation*, 1:197-217.

Taylor, D. and Wang, G. (1999). "A critical distance theory which unifies the prediction of fatigue limits for large and small cracks and notches", *Proceedings of the 7th International Fatigue Congress*. Higher Education Press. Beijing, China, 579-584.

Thompson, N., Wadsworth, N. and Louat, N. (1956). "Xi. The origin of fatigue fracture in copper", *Philosophical Magazine*, 1(2): 113-126.

Timoshenko, S. (1954). "Stress concentration in the history of strength of materials", *Proceedings of the Society for Experimental Stress Analysis*. 12(1), 11-12.

Tong, G., Aiqun, L. and Jianhui, L. (2008). "Fatigue life prediction of welded

joints in orthotropic steel decks considering temperature effect and increasing traffic flow", *Structural Health Monitoring*, 7(3): 189-202.

Train, K.E. (2008). "EM algorithms for nonparametric estimation of mixing distributions", *Journal of Choice Modeling*, 1(1): 40-69.

Tunna, J. (1986). "Fatigue life prediction for Gaussian random loads at the design stage", *Fatigue and Fracture of Engineering Materials*, 9(3): 169-184.

Turner, J. D. and Hill, M. (1999). *Instrumentation for Engineers and Scientists*. Oxford University Press, UK.

Tyrell, D., Severson, K., Perlman, A.B., Brickle, B. and Vanlngen-Dunn, C. (2000). "Rail passenger equipment crashworthiness testing requirements and implementation", *Proceedings of the International Mechanical Engineering Congress and Exposition*. Orlando, USA.

UIC. (2011). *Wider Benefits of High-speed Rail Confirmed in New Studies.* UIC Publications, Paris.

UIC. (2015). *High Speed Rail Fast Track to Sustainable Mobility*. UIC Publications, Paris.

UIC. (2017). *High Speed Lines in the World*. UIC Publications, Paris.

UIC. 615-4 (1994). *Motive Power Units: Bogies and Running Gear-Bogie Frame Structure Strength Tests*. UIC Publications, Paris.

Ureña, J.M., Menerault, P. and Garmendia, M. (2009). "The high-speed rail challenge for big intermediate cities: a national, regional and local

perspective", *Cities*, 26(5): 266-279.

Vassilopoulos, A.P., Manshadi, B.D. and Keller, T. (2010). "Influence of the constant life diagram formulation on the fatigue life prediction of composite materials", *International Journal of Fatigue, 3*2(4): 659-669.

Vincent, L., Jean-Christophe, L.R. and Said, T. (2012). "On the high cycle fatigue behavior of a type 304L stainless steel at room temperature", *International Journal of Fatigue*, 38: 84-91.

Wan, H.P. and Ni, Y.Q. (2017). "Bayesian modeling approach for forecast of structural stress response using structural health monitoring data", *Journal of Structural Engineering*, accepted.

Wan, H.P. and Ren, W.X. (2016). "Stochastic model updating utilizing Bayesian approach and Gaussian process model", *Mechanical Systems and Signal Processing*, 70-71: 245-268.

Wang, J., Liu, X.Z., and Ni, Y.Q. "A Bayesian probabilistic approach for acoustic emission-based rail condition assessment", *Computer-Aided Civil and Infrastructure Engineering*, 0: 1-14.

Wang, X., Benjamin, Q., Jean-Daniel, C. and Jérôme, A. (2016). "Estimation of multiple sound sources with data and model uncertainties using the EM and evidential EM algorithms", *Mechanical Systems and Signal Processing*, 66: 159-177.

Watson, A.S. and Timmis, K. (2011). "A method of estimating railway axle stress

spectra", *Engineering Fracture Mechanics*, 78(5): 836-847.

Web News (2011). http://english.caixin.com/2011-08-25/100294816.html, China.

Wei, Z. and Dong, P. (2010). "Multiaxial fatigue life assessment of welded structures", *Engineering Fracture Mechanics*, 77(15): 3011-3021.

Weston, P., Li, P. and Ling, C. (2006). "Track and vehicle condition monitoring during normal operation using reduced sensor sets", *HKIE Transactions*, 13: 47-54.

Williams, J.K. (1965). "The airworthiness approach to structural fatigue, in fatigue design procedures", *Proceedings of the 4th lCAF-Symposium*. Munich, Turkey.

Wirsching, P. and Chen, Y. (1988). "Considerations of probability-based fatigue design for marine structures", *Marine Structures*, 1: 31-43.

Worden, K. and Dulieu-Barton, J.M. (2004). "An overview of intelligent fault detection in systems and structures", *Structural Health Monitoring*, 3(1): 85-98.

Xiao, Z.G. and Yamada, K. (2004). "A method of determining geometric stress for fatigue strength evaluation of steel welded joints", *International Journal of Fatigue*, 26(12): 1277-1293.

Yamamoto, M. (2013). *Research and Development of Fatigue Issues for Railway Steel Products and Future Prospects*. Nippon Steel and Sumitomo Metal Technical Report, Tokyo, Japan, 105.

Yang, A. (2013). “On the common conceptual and computational frameworks for multiscale modeling”, *Industrial and Engineering Chemistry Research*, 52(33): 11451-11462.

Ye, X.W. (2010). *Fatigue Reliability Assessment of Steel Bridges Instrumented with Structural Health Monitoring System*. PhD Thesis, The Hong Kong Polytechnic University, Hong Kong, China.

Yildirim, H.C. and Marquis, G.B. (2012). “Overview of fatigue data for high frequency mechanical impact treated welded joints”, *Welding in the World*, 56 (7-8): 82-96.

Yung, J.Y. and Lawrence, F.V. (1985). “Analytical and graphical aids for the fatigue design of weldments”, *Fatigue and Fracture of Engineering Materials and Structures*, 8(3): 223-241.

Zerbst, U., Mädler, K. and Hintze, H. (2005). “Fracture mechanics in railway applications: an overview”, *Engineering Fracture Mechanics*, 72(2): 163-194.

Zhai, W., Wang, K. and Cai, C. (2009). “Fundamentals of vehicle-track coupled dynamics”, *Vehicle System Dynamics*, 47(11): 1349-1376.

Zhang, L.H., Ni, Y.Q., Wang, X. and Liu, X.Z. (2016). “Monitoring of dynamic behavior and ride quality of an in-service high-speed train under different operation conditions”, *presented at the 1st International Workshop on Structural Health Monitoring for Railway System*. Qingdao, China.

Zhang, S. (2008). “Study on testing and establishment method for the load spectrum of bogie frame for high-speed trains”, *Science in China Series E: Technological Sciences*, 51(12): 2142-2151.

Zhao, G.T. (2016). “Health management technology for high-speed railway infrastructures”, *Workshop on New Monitoring Technology for Construction and Operation Safety of High-speed Railway*. Chengdu, China.

Zhao, W., Li, X. and Dong, P. (2017). *Theory and Method for Fatigue Design of Welded Structures*. China Machine Press, China.

Zhu, J. and Collette, M. (2016). “A Bayesian approach for shipboard lifetime wave load spectrum updating”, *Structure and Infrastructure Engineering*, 2479: 11-15.