RESEARCH REPORT

Using Integrated Reliability Analysis to Optimise Maintenance Strategies

A Bayesian Integrated Reliability Analysis of Locomotive Wheels

Jing Lin

LULEÅ UNIVERSITY OF TECHNOLOGY

LTU 2013:05

Järnvägstekniskt centrum, Luleå tekniska universitet
Besöksadress: Laboratorievägen, Porsön, Luleå. *Postadress:* S-971 87 Luleå
Telefon: 0920-49 10 00, *Fax:* 0920-49 19 35
Hemsida: http://www.jvtc.ltu.se

Using Integrated Reliability Analysis to Optimise Maintenance Strategies

A Bayesian Integrated Reliability Analysis of Locomotive Wheels

JVTC Projectnr 274

Jing (Janet) Lin

Division of Operation and Maintenance Engineering
JVTC - Luleå Railway Research Centre

Tryck: Universitetstryckeriet, Luleå, 2013

ISSN: 1402-1528
ISBN 978-91-7439-600-3 (tryckt)
ISBN 978-91-7439-601-0 (pdf)

Luleå 2013

www.ltu.se

Summary

1 Purpose

The goal of the research presented in this report is to propose, develop and test an integrated reliability analysis to optimise the maintenance strategies of the railway industry. This integrated analysis applies traditional statistics theories as well as Bayesian statistics using Markov Chain Monte Carlo (MCMC) methodologies. Using the Bayesian inference leads to greater flexibility because such analysis can simultaneously accommodate the following:

- Small sample data;
- Incomplete data set, including censored or truncated data;
- Complex operational environments.

In this report, an integrated procedure for Bayesian reliability inference using MCMC is applied to a number of case studies using locomotive wheel degradation data from Iron Ore Line (Malmbanan), Sweden. The research explores the impact of a locomotive wheel's installed position on its service lifetime and attempts to predict its reliability characteristics by using parametric models, non-parametric models, frailty factors, etc.

2 Preface

This research presented in this report was carried out at the Division of Operation and Maintenance Engineering at Luleå University of Technology between 2012 and 2013. The report contains five papers.

I would like to thank Luleå Railway Research Centre (Järnvägstekniskt Centrum, Sweden) at Luleå University of Technology and Swedish Transport Administration (Trafikverket) for initiating the research study and providing financial support.

I would like to express my gratitude to all of my colleagues. In particular, I would like to thank Professor Uday Kumar for being the project's supervisor and supporting me during my research at the Division of Operation and Maintenance Engineering. I would also like to thank Veronica Jägare and Cecilia Glover for their support.

I would like to express my gratitude to Matthias Asplund and Per-Olof Larsson-Kråik at Trafikverket/LTU, Thomas Nordmark, Ove Salmonsson and Hans-Erik Fredriksson at LKAB for their support and for discussing the locomotive wheels with me.

To my family

Janet (Jing) Lin, 05/05/2013, Luleå

3 Content

4 Abbreviations and Symbols

CMMS	Computerized Maintenance Management System
MCMC	Markov Chain Monte Carlo
MC error	Monte Carlo error
CPH	Cox Proportional Hazard
BOM	Bill of Material
RQ	Research Question
SD	Standard Deviation
HPD	Highest Posterior Distribution Density
MTTF	Mean Time To failure
BF	Bayesian Factor
BIC	Bayesian Information Criterion
DIC	Deviance Information Criterion
RCF	Rolling Contact Fatigue
KPI	Key Performance Indicator
LCC	Life Cycle Cost
i	Index
n	the number of system/units
x	covariate
β	the coefficient of the covariate
λ	the failure rate in exponential model
γ	the rate parameter in Weibull model
μ	the logarithmic mean in log-normal model
$R(\cdot)$	Reliability function
H	Inspection level
t	lifetime
$\lambda(t)$	Baseline hazard rate
$\Lambda(t)$	Cumulative hazard rate
κ	Rate parameter in Gamma distribution
ω	Frailty parameter
b	Piecewise interval
k	the number of the intervals in the piecewise constant model
Rd	Diameter of the wheel
Sd	Flange thickness
Rr	Radial runout
Rx	Axial runout

5 Introduction

This section presents the introduction of this research.

5.1 Background

The railway industry needs a more flexible decision-making strategy for maintenance optimisation, one that will accommodate the following three weaknesses in current reliability studies:

- **Small sample data for analysis.** The foundation of any classical statistics method is the law of large numbers where the data sample is expected to contain an infinite number of data. However, in real world situations, the number of data samples may be limited, or the samples may be small. For instance, in the railway industry, the real running data are often very limited for decision making, especially for new infrastructure. Therefore, prior information, including data collected through different sources, should be considered.
- **Incomplete data set, including censored or truncated data.** For economic reasons, we are seldom able to run infrastructures to failure to get the required data; we only know when their lifetimes exceed certain time points. In addition, it is common in the railway industry to replace or repair infrastructure before or after its lifetime endings, making it difficult to get the exact lifetime from real running data. Take the degradation data for locomotive wheels, for example. If the degradation data are less than the pre-specified diameter, the corresponding predicted lifetime is viewed as right-censored. Finally, the quality of the data obtained from Computerized Maintenance Management System (CMMS) may not be high, leaving the data set uncompleted.
- **Complex operational environments.** In the real world, systems run in different operational environments, which will significantly influence the system's reliability and maintenance strategies. For instance, in the railway industry, when replacement strategies are considered, in addition to wear, the failure of a technical system is affected by maintainability issues and the implementation of preventive measures. Because of the severe winter conditions in the Norrbotten region of Sweden (the area studied in this report), the influence of the weather should be seriously considered as well.

These factors, as well as the wear behaviour, should be considered as covariates for developing replacement or maintenance strategies. However, to date, no integrated reliability method considers all aspects simultaneously.

The recent proliferation of Markov Chain Monte Carlo (MCMC) approaches has led to the use of the Bayesian inference in a wide variety of fields, including behavioural science, finance, human health, process control, ecological risk assessment, and risk assessment of engineered systems (Kelly & Smith, 2009, and the references therein). Discussions of MCMC related methodologies and their applications in Bayesian Statistics now appear throughout the literature (Congdon, 2001; 2003). For the most part, studies in reliability analysis focus on the following topics and their cross-applications: 1) hierarchical reliability models (Robinson, 2001; Johnson, et al., 2003; Wilson, 2008; Lin, et al., 2013a); 2) complex system reliability analysis (Tont, et al., 2010; Lin, 2008; Lin, et al. 2011); 3) faulty tree analysis (Hamada, et al., 2004; Graves, et al., 2007); 4) accelerated failure models (Kuo & Mallick, 1997; Walker & Mallick, 1999; Hanson & Johnson, 2004; Ghosh & Ghosal, 2006; Komarek & Lesaffre, 2009); 5) reliability growth models (Li, et al., 2002; Tao, et al., 2004); 6) Masked system reliability (Kuo & Yang, 2000); 7) software reliability engineering (Ilkka, 2006; Tamura, et al. 2011); 8) Reliability benchmark problems (Schueller & Pradlwarter, 2007; Au, et al., 2007). Most of the literature emphasises the model's development; no studies offer a full framework to accommodate academic research and engineering applications seeking to implement modern computational-based Bayesian approaches, especially in the area of reliability.

To fill the gap and to facilitate MCMC applications from a reliability perspective, this research proposes an integrated procedure for the Bayesian inference to consider the defects in the railway industry's reliability studies.

The service life of a train wheel can be significantly reduced due to failure or damage, leading to excessive cost and accelerated deterioration, a point which has received considerable attention in recent literature. In order to monitor the performance of wheels and make replacements in a timely fashion, the railway industry uses both preventive and predictive maintenance. By predicting the wear of train wheels (Johansson & Andersson, 2005; Braghin et al., 2006; Tassini et al., 2010), fatigue (Bernasconi et al., 2005; Liu, et al., 2008), tribological aspects (Clayton, 1996), and failures (Yang & Letourneau, 2005), the industry can design strategies for different types of preventative maintenance (re-profiling, lubrication, etc.) for various periods (days, months, seasons, running distance, etc.). Software dedicated to predicting wear rate has also been proposed (Pombo et al., 2010). Finally, condition monitoring data have been studied with a view to increasing the wheels' lifetime (Skarlatos et al., 2004; Donato et al., 2006; Stratman et al., 2007; Palo, 2012).

One common preventive maintenance strategy (used in the case study) is re-profiling wheels after they run a certain distance. Re-profiling affects the wheel's diameter; once the diameter is reduced to a pre-specified length, the wheel is replaced by a new one. Seeking to optimise this maintenance strategy, researchers have examined wheel degradation data to determine wheel reliability and failure distribution (Freitas et al., 2009, 2010; and the references therein). However, these studies cannot solve the combined problem of small data samples and incomplete datasets while simultaneously considering the influence of several covariates. For example, to avoid the potential influence of wheel location, Freitas et al. (2009, 2010) only consider those on the left side of specified axle and on certain specified cars, but point out that "the degradation of a given wheel might be associated with its position on a given car". Yang and Letourneau (2005) suggest that certain attributes, including a wheel's installed position (right or left), might influence its wear rate, but they do not provide case studies. Palo et al. (2012) conclude that "different wheel positions in a bogie show significantly different force signatures". In a recent seminar in Sweden (Kiruna, April 2012), experts from Norway illustrated their new findings that in a given topography, the wheels installed on the right and the left sides of a locomotive experience different force. Unfortunately, they did not include signal charts derived from condition monitoring tools, nor did they consider the influence of wheel position on degradation.

To address the above issues, in this project, we explore the influence of locomotive wheels' positioning on reliability using a proposed integrated Bayesian procedure with MCMC methods. The research first considers parametric models, including the Bayesian Exponential Regression Model, Bayesian Weibull Regression Model and Bayesian Log-normal Regression Model. Second, since semi-parametric Bayesian methods offer a more general modelling strategy that contains fewer assumptions (Ibrahim et al., 2001), we adopt the piecewise constant hazard model to establish the distribution of the locomotive wheels' lifetime. Most reliability studies are implemented under the assumption that individual lifetimes are independent identified distributed (i.i.d), but at times, Cox proportional hazard (CPH) models cannot be used because of the dependence of data within a group. Therefore, in this project we also considered frailty models, in which the data are conditionally independent. Through a gamma shared frailty, the dependence within subgroups can be considered an unknown and unobservable risk factor (or explanatory variable) of the hazard function, allowing us to determine reliability more flexibly for the wheel. A similar frailty model is studied later in this project with a Weibull frailty factor. To utilise both degradation data and the re-profiling performance of the wheels, statistics on the wheels are compared in the following categories: 1) degradation analysis using a Weibull frailty model; 2) work orders for re-profiling; 3) the performance of re-profiling parameter; and 4) wear rates.

5.2 Description of Data

In this project, all the case studies come from Sweden's Iron Ore Line. The data focus on the heavy haul cargo trains' locomotive wheels and were collected by LKAB\MTAB from October 2010 to January 2012. This section gives background information on the Iron Ore Line. It also introduces the degradation data and the re-profiling parameters for the locomotive wheels being studied.

5.2.1 Iron Ore Line (Malmbanan)

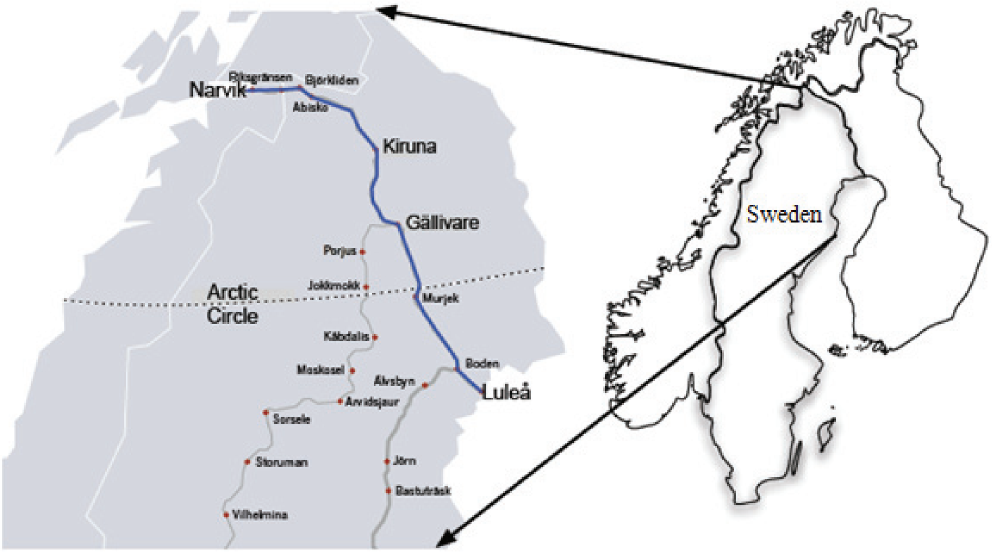


Fig.1 Geographical location of Iron Ore Line (Malmbanan)

The Iron Ore Line (Malmbanan) is the only existing heavy haul line in Europe; it stretches 473 kilometres and has been in operation since 1903. As Fig. 1 shows, it is mainly used to transport iron ore and pellets from the mines in Kiruna (also Malmberget, close to Kiruna, in Sweden) to Narvik Harbour (Norway) in the northwest and Luleå Harbour (Sweden) in the southeast. The track section on the Swedish side is owned by the Swedish government and managed by Trafikverket (Swedish Transport Administration), while the iron ore freight trains are owned and managed by the freight operator (a Swedish company). Each freight train consists of two IORE locomotives accompanied by 68 wagons with a maximum length of 750 metres and a total train weight of 8500 metric tonnes. The trains operate in harsh conditions, including snow in the winter and extreme temperatures ranging from - 40 °C to + 25 °C. Because carrying iron ore results in high axle loads and there is a high demand for a constant flow of ore/pellets, the track and wagons must be monitored and maintained on a regular basis. The condition of the locomotive wheel profile is one of the most important aspects to consider.

5.2.2 Degradation data and re-profiling parameters

In this study, we use the degradation data from selected heavy haul cargo locomotives, collected from October 2010 to January 2012. For each locomotive, see Fig.2, there are two bogies (incl., Bogie I, Bogie II); and each bogie contains six wheels. The installed position of a wheel on a particular locomotive is specified by the bogie number (I, II-number of bogies on the locomotive), an axel number (1, 2, 3-number of axels for each bogie) and the position of the axle (right or left) where each wheel is mounted.

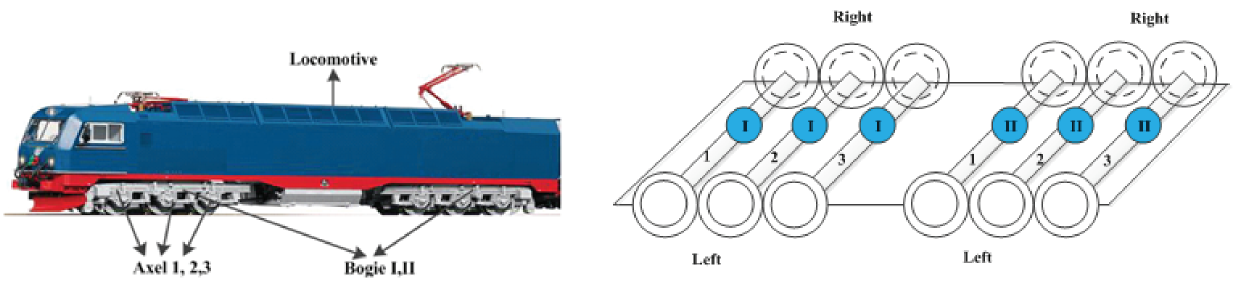


Fig.2 Wheel positions specified in this study

The diameter of a new locomotive wheel in this study is about 1250 mm. Following the current maintenance strategy, a wheel's diameter is measured after it runs a certain distance. If it is reduced to 1150 mm, the wheel set is replaced by a new one. Otherwise, it is re-profiled (see Fig.3). Therefore, a threshold level for failure, denoted as y_0, is defined as 100 mm (y_0 = 1250 mm -1150 mm). The wheel's failure condition is assumed to be reached if the diameter reaches y_0. The dataset includes the diameters of all locomotive wheels at a given inspection time, the total running distances corresponding to their "mean time between re-profiling", and the wheels' bill of material (BOM) data, from which we can determine their positions.

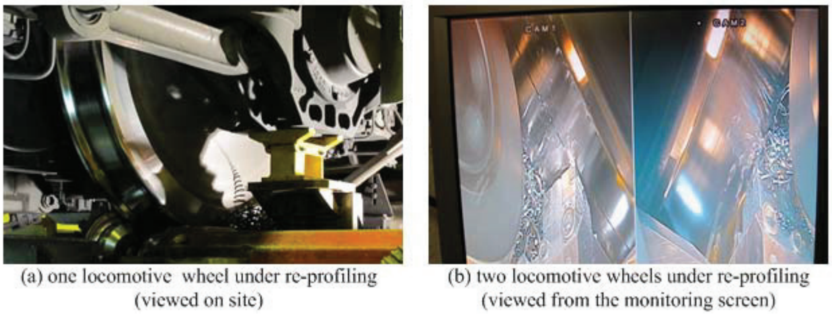


Fig.3 Locomotive wheels on-site re-profiling

During the re-profiling process, the re-profiling parameters include but are not limited to: 1) the diameters of the wheels; 2) the flange thickness; 3) the radial run-out; 4) the lateral run-out.

5.3 Purpose and objectives

Following the above discussion, the overall goal of the research presented in this report is to propose, develop and test an integrated reliability analysis to optimise the railway industry's maintenance strategies. This integrated analysis applies both traditional statistics theories and Bayesian statistics using Markov Chain Monte Carlo (MCMC) methodologies. Using the Bayesian inference adds flexibility to the analysis as it can accommodate the following aspects simultaneously:

- Small sample data;
- Incomplete data set, including censored or truncated data;
- Complex operational environments.

The research also explores the impact of a locomotive wheel's installed position on its service lifetime and attempts to predict its reliability characteristics by using parametric models, non-parametric models, frailty factors, etc.

5.4 Research questions and appended papers

To fulfil the purpose and objectives of this project, we ask the following research questions:

- **RQ1** What is the integrated procedure for Bayesian Reliability Inference using MCMC methodologies?
- **RQ2** How can the Bayesian parametric models be applied to integrated reliability analysis and maintenance strategies optimisation?
- **RQ3** How can the Bayesian non-parametric models be applied to integrated reliability analysis and maintenance strategies optimisation?
- **RQ4** How can the frailty factors be applied to integrated reliability analysis and maintenance strategies optimisation?
- **RQ5** How can the other analyses be applied to integrated reliability analysis?

These research questions are answered by the five appended papers:

- Lin J. *An Integrated Procedure for Bayesian Reliability Inference using Markov Chain Monte Carlo Methods*. Submitted to Journal. (**Paper I**)
- Lin J, Asplund M, Parida A. *Reliability Analysis for Degradation of Locomotive Wheels using Parametric Bayesian Approach.* Accepted by Quality and Reliability Engineering International. Will be published in 2013. DOI: 10.1002/qre.1518 (**Paper II**)
- Lin J, Asplund M. *Bayesian Semi-parametric Analysis for Locomotive Wheel Degradation using Gamma Frailties*. Submitted to Journal. Under Review. (**Paper III**)
- Lin J, Asplund M. *A Comparison Study for Locomotive Wheels' Reliability Assessment using the Weibull Frailty Model*. Revised Manuscript has been submitted to Journal of Rail and Rapid and Transit, in April, 2013. **(Paper IV)**
- Lin J, Asplund M, Parida A. *Bayesian Parametric Analysis for Reliability Study of Locomotive Wheels*. Conference Proceedings. The 59th Annual Reliability and Maintainability Symposium (RAMS® 2013). January 28-31, Orlando, FL, USA. **(Paper V)**

Each paper makes its own contribution to the research questions. Table 1 shows the relationship between the papers and the research questions.

Table.1 Relationship between the appended papers and research questions

Research questions	RQ1	RQ2	RQ3	RQ4	RQ5
Paper I	+				
Paper II	+	+			
Paper III	+		+	+	
Paper IV	+				+
Paper V	+	+			

5.5 Scope and limitations

The scope and limitations of this project include:

- The research is applicable to the reliability analysis of any other lifetime data.
- The proposed Bayesian reliability inference procedure has been applied to all case studies in this research.
- The data of the case studies focus on the heavy haul cargo trains' locomotive wheels and were collected by LKAB from October 2010 to January 2012.
- Other assumptions in each study are discussed in the five appended papers, separately.

6 Results and discussions

This section discusses the research findings for each question.

6.1 An Integrated Procedure for Bayesian Reliability Inference using MCMC

RQ1 *What is the integrated procedure for Bayesian Reliability Inference using MCMC methodologies?*

The first research question is answered in Paper I; findings are applied to Papers II, III, IV, and V.

The recent proliferation of MCMC approaches has led to the use of the Bayesian inference in a wide variety of fields. To facilitate MCMC applications, this research has proposed an integrated procedure for Bayesian inference using MCMC methods, from a reliability perspective. The proposed procedure uses the Bayesian reliability inference to determine system (or unit) reliability and failure distribution, and to support the optimisation of maintenance strategies, etc.

The general procedure begins with the collection of reliability data (see Fig.4.). These are the observed values of a physical process, such as various “lifetime data”. The data may be subject to uncertainties, such as imprecise measurement, censoring, truncated information, and interpretation errors. Reliability data are found in the “current data set”; they contain original data and include the evaluation, manipulation, and/or organisation of data samples. At a higher level in the collection of data, a wide variety of “historical information” can be obtained, including the results of inspecting and integrating this “information”, thereby adding to “prior knowledge”. The final level is reliability inference, which is the process of making a conclusion based on “posterior results”.

Using the above definitions, we propose an integrated procedure which constructs a full framework for the standardised process of Bayesian reliability inference. As shown in Fig.4, the procedure is composed of a continuous improvement process including four stages (Plan, Do, Study, Action) and the following 11 sequential steps which are discussed in more detail in Paper I: 1) data preparation; 2) prior inspection and integration; 3) prior selection; 4) model selection; 5) posterior sampling; 6) MCMC convergence diagnostic; 7) Monte Carlo error diagnostic; 8) model improvement; 9) model comparison; 10) inference making; 11) data updating and inference improvement.

- Step 1: Data preparation. The original data sets for “history information” and “current data” related to reliability studies need to be acquired, evaluated, and merged. In this way, “history information” can be transferred to “prior knowledge”, and “current data” can become “reliability data” in later steps.
- Step 2: Prior inspection and integration. During this step, “prior knowledge” receives a second and more extensive treatment, including a reliability consistency check, a credence test, and a multi-source integration. This step improves prior reliability data.

- Step 3: Prior selection. This step uses the results achieved in Step 2 to determine the model's form and parameters, for instance, selecting informative or non-informative priors, or unknown parameters and their distributed forms.
- Step 4: Model selection. This step determines a reliability model (parametric or non-parametric), selecting from n candidates for the studied system/units. It considers both "reliability data" and the inspection, integration, and selection of priors to implement the i^{th} ($i = 1, \cdots i+1, \cdots n$).
- Step 5: Posterior sampling. In this step, we determine a sampling method (for instance, Gibbs sampling, Metropolis-Hastings sampling, etc.) to implement MCMC simulation for the model's posterior calculations.
- Step 6: MCMC convergence diagnostic. In this step, we check whether the Markov chains have reached convergence. If they have, we move on to the next step; if they have not, we return to Step 5 and re-determine the iteration times of posterior sampling or re-choose the sampling methods; if the results still cannot be satisfied, we return to Steps 3 and 4 and re-determine the prior selection and model selection.
- Step 7: Monte Carlo error diagnostic. We need to decide if the Monte Carlo error is small enough to be accepted in this step. As discussed in Step 6, if it is accepted, we go on to the next step; if it is not, we return to Step 5 and re-decide the iteration times of the posterior sampling or re-choose the sampling methods; if the results still cannot be accepted, we go back to Steps 3 and 4 and recalculate the prior selection and model selection.
- Step 8: Model improvement. Here, we choose the $i+1^{th}$ candidate model and restart from Step 4.
- Step 9: Model comparison. After implementing n candidate models, we need to: 1) compare the posterior results to determine the most suitable model; or 2) adopt the average posterior estimations (using the Bayesian model average or the MCMC model average) as the final results.
- Step 10: Inference making. After achieving the posterior results in Step 9, we can perform Bayesian reliability inference to determine system (or unit) reliability, find the failure distribution, and optimise maintenance strategies, etc.;
- Step 11: Data updating and inference improvement. Along with the passage of time, new "current data" can be obtained, relegating "previous" inference results to "historical data". By updating "reliability data" and "prior knowledge", and restarting at Step 1, we can improve the reliability inference.

In summary, by using this step-by-step method, we can create a continuous improvement process for the Bayesian reliability inference.

Note that Steps 1, 2, and 3 are assigned to the "Plan" stage when data for MCMC implementation are prepared. In addition, a part of Steps 1, 2, and 3 refers to the elicitation of prior knowledge. Steps 4 and 5 are both assigned to the "Do" stage, where the MCMC sampling is carried out. Steps 6 to 9 are treated as the "Study" stage; in these steps, the sampling results are checked and compared; in addition, knowledge is accumulated and improved upon by implementing various candidate reliability models. The "Action" stage consists of Steps 10 and 11; at this point, a continuously improved loop can be obtained. In other words, by implementing the step-by-step procedure, we can accumulate and gradually update prior knowledge. Equally, posterior results will be improved upon and become increasingly robust, thereby improving the accuracy of the inference results.

Also note that Paper I has focused on six steps and their relationship to MCMC inference implementation: 1) prior elicitation; 2) model construction; 3) posterior sampling; 4) MCMC convergence diagnostic; 5) Monte Carlo error diagnostic; 6) model comparison.

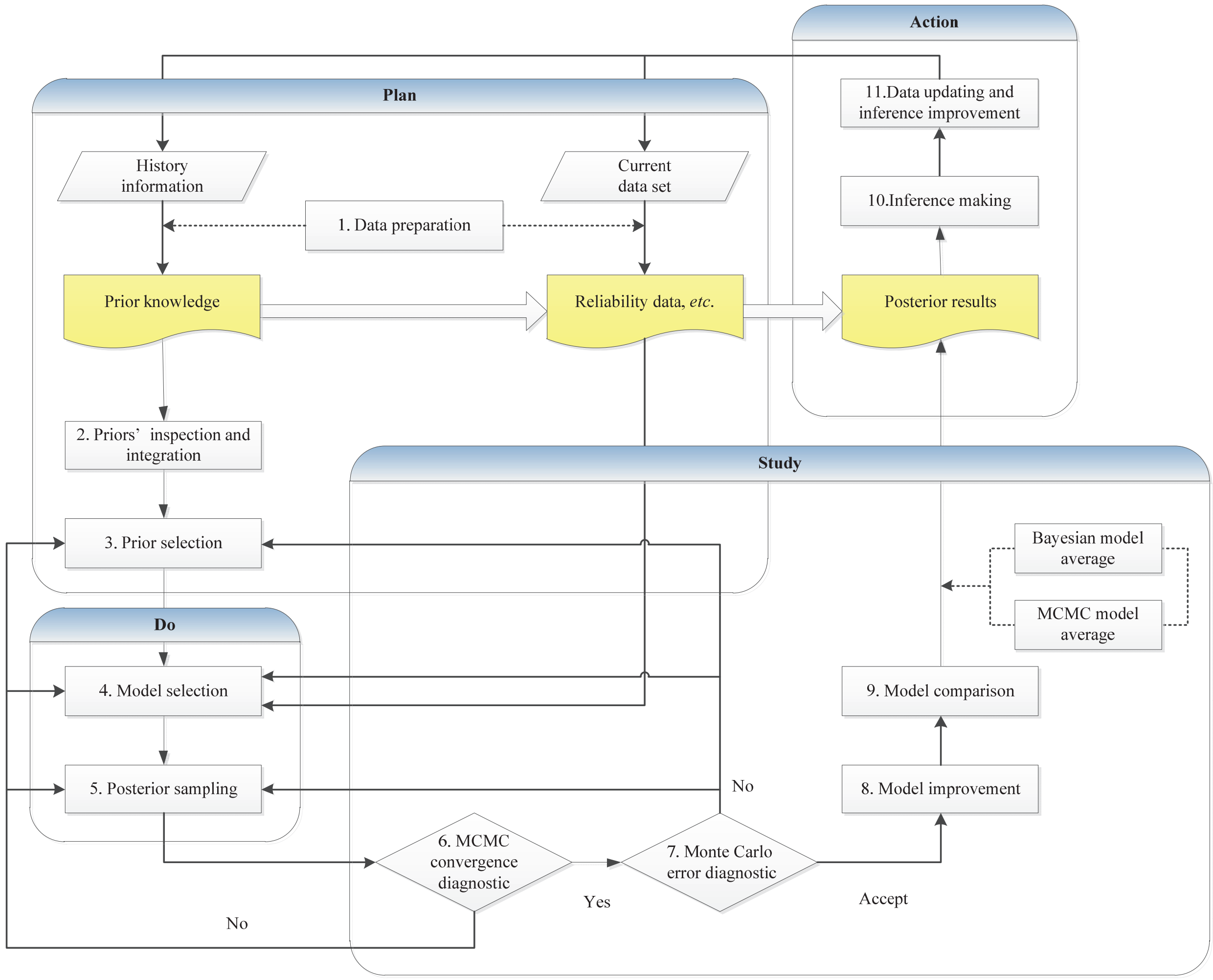


Fig.4 An Integrated Procedure for Bayesian Reliability Inference via MCMC

6.2 Bayesian Parametric Analysis for Locomotive Wheel Degradation

RQ2 *How can the Bayesian parametric models be applied in integrated reliability analysis and maintenance strategies optimisation?*

The second research question is addressed in Paper II and Paper V.

The service life of a railroad wheel can be significantly reduced due to failure or damage, leading to excessive cost and accelerated deterioration. Damage data show that a major proportion of wheel damage stems from degradation. To monitor the performance of wheels and make replacements before adverse effects occur, the railway industry uses both preventive and predictive maintenance. In one common preventive maintenance policy in the Swedish railway company studied, a wheel's diameter is measured after running a certain distance. If it is reduced to a pre-specified height, the wheel is replaced. Otherwise, it is re-profiled or other maintenance strategies are adopted.

This study undertakes a reliability study using a Bayesian survival analysis framework (Ibrahim et al, 2001; Congdon, 2001 & 2003; Jing, 2008) to explore the impact of the wheel's installed position on its service lifetime and to predict its reliability characteristics. The Bayesian Exponential Regression Model, Bayesian Weibull Regression Model and Bayesian Log-normal Regression Model are used to analyse the lifetime of locomotive wheels using degradation data and taking into account the position of the wheel.

In this study, the wheel position is described by three different discrete covariates: the bogie, the axle and the side of the locomotive where the wheel is mounted. By introducing the covariate $\mathbf{x}_i$'s linear function $\mathbf{x}_i'\boldsymbol{\beta}$, these three parameter models are constructed depending on the failure rate λ_i in the exponential model, the log of the rate parameter $\ln(\gamma_i)$ in the Weibull model and the logarithmic mean μ_i in the log-normal models. Following the convergence diagnostics (i.e., checking dynamic traces in Markov chains, time series, and Gelman-Rubin-Statistics, and comparing the MC error with Standard Deviation (SD)), we consider the following posterior distribution summaries (shown in Tables 2, 3 and 4), for our models (Bayesian Exponential Regression Model, Bayesian Weibull Regression Model, and Bayesian Log-normal Regression Model), including the parameters' posterior distribution mean, standard deviation, Monte Carlo error, and 95% HPD (highest posterior distribution density) interval.

Table.2 Posterior Distribution Summaries for Exponential Regression Model

Parameter	Mean	SD	MC error	95 % HPD Interval
β_0	-5.862	0.7355	0.02299	(-7.366,-4.452)
β_1	-0.07207	0.3005	0.007269	(-0.6672,0.5104)
β_2	-0.03219	0.1858	0.003797	(-0.3889,0.3325)
β_3	-0.0124	0.2973	0.00726	(-0.5954,0.5787)

Table.3 Posterior Distribution Summaries for Weibull Regression Model

Parameter	Mean	SD	MC error	95 % HPD Interval
α	10.08	0.9674	0.05559	(8.234,11.76)
β_0	-60.47	5.977	0.3434	(-71.01,-49.16)
β_1	-0.07775	0.306	0.008339	(-0.6845,0.5156)
β_2	-0.146	0.2231	0.005801	(-0.5878,0.2856)
β_3	-0.05026	0.2982	0.007143	(-0.6356,0.5324)

Table.4 Posterior Distribution Summaries for Log-normal Regression Model

Parameter	Mean	SD	MC error	95 % HPD Interval
β_0	5.864	0.05341	0.001622	(5.76,5.97)
β_1	0.06733	0.02174	5.042E-4	(0.02492,0.1103)
β_2	0.02077	0.01373	2.765E-4	(-0.006291,0.04781)
β_3	0.001102	0.02175	5.007E-4	(-0.0412,0.04444)
τ	187.5	39.84	0.3067	(118.3,273.5)

Accordingly, the locomotive wheels' reliability functions can be written as:

- Bayesian Exponential Regression Model:

$$R(t_i \mid \mathbf{X}) = \exp\left[-\exp(-5.862 - 0.072x_1 - 0.032x_2 - 0.012x_3) \times t_i\right]$$

- Bayesian Weibull Regression Model:

$$R(t_i \mid \mathbf{X}) = \exp\left[-\exp(-60.47 - 0.078x_1 - 0.146x_2 - 0.050x_3) \times t_i^{10.08}\right]$$

- Bayesian Log-normal Regression Model:

$$R(t_i \mid \mathbf{X}) = 1 - \Phi\left[\frac{\ln(t_i) - (5.864 + 0.067x_1 + 0.02x_2 + 0.001x_3)}{(187.5)^{-1/2}}\right]$$

Obviously, other reliability characteristics of lifetime distribution, including Mean Time to Failure (MTTF), can also be determined.

Other findings on model comparison, maintenance predictions, and maintenance inspection levels are briefly discussed in the following subsections.

6.2.1 Model Comparison

Traditional technologies for model comparison consider two main aspects: the model's measure of fit and its complexity. Usually, improving the model's complexity will improve its fit. For instance, by considering more unknown parameters, the SD and MC error of the model's posterior can be reduced and the model's measure of fit can be improved. However, the complexity of the model will be increased at the same time. Therefore, most model comparison studies focus on the balance between them. When comparing Bayesian models, both the Bayesian Factor (BF) and Bayesian Information Criterion (BIC) can be used, but for complex Bayesian hierarchical models, this becomes more difficult. Spiegelhalter et al. (2002) have proposed the Deviance Information Criterion (DIC), which utilises the model's deviance to evaluate its measure of fit and the effective number of parameters to evaluate its complexity.

We calculate the DIC values for the above three Bayesian parametric models separately, as shown in Table 6.2.1.

Table.5 DIC Summaries

Model	$\overline{D}(\theta)$	$D(\overline{\theta})$	p_d	DIC
Exponential Regression	648.98	645.03	3.95	652.93
Weibull Regression	472.22	467.39	4.83	477.05
Log-normal Regression	442.03	436.87	5.16	447.19

Based on Celeux et al. (2006) and the related discussions above, we choose the model with the lowest DIC value. When DIC<5, the difference among models can be ignored. Our results show that the DIC for the Log-normal Regression Model is the lowest (447.19); following the arguments above, it is more suitable than the other two. Using the same model, when we analyse other locomotives' wheels running under similar conditions, we reach similar conclusions. However, when we compare the DIC values for the Weibull Regression Model and the Exponential Regression Model, 477.05 and 652.93, respectively, we see that the performance of the Weibull Regression Model is close to the Log-normal Regression Model, making it a suitable choice in certain specified situations.

6.2.2 Maintenance Predictions

All Bayesian parameter models reach a common conclusion: the installation positions influence the wheels' lifetimes. In addition, considering the covariates' coefficients in our case study, we find the following: 1) the lifetime of the wheel installed in the second bogie is longer than that of the wheel installed in the first one; 2) the wheel installed in the third axel has a longer lifetime than that installed in the second axel, and the wheel in the second axel has a longer lifetime than the one in the first axel; 3) the right side wheel's lifetime is shorter than the left side. (Researchers from Norwegian National Rail Administration cited previously concur with this. Using condition monitoring methods on train wheels operating on the same route, they found that the wheel forces on the right and the left sides can be different, even for wheels in the same axel.). Possible causes for the differences include the influence of the topographical complexity, and the position of the locomotive's centre of gravity.

The three Bayesian parametric regression models presented here are all effective in their Markov chain convergence and other diagnostic tools; see, for example, Spiegelhalter et al. (2002) who compare the computation process, including checking Markov chains' dynamic traces, time series and Gelman-Rubin-Statistics, and comparing the MC error with Standard Deviation (SD). However, we prefer the Bayesian Lognormal Regression Model because of its DIC values. The prediction of the locomotive wheels' MTTF, following this model, appears in Table.6.

It should be pointed out that the 95% HPD interval in the Bayesian Lognormal Regression Model for β_2 and β_3 includes 0 (Table.4). This means that although the positioning does have an influence, in some instances, the impact on the wheel's service lifetime is not significantly strong. In our case, the bogies have more impact on service lifetime than axels or sides. Given this knowledge, we can deal with such covariates better in our future research.

Table.6 MTTF statistics based on Bayesian Lognormal Regression Model

Bogie	Axel	Side	μ_i	MTTF($\times 10^3$ km)
I (x_1 =1)	1 (x_2 =1)	Right (x_3 =1)	5.9532	387.03
		Left (x_3 =2)	5.9543	387.46
	2 (x_2 =2)	Right (x_3 =1)	5.9740	395.16
		Left (x_3 =2)	5.9751	395.60
	3 (x_2 =3)	Right (x_3 =1)	5.9947	403.43
		Left (x_3 =2)	5.9958	403.87
II (x_1 =2)	1 (x_2 =1)	Right (x_3 =1)	6.0205	413.97
		Left (x_3 =2)	6.0216	414.43
	2 (x_2 =2)	Right (x_3 =1)	6.0413	422.67
		Left (x_3 =2)	6.0424	423.14
	3 (x_2 =3)	Right(x_3 =1)	6.0621	431.56
		Left (x_3 =2)	6.0632	432.03

6.2.3 Maintenance Inspection Level

According to the assumptions in Paper II, the maintenance inspection level H_2 (where $0 \le H_2 \le H_1$) determines how many lifetime data are "right-censored". Obviously, the higher the maintenance inspection level, the more data are considered "right-censored" and vice versa. For instance, in Fig.5, we show a higher maintenance inspection level (80 mm) and a lower one (20 mm). We denote the area between H_1 and H_2 as Zone I, and the area between H_2 and zero degradation level as Zone II. Therefore, based on the likelihood functions discussed Paper II, the MTTF statistics which are achieved from the higher H_2 (the left picture in Fig.5, where H_2 = 80 mm) will be higher than those obtained from the lower H_2 (the right picture, where H_2 =20 mm), because fewer degradation data are considered right-censored. In other words, the results achieved from the former are more "optimistic", and the results obtained from the latter are more "pessimistic". An extreme condition is to suppose H_2 =0 mm.

For this reason, we can get an interval prediction between "optimistic" and "pessimistic" with different maintenance inspection levels which actually reflect the different risk confidence levels.

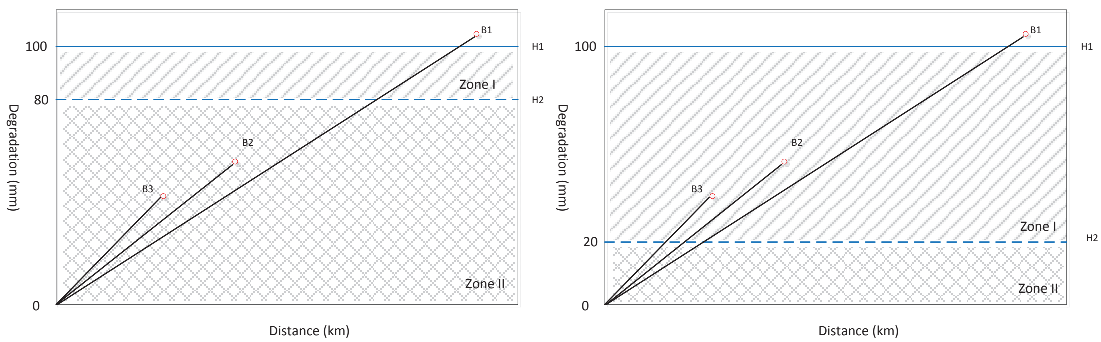


Fig.5 Maintenance Inspection Level with Zone I and Zone II

6.3 Bayesian Non-parametric Analysis for Locomotive Wheel Degradation using Frailties

RQ3 *How can the Bayesian non-parametric models be applied in integrated reliability analysis and maintenance strategies optimisation?*

RQ4 *How can the frailty factors be applied in integrated reliability analysis and maintenance strategies optimisation?*

The third question is studied in Paper III; the fourth is studied in Papers III and IV.

Since semi-parametric Bayesian methods offer a more general modelling strategy that contains fewer assumptions (Ibrahim et al., 2001), in Paper III, we adopt the piecewise constant hazard model to establish the distribution of the locomotive wheels' lifetime. The applied hazard function is sometimes referred to as a piecewise exponential model; it is convenient because it can accommodate various shapes of the baseline hazard over the intervals.

In addition, as mentioned in Paper III, most reliability studies are implemented under the assumption that individual lifetimes are independent identified distributed (i.i.d). However, sometimes CPH models cannot be used because of the dependence of data within a group. For instance, because they have the same operating conditions, the wheels mounted on a particular locomotive may be dependent. In a different context, some data may come from multiple records which actually belong to the wheels installed in the same position but on another locomotive. Modelling dependence in multivariate survival data has received considerable attention, especially in cases where the datasets may come from subjects of the same group which are related to each other (Sahu et al., 1997; Aslanidou et al., 1998). A key development in modelling such data is to consider frailty models, in which the data are conditionally independent. When frailties are considered, the dependence within subgroups can be considered an unknown and unobservable risk factor (or explanatory variable) of the hazard function.

In Paper III, we consider a gamma shared frailty, first discussed by Clayton (1978) and later developed by Sahu et al. (1997), to explore the unobserved covariates' influence on the wheels on the same locomotive. In this study, the gamma shared frailties $\boldsymbol{\omega}_i$ are used to explore the influence of unobserved covariates within the same locomotive. By introducing covariate $\mathbf{x}_i$'s linear function $\mathbf{x}_i^{'}\boldsymbol{\beta}$, the influence of the bogie in which a wheel is installed can be taken into account.

The results of the case study suggest that the wheels' lifetimes differ according to where they are installed on the locomotive. The wheel installed in the second bogie has a longer lifetime than the one in the first bogie. The differences could be influenced by the real running situation (e.g. topography), and the locomotive's centre of gravity. The gamma frailties help with exploring the unobserved covariates and thus improve the model's precision. Results also indicate a close positive relationship between the wheels mounted on the same locomotive; the heterogeneity between locomotives is significant as well. We can determine the wheel's reliability characteristics, including the baseline hazard rate $\lambda(t)$, reliability $R(t)$, and cumulative hazard rate $\Lambda(t)$, etc.

The results indicate the existence of change points. As Fig.6, Fig.7 and Fig.8 show, wheel reliability declines sharply at the fourth piecewise interval, while at the fifth piecewise interval, the cumulative hazard increases dramatically. The results allow us to evaluate and optimise wheel replacement and maintenance strategies (including the re-profiling interval, inspection interval, lubrication interval, depth and optimal sequence of re-profiling, and so on). Finally, the approach discussed in this study can be applied to cargo train wheels or to other technical problems (e.g. other industries, other components).

Statistics summaries (Table.7) in this study include the parameters' posterior distribution mean, SD, MC error, and the 95% HPD interval. In Table.7, $\beta_1 > 0$ means that the wheels mounted in the first bogie (as $x = 1$) have a shorter lifetime than those in the second (as $x = 2$). However, the influence could possibly be reduced as more data are obtained in the future, because the 95% HPD interval includes 0 point. Because $\kappa > 0.5$, there is a positive relationship between the wheels mounted on the same locomotive; in addition, the heterogeneity among the locomotives is significant. Meanwhile, $\omega_1 < 1$ suggests that the predictive lifetimes for those wheels mounted on the first locomotive are longer than if the frailties are not considered; in fact, $\omega_2 > 1$ indicates the opposite conclusion.

Table.7 Posterior Distribution Summaries

Parameter	mean	SD	MC error	95% HPD Interval
β_0	-10.39	2.888	0.2622	(-16.61, -4.79)
β_1	0.3293	0.4927	0.02016	(-0.661, 1.271)
κ	0.563	0.269	0.01038	(0.1879, 1.225)
ω_1	0.1441	0.1374	0.004822	(0.01192, 0.5258)
ω_2	1.866	1.016	0.03628	(0.3846, 4.308)
b_1	0.1361	1.595	0.1037	(-3.196, 3.364)
b_2	0.758	2.182	0.1672	(-3.7, 5.248)
b_3	1.94	2.514	0.2105	(-3.126, 7.342)
b_4	4.447	2.668	0.2389	(-0.5652, 10.48)
b_5	6.342	2.684	0.2415	(1.126, 12.29)
b_6	8.159	2.724	0.2417	(2.843, 14.15)

Baseline hazard rate statistics based on the above results are shown in Table.8 and Fig.6. At the fourth piecewise interval, the wheels' baseline hazard rate increases dramatically.

Table.8 Baseline Hazard Rate Statistics

Piecewise Intervals($\times$ 1000km)	1	2	3	4	5	6
	(0, 60]	(60, 120]	(120, 180]	(180, 240]	(240, 300]	(300, 360]
λ_k	1.15	2.13	6.96	85.37	567.93	3494.69

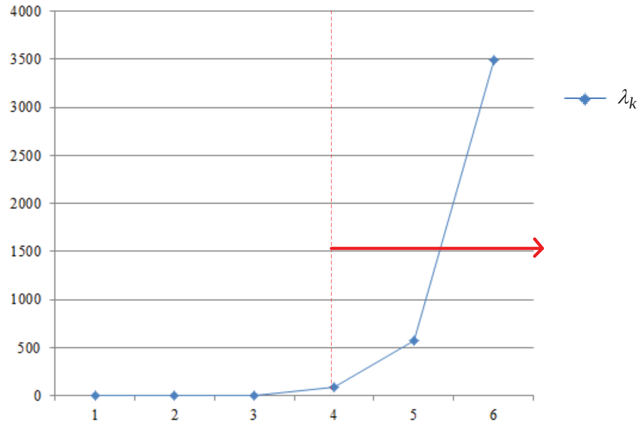


Fig.6 Plot of Baseline Hazard Rate

By considering the random effects resulting from the natural variability (explained by covariates) and from the unobserved random effects within the same group (explained by frailties), we can determine other reliability characteristics of lifetime distribution. The statistics on reliability $R(t)$ and cumulative hazard rate $\Lambda(t)$ for the two wheels mounted in different bogies are listed in Table.9, Fig.7 and Fig.8.

Fig.6 and Fig.7 show frailties between Locomotive 1 and Locomotive 2. In addition, for these locomotives, the wheels mounted in the first bogie ($x=1$) have lower reliability and a higher cumulative hazard rate than those mounted in the second one ($x=2$).

Table.9 Reliability and Cumulative hazard statistics

Distance (1000 km)	Reliability $R(t)$				Cumulative hazard $\Lambda(t)$			
	Locomotive 1		Locomotive 2		Locomotive 1		Locomotive 2	
	Bogie I	Bogie II	Bogie I	Bogie II	Bogie I	Bogie II	Bogie I	Bogie II
60	0.999577	0.999412	0.994534	0.99241	0.000184	0.000256	0.00238	0.003309
120	0.998425	0.997811	0.97979	0.97202	0.000685	0.000952	0.008867	0.012325
180	0.992318	0.989338	0.90496	0.870393	0.003349	0.004655	0.04337	0.060285
240	0.881485	0.839169	0.195241	0.103252	0.054785	0.076151	0.709428	0.986101
300	0.350289	0.232678	1.26E-06	6.31E-09	0.455574	0.633245	5.899379	8.200106
360	0.000433	2.11E-05	2.75E-44	2.82E-61	3.363977	4.67591	43.56128	60.54995

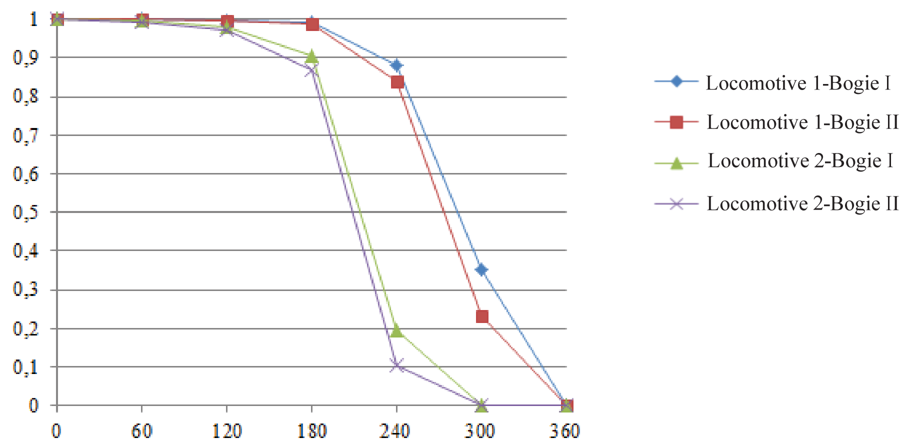


Fig.7 Plot of the reliabilities for Locomotive 1 and Locomotive 2

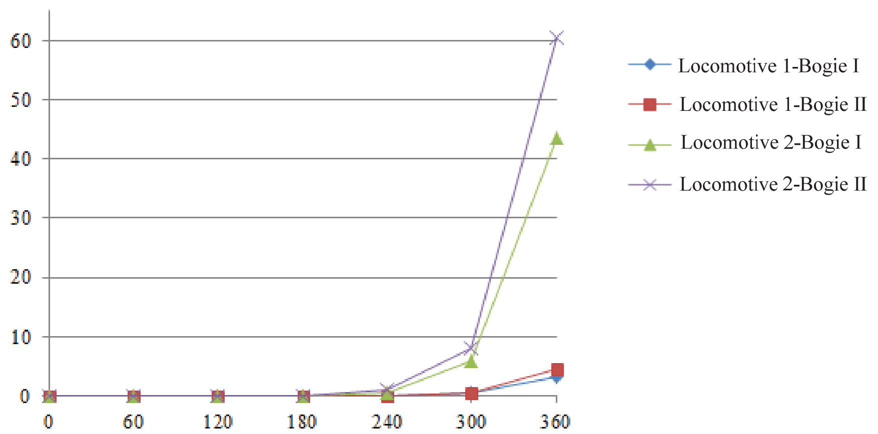


Fig.8 Plot of the Cumulative hazard for Locomotive 1 and Locomotive 2

Fig.7 and Fig.8 show change points in the wheels. For example, the reliability declines sharply at the fourth piecewise interval, and at the fifth piecewise interval, the cumulative hazard increases dramatically.

In what follows, several other finds are briefly discussed. These include: lifetime prediction and replacement optimisation, preventative maintenance optimisation, and re-profiling optimisation.

First, determining reliability characteristics distributed over the wheels' lifetime could be used to optimise replacement strategies. The results could also support related predictions for spares inventory.

Second, the change points (Fig.6, Fig.7, Fig.8) appearing in the fourth and fifth piecewise interval (from 180 000 to 300 000 kilometres) indicate that after running about 180 000 kilometres, the locomotive wheel has a high risk of failure. Rolling contact fatigue (RCF) problems could start at the fifth interval (after 240 000 kilometres). Therefore, special attention should be paid if the wheels have run longer than these change points. Finally, because re-profiling may leave cracks over time and reduce the wheel's lifetime, cracks should be checked after re-profiling to improve the lifetime.

Third, the wheels installed in the first bogie should be given more attention during maintenance. Especially when the wheels are re-profiled, they should be checked starting with the first bogie to avoid duplication of effort. Note that in the studied company, the inspecting sequences are random; this means that the first checked wheel could belong in the second bogie. After the second checked wheel is lathed or re-profiled, if the diameter is less than predicted, the first checked wheel might need to be lathed or re-profiled again. Therefore, starting with the wheel installed in the first bogie could improve maintenance effectiveness.

Last but not least, the frailties between locomotives could be caused by the different operating environments (e.g., climate, topography, and track geometry), configuration of the suspension, status of the bogies or spring systems, operation speeds and applied loads. Specific operating conditions should be considered when designing maintenance strategies because even if the locomotives and wheel types are the same, the lifetimes and operating performance could differ.

Note that in section 3 of Paper IV, we consider a gamma frailty but in a Weibull frailty model. More details appear in Paper IV. The difference between the frailty models used in Papers III and IV is that the former is studied in a non-parametric formulate (the baseline hazard rate is piecewise), while the latter is studied in a parametric formulate (the baseline hazard rate is Weibull).

6.4 Comparison Study of the Reliability Assessment of Locomotive Wheels

RQ5 *How the other analyses can be applied in integrated reliability analysis?*

The fifth research question is discusses in Paper IV.

The service life of different railroad wheels can vary greatly. Take a Swedish railway company, for example. For the wheels of its 26 locomotives, statistics show that from 2010 to 2011, the longest mean time between re-profiling was around 59 000 kilometres and the shortest was about 31 000 kilometres. The large difference can be attributed to the non-heterogeneous nature of the wheels; each differs according to its installed position, operating conditions, re-profiling characteristics, etc.

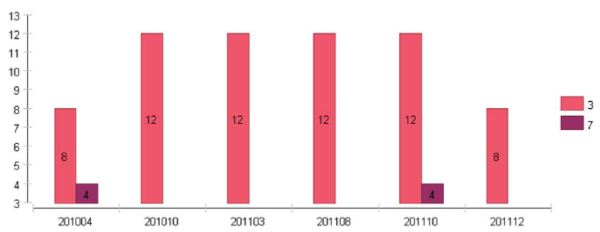
In this study, we compare the wheels on two selected locomotives to explore some of these differences. We propose integrating reliability assessment data with both degradation data and re-profiling performance data. Our case study compares: 1) degradation analysis using a Weibull frailty model; 2) work orders for re-profiling; 3) the performance of re-profiling parameter; and 4) wear rates.

6.4.1 Results from comparing re-profiling work orders

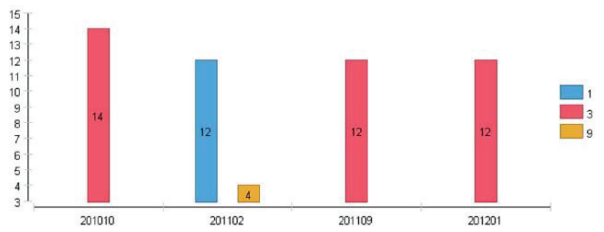
This section shows some results from comparing the work orders for wheel re-profiling by date (denoted as “by date” in Fig.9) and the corresponding bogies’ total number of kilometres in operation (denoted as “by kilometres” in Fig.10), separately.

In Fig.9, the work order statistics for re-profiling are listed by date. The colour and the number of the bar represent the type of work order reported in the system. For instance, number 1 (blue) means the reason for re-profiling is a high flange; number 3 (red) represents the RCF problem; number 7 (purple) means the re-profiling is due to the dimension difference between wheels in a bogie; number 9 (yellow) denotes a thick flange.

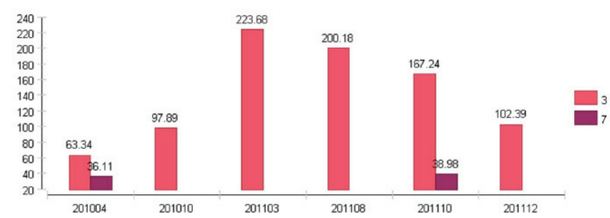
The work orders have 14 categories for re-profiling: high flange, thin flange, RCF, unbalanced wheel, QR measurements, out-of-round wheel, dimension difference in between wheels in same bogie, vibrations, thick flange, cracks, remarks from measurement of the wheel by Miniprof, other defects, to plant for re-profiling, and hollowware. These categories are determined by the operator and are listed in Appendix A of Paper IV. Take Fig.9 (a) for example. By April 2010, the wheels of Locomotive 1 have been re-profiled 12. Eight times it was related to category 3 (RCF problem), and four times it was in category 7 (the dimension difference between wheels in a bogie).



(a)Work order numbers for Locomotive 1

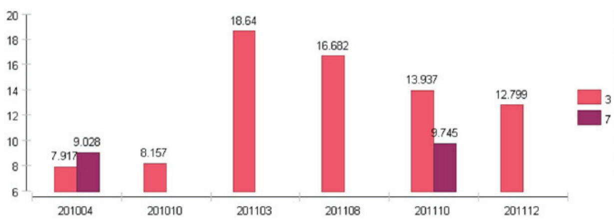


(b) Work order numbers for Locomotive 2

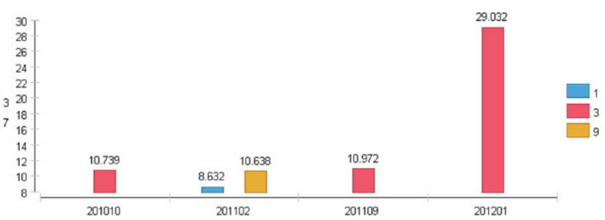


(c)Total re-profiling for Locomotive 1 /mm

(d) Total re-profiling for Locomotive 2 /mm



(e)Average re-profiling for Locomotive 1 /mm

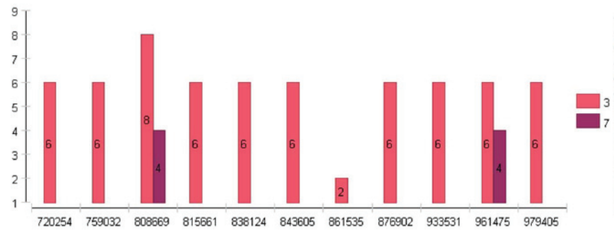


(f) Average re-profiling for Locomotive 2 /mm

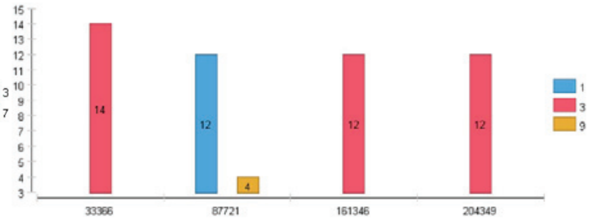
Fig.9 Work orders statistics on re-profiling by date

In Fig. 9 and Fig.10, the figures on the left side provide the statistics for locomotive 1, while those on the right are for locomotive 2. Note that in Fig.10, the work order statistics on re-profiling are listed by the corresponding bogies' total number of kilometres in operation on the reported date. In Fig.7 (b), the wheels have run 87721 kilometres and been re-profiled 16 times, 12 times due to category 1 (high flange) and 4 times due to category 9 (thick flange).

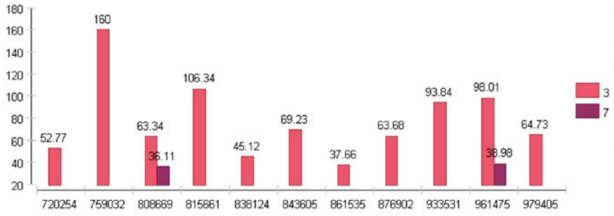
It should be pointed out that since October 2010, new wheels have been mounted on both locomotives. However, the selected work orders are from the beginning of 2010; therefore, more re-profiling has been done on locomotive 1.



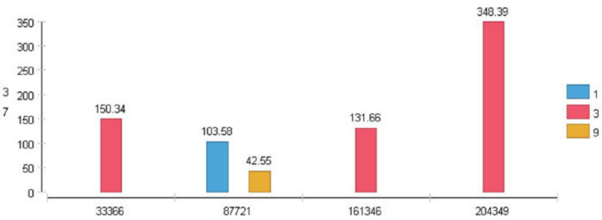
(a)Work order numbers for Locomotive 1



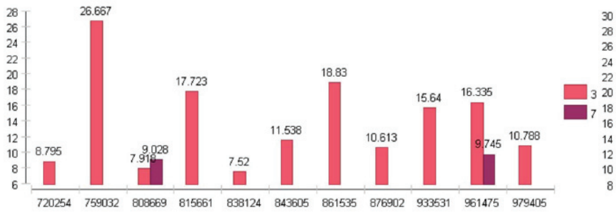
(b) Work order numbers for Locomotive 2



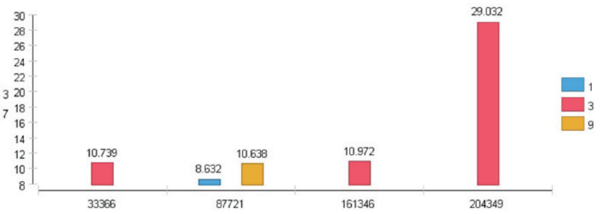
(c)Total re-profiling for Locomotive 1



(d) Total re-profiling for Locomotive 2



(e)Average re-profiling for Locomotive 1



(f) Average re-profiling for Locomotive 2

Fig.10 Work orders statistics on re-profiling by kilometres

For locomotive 1, there are two failure modes: RCF and dimensional differences for wheels in the same bogie. The number of re-profiling work orders due to RCF is 64; the number due to dimensional differences for wheels in the same bogie is 8. Locomotive 2 shows three failure modes, high flange, RCF and thick flange. Again, the dominant failure mode is RCF with 38 re-profilings, followed by high flange with 12 re-profilings and thin flange with 4; see Fig.9 (b). Figs. 9 (c) and (d) show the amount of material removed at each re-profiling for all wheels. Even here, the RCF failure dominates with more material lost in re-profiling. Figs. 9 (e) and (f) show the mean cut deep for each re-profiling. The RCF failure mode has deeper cuts than other modes; the high flange failure mode has the smallest mean cut depth.

Fig. 10 shows the same information but uses the global traveling distance in kilometres (km). It should be pointed out that for Locomotive 1, Fig.10 has more bars on the left hand side because the axels have been changed and the recorded kilometres are different.

Generally speaking, RCF is the main type of work order for both locomotives. What should also be pointed out is that in the work order statistics, natural wear and the amount of re-profiling are considered simultaneously. Yet the trends in the amount of re-profiling are different. For instance, for locomotive 1, there is a decreasing trend for new wheels, while locomotive 2 shows an increasing trend.

During this investigation, we have discovered a number of problems in the work orders. For example, some reported data cannot be recognised (e.g., some wheels are apparently re-profiled twice on one date; some reported wheel diameters after re-profiling are larger than before re-profiling). We suggest applying related Key Performance Indicators (KPIs) to monitor the re-profiling work and the wheel performance in the future.

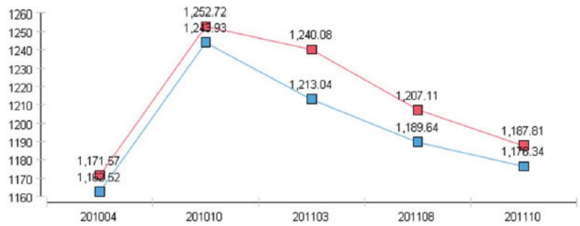
6.4.2 Results from comparing re-profiling parameters

This section provides some results from comparing the re-profiling parameters (the statistics before and after each re-profiling), including the diameter of the wheel (denoted as Rd), the flange thickness (denoted as Sd), the radial runout (denoted as Rr), and the axial runout (denoted as Rx).

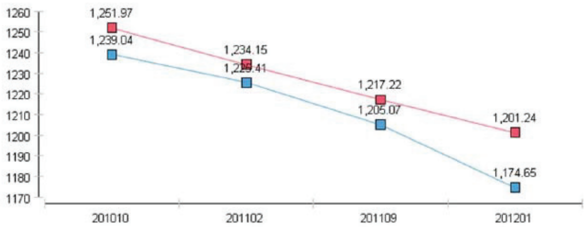
6.4.2.1 Assessment of re-profiling parameters (Rd)

Starting in this section, we only include statistics by re-profiling date. In addition, due to the similarities of the wheels installed in the same bogie, we only list statistics for the chosen wheel within each bogie. The red line represents the statistics obtained before re-profiling; the blue line represents statistics after re-profiling.

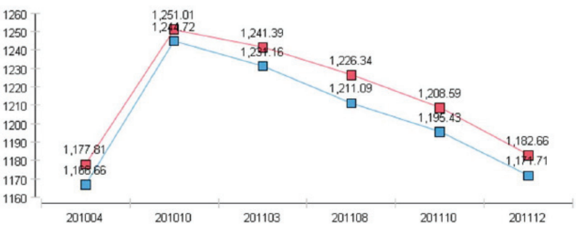
Fig.11 shows locomotive 1 on the left hand side and locomotive 2 on the right; for the graphs, the y-axle is the wheel diameter and the x-axle is the re-profiling date. For locomotive 1, the graphs start with the last re-profiling of an old wheel; step two is the first re-profiling with new wheels.



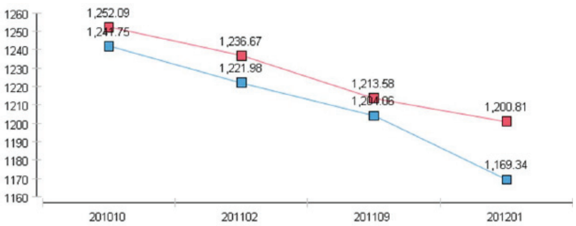
(a)Rd statistics for Locomotive 1 (I1H)



(b) Rd statistics for Locomotive 2 (I1H)



(c)Rd statistics for Locomotive 1 (II1H)



(d) Rd statistics for Locomotive 2 (II1H)

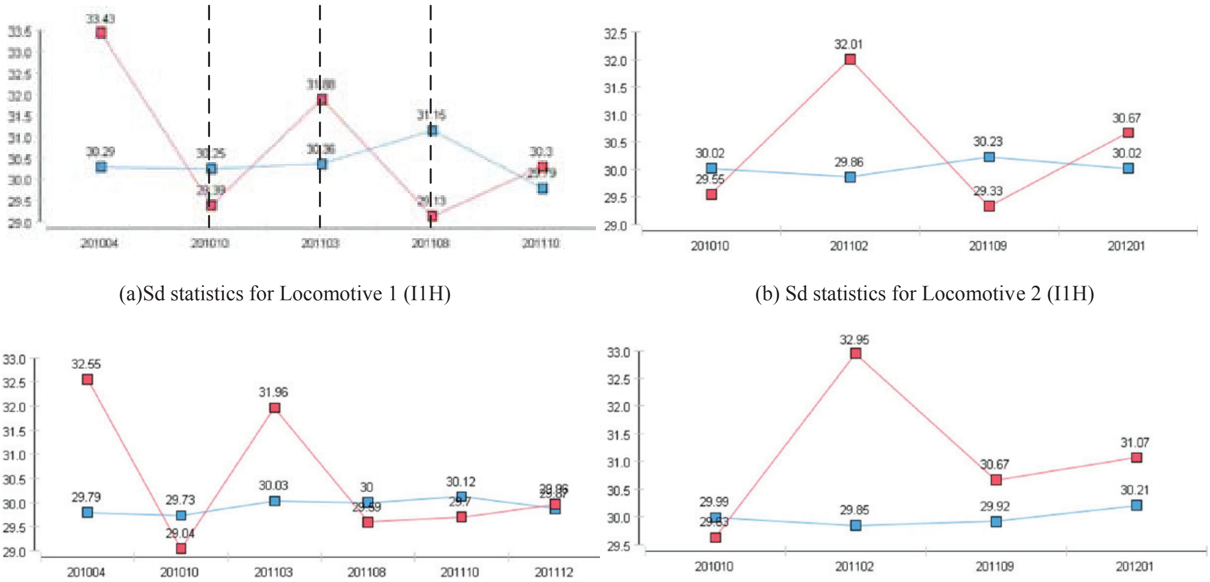
Fig.11 Rd statistics by date (before and after re-profiling): one example (I1H & II1H)

The wheels installed in the same bogie show similar trends in before and after re-profiling (denoted as Δ Rd). Δ Rd is decreasing for locomotive 1 and increasing for locomotive 2.

6.4.2.2 Assessment of re-profiling parameter (Sd)

Fig.12 shows some statistics of the Sd for the selected wheels.

Locomotive 1 is represented on the left hand side, with locomotive 2 on the rightt. For both, the flange thickness increases during winter and decreases in summer; this phenomenon is especially pronounced for locomotive 1 and the first bogie and first axle; see the dotted lines in Fig.12a.



(a)Sd statistics for Locomotive 1 (I1H)

(b) Sd statistics for Locomotive 2 (I1H)

(c)Sd statistics for Locomotive 1 (II1H)

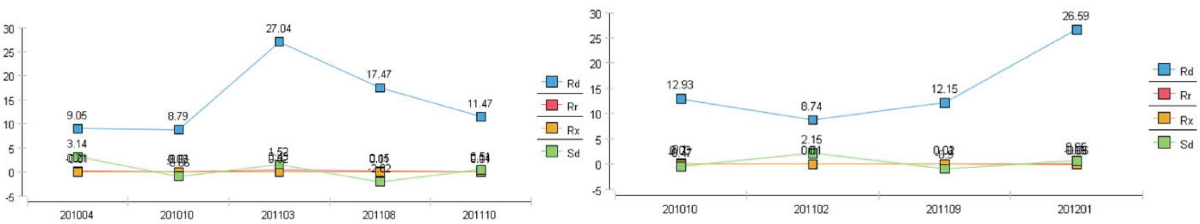
(d) Sd statistics for Locomotive 2 (II1H)

Fig.12 Sd statistics by date (before and after re-profiling): one example (I1H & II1H)

Like the Rd statistics, the Sd statistics for the wheels installed in the same bogie are quite similar. The "after" statistics (in blue) are stable. The "before" statistics (in red) are gradually becoming stable, which means the gap (denoted as Δ Sd) is decreasing. Note that if we check the before and after statistics in different seasons, we see that the red line is decreasing in the summer and increasing in the winter; see Fig.12 (a).

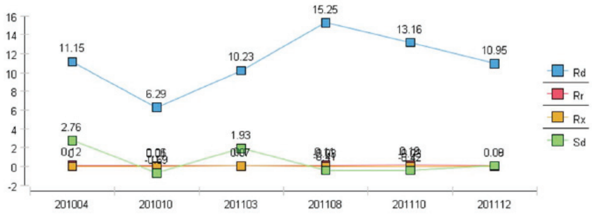
6.4.2.3 Assessment of re-profiling parameter (ΔRd, ΔSd, ΔRr, ΔRx)

In this section, we simultaneously consider the gaps of the four parameters discussed above: ΔRd (blue), ΔSd (green), ΔRr (red), and ΔRx (yellow).

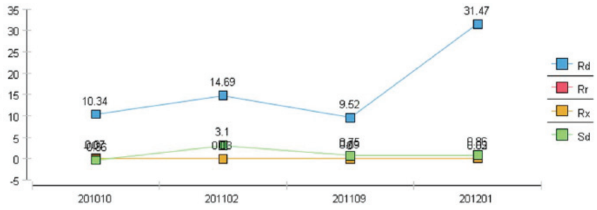


(a) Gap statistics for Locomotive 1 (I1H)

(b) Gap statistics for Locomotive 2 (I1H)



(c) Gap statistics for Locomotive 1 (II1H)



(d) Gap statistics for Locomotive 2 (II1H)

Fig.13 gap statistics by date (before and after re-profiling): one example (I1H & II1H)

As discussed above, the statistics for the wheels installed in the same bogie are quite similar. Among these four parameters, the changing of ΔRd is the most obvious, with ΔSd coming second. The changing of ΔRr and ΔRx are random and the amount is quite small compared to the first two parameters. Therefore, we suggest applying the first two parameters to monitor the wheels' re-profiling performance in the future.

6.4.3 Results from comparing wear rate

This section shows some results from comparing the wheels' wear rates (Tables.10 to 13). More details appear in Appendix B (Table B.1- B.20) of Paper IV.

Table 10 shows locomotive 1, bogie 1 and the first axle on the right side; Table 11 shows locomotive 1, bogie 2 and the first axle on the right side; see Fig. 2 for the position of the bogies and axles. The number of re-profiling work orders is different between bogies: bogie 1 has 4 and bogie 2 has 5. The reason for the difference may be that bogie 1 was changed after the fourth re-profiling. The re-profiling at times 1 to 4 was done at the same time for both bogies, extending over 12 months.

As for locomotive 1, Table 10 shows that it has been running for 123.351 km; the mean distance between re-profiling is 41.117 km. The distance after the last re-profiling for bogie 2 was only 17.930 km, less than half of the average distance for re-profiling numbers 1 to 4; see Table 11. Tables 10 and 11 also show the diameter of the wheel before and after re-profiling and the amount of material removed at each re-profiling. The mean amount of material removed during re-profiling for bogie 1 is 16.193 mm and for bogie 2, 11.176 mm. Remarkably, the amount of re-profiling for bogie 2, step 2 is 27.04 mm, much more than the others; as noted above, the mean is 16.193 mm. If we compare natural wear with artificial wear, the former is between 15 mm and 20% of the total wear. In addition, the total wear rate for locomotive 2, bogie 1, is 0.619 mm/1000 km; for bogie 2, it is 0.393 mm/1000km.

As mentioned, locomotive 1 and locomotive 2 have the same operating conditions (see Fig. 10 for the comparison), but the figures in Tables 12 and 13 show different results. Table 12 shows locomotive 2, the first bogie, the first axle, and the right hand side wheel; Table 13 shows the second bogie, the first axle, and the right hand side wheel. This locomotive has been re-profiled 4 times in 15 months; the mean distance between re-profiling is 56.990 km. The mean amount of material removed for re-profiling for bogie 1 is 15.10 mm; for bogie 2 it is 16.51 mm. The last re-profiling for the first bogie removed 26.59 mm and for the second bogie 31.47 mm. Finally, the total wear rate for locomotive 2, bogie 1, is 0.452 mm/1000 km and for bogie 2, 0.484 mm/1000km

Table.10 Statistics for wear rate: an example (locomotive 1, I1H)

Locomotive	1	Position	I1H		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201103	201108	201110	12 months
Reported kilometres /1000km	720.254	759.032	815.661	843.605	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	123.351
Diameters (before)/mm	1252.72	1240.08	1207.11	1187.81	/
Diameters (after)/mm	1243.93	1213.04	1189.64	1176.34	/
Re-profiling Amount/mm	8.79	27.04	17.47	11.47	64.77
Natural Wear/mm	0	3.85	5.93	1.83	11.61
Total Wear/mm	8.79	30.89	23.4	13.3	76.38
Re-profiling Amount %	1	0.875	0.747	0.862	0.848
Natural Wear %	0	0.125	0.253	0.138	0.152
WearRate re-profiling	/	0.697	0.308	0.41	0.525
WearRate Natural	/	0.099	0.105	0.065	0.094
WearRate Total	/	0.797	0.413	0.476	0.619

Table.11 Statistics for wear rate: an example (locomotive 1, II1H)

Locomotive	1	Position	II1H			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres	838.124	876.902	933.531	961.475	979.405	/
Absolute kilometres	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1251.01	1241.39	1226.34	1208.59	1182.66	/
Diameters (after)/mm	1244.72	1231.16	1211.09	1195.43	1171.71	/
Re-profiling Amount/mm	6.29	10.23	15.25	13.16	10.95	44.93
Natural Wear/mm	0	3.33	4.82	2.5	12.77	10.65
Total Wear/mm	6.29	13.56	20.07	15.66	23.72	55.58
Re-profiling Amount %	1	0.754	0.76	0.84	0.462	0.808
Natural Wear %	0	0.246	0.24	0.16	0.538	0.192
WearRate re-profiling	/	0.264	0.269	0.471	0.611	0.318
WearRate Natural	/	0.086	0.085	0.089	0.712	0.075
WearRate Total	/	0.35	0.354	0.56	1.323	0.393

Table.12 Statistics for wear rate: an example (locomotive 2, I1H)

Locomotive	2	Position	11H		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1251.97	1234.15	1217.22	1201.24	/
Diameters (after)/mm	1239.04	1225.41	1205.07	1174.65	/
Re-profiling Amount/mm	12.93	8.74	12.15	26.59	60.41
Natural Wear/mm	0	4.89	8.19	3.83	16.91
Total Wear/mm	12.93	13.63	20.34	30.42	77.32
Re-profiling Amount %	1	0.641	0.597	0.874	0.781
Natural Wear %	0	0.359	0.403	0.126	0.219
WearRate re-profiling	/	0.161	0.165	0.618	0.353
WearRate Natural	/	0.09	0.111	0.089	0.099
WearRate Total	/	0.251	0.276	0.707	0.452

Table.13 Statistics for wear rate: an example (locomotive 2, II1H)

Locomotive	2	Position	21H		Total/Average
Number	1	2	3	4	4 times
Date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1252.09	1236.67	1213.58	1200.81	/
Diameters (after)/mm	1241.75	1221.98	1204.06	1169.34	/
re-profiling Amount/mm	10.34	14.69	9.52	31.47	66.02
Natural Wear/mm	0	5.08	8.4	3.25	16.73
Total Wear/mm	10.34	19.77	17.92	34.72	82.75
re-profiling Amount %	1	0.743	0.531	0.906	0.798
Natural Wear %	0	0.257	0.469	0.094	0.202
WearSpeed re-profiling	/	0.27	0.129	0.732	0.386
WearSpeed Natural	/	0.093	0.114	0.076	0.098
WearSpeed Total	/	0.364	0.243	0.807	0.484

By comparing the interval of the re-profiling date, we can simply divide each re-profiling episode into seasons (for instance, the summer and warmer times, the winter and cooler times).

In Table.14, we list the statistics for the WearRate_total of all the wheels for the two locomotives. The mean wear rates are 0.516 mm/1000km for locomotive 1 and 0.480 mm/1000km for locomotive 2; in other words, locomotive 1 has a 75% higher wear rate. Axles 1, 2 and 5 have 11.6 % higher wear rate than axles 3, 4 and 6.

Table.14 Statistics for total wear rates

	WearRate_total											
	11H	11V	12H	12V	13H	13V	21H	21V	22H	22V	23H	23V
Locomotive 1	0.619	0.607	0.614	0.605	0.542	0.533	0.393	0.404	0.467	0.467	0.467	0.472
Locomotive 2	0.452	0.439	0.448	0.448	0.449	0.448	0.484	0.482	0.568	0.575	0.487	0.476

By comparing the above parameters of the wheels installed in different positions on the locomotives, as well as the results shown in Appendix B (Table B.1 – B.20) of Paper IV, we reach the following additional conclusions:

- the average wear rate of the wheels on locomotive 1 is greater than for locomotive 2;
- the natural wear is about 10% ~ 25 % of the total wear; the re-profiling is about 75 %~ 90% of the total;
- the natural wear in winter time is slower than in summer;
- the re-profiling rate in winter is faster than in summer;
- the wheels installed on the second axel in the second bogie have an abnormally higher wear rate than the wheels installed in the same bogie but on the other axel; this requires more attention;
- The wheels installed in the same bogie perform similarly.

Generally speaking, the results in this study show that for the two locomotives: 1) under the specified installation position and operating conditions, the Weibull frailty model is a useful tool to determine wheel reliability; 2) RCF is the principal reason for re-profiling work orders; 3) the re-profiling parameters can be applied to monitor both the wear rate and the re-profiling loss; 4) the total wear of the wheels can be investigated by considering natural wear and re-profiling loss separately, but natural wear and re-profiling loss differ depending on the locomotive and the operating conditions; and 5) the bogie in which a wheel is installed influences wheel reliability.

7 Conclusions

From the discussion of the research questions, we reach the following conclusions.

First, the proposed integrated procedure for Bayesian reliability inference using MCMC methods has built a full framework for related academic research and engineering applications to implement modern computational-based Bayesian approaches, especially for reliability inference. The suggested procedure is a continuous improvement process with four stages (Plan, Do, Study, and Action) and 11 sequential steps, which can deal with small and incomplete datasets and simultaneously consider the influence of different covariates.

Second, the parametric Bayesian models (including Bayesian Exponential Regression Model, Bayesian Weibull Regression Model, and Log-normal Regression Model, etc.), non-parametric Bayesian models (piecewise constant hazard rate, etc.), frailty models (gamma frailty, etc), as well as the comparison study, are all useful tools for locomotive wheels' reliability analysis using degradation data. Utilising the MCMC technique via the Gibbs sampler can facilitate the integration of high-dimensional probability distributions to make inferences about model parameters and to make predictions.

Third, the results of the case studies show that with the above models, we can determine the locomotive wheel's reliability characteristics, including the baseline hazard rate $\lambda(t)$, reliability $R(t)$, and cumulative hazard rate $\Lambda(t)$, etc. The results also allow us to evaluate and optimise wheel replacement and maintenance strategies (including the re-profiling interval, inspection interval, lubrication interval, depth and optimal sequence of re-profiling, and so on).

Fourth, the case studies' results also reveal that, the wheels' lifetimes differ according to where they are installed on the locomotive. The differences could be influenced by the real running situation (e.g. topography), the locomotive's centre of gravity, the braking forces and the curving forces should also be considered.

Fifth, considering frailties can help with exploring the unobserved covariates and thus improve the model's precision. Results indicate a close positive relationship between the wheels mounted on the same locomotive; the heterogeneity between locomotives is also significant. The results indicate the existence of change points which allow us to evaluate and optimise wheel replacement and maintenance strategies.

Sixth, in the comparison study which takes an integrated data approach to reliability assessment by considering both degradation data and re-profiling data, we reach the following conclusion: 1) rolling contact fatigue (RCF) is the main type of re-profiling work order; 2) the re-profiling parameters can be applied to monitor both the wear rate and the re-profiling loss; 3) the total wear of the wheels can be determined by investigating natural wear and/or loss of wheel diameter through re-profiling loss, but these differ across locomotives and under different operating conditions; 4) the bogie in which a wheel is installed is a key factor in assessing the wheel's reliability.

Last but not least, the approach studied in this report can be applied to cargo train wheels or to other technical problems (e.g. other industries, other components).

8 Future research

Based on the research conducted for this report, we suggest the following research:

- In this research, the case studies only focus on locomotive wheels. We should consider more applications, for instance, cargo train wheels, or other technical problems (e.g. other industries, other components).
- The results achieved by this study, could be extended to other train wheels research topics, e.g., Wheels' "health diagnostic".
- The covariates considered in this report are limited to locomotive wheels' installed positions; more covariates must be considered. These include such factors as operating environment (e.g., climate, topography, track geometry, the braking forces and the curving forces), configuration of the suspension, status of the bogies and the spring systems, operation speeds and the applied loads, etc.
- We have chosen vague prior distributions for most of the case studies. Other prior distributions, including both informative and non-information prior distributions, should be studied.
- In subsequent research, we plan to consider using our results to optimise maintenance strategies and the related LCC (Life Cycle Cost) problem considering maintenance costs, particularly with respect to different maintenance inspection levels and inspection periods (long term, medium term and short term).

9 References

1. Aslanidou H, Dey D K, Sinha D. Bayesian analysis of multivariate survival data using Monte Carlo methods. Canadian Journal of Statistics. 1998, 26: 33-48
2. Au S K, Ching J, Beck J L. Application of Subset Simulation methods to reliability benchmark problems. Structural Safety.2007, 29(3): 183-193
3. Bernasconi A, et al. An Integrated Approach to Rolling Contact Sub-surface Fatigue assessment of Railway Wheels. Journal of Wear. 2005. 258: 973-980
4. Braghin F, et al. A Mathematical Model to Predict Railway Wheel Profile Evolution Due to Wear. Journal of Wear. 2006. 261: 1253-1264
5. Celeux G, et al. Deviance Information Criteria for Missing Data Models. Journal of Bayesian Analysis. 2006, 1(4):651--674
6. Clayton D G. A Model for Association in Bivariate Life Tales and its Application in Epi-demiological Studies of Familial Tendency in Chronic Disease Incidence. *Biometicka.* 1978, 65: 141-151
7. Clayton P. Tribological Aspects of Wheel-Rail Contact: A Review of Recent Experimental Research. Journal of Wear. 1996. 191: 170-183
8. Congdon P. Applied Bayesian Modeling. England: John Wiley and Sons. 2003
9. Congdon P. Bayesian Statistical Modeling. England: John Wiley and Sons. 2001
10. Donato P, et al. Design and Signal Processing of A Magnetic Sensor Array for Train Wheel Detection. Journal of Senors and Actuators A. 2006. 132: 516-525
11. Freitas M A, et al. Reliability assessment using degradation models: Bayesian and classical approaches. Pesquisa Operacional. 2010,30 (1):195-219
12. Freitas M A, et al. Using Degradation Data to Assess Reliability: A Case Study on Train Wheel Degradation. Journal of Quality and Reliability Engineering International. 2009, 25: 607-629
13. Ghosh S K, Ghosal S. Semiparametric Accelerated Failure Time Models for Censored Data. In: Upadhyay S K, *et al.*, editors. Bayesian Statistics and Its Applications. New Delhi: Anamaya Publishers. 2006
14. Graves T L, *et al.* A fully Bayesian Approach for Combining Multi-level Information in Multi-state Fault Tree Quantification. Reliability Engineering and System Safety. 2007, 92: 1476-1483
15. Hamada M, *et al.* A Fully Bayesian Approach for Combining Multilevel Failure Information in Fault Tree Quantification and Optimal Follow-on Rewource Allocation. Reliability Engineering and System Safety. 2004, 86: 297-305
16. Hanson T, Johnson W O. A Bayesian Semiparametric AFT Model for Interval-Censored Data. Journal of Computational and Graphical Statistics. 2004, 13(2): 341-361
17. Ibrahim J G, Chen M H, Sinha D. Bayesian Survival Analysis. New York: Berlin Heidelberg, 2001

18. Ilkka K. Methods and problems of software reliability estimation. VTT Technical Research Centre of Finland. 2006. VTT Working Papers 1459-7683

19. Johansson A, Andersson C. Out-of-round Railway Wheels- a Study of Wheel Polygonalization through Simulation of Three-dimensional Wheel-Rail Interaction and Wear. Journal of Vehicle System Dynamics. 2005, 43(8):539-559

20. Johnson V E, *et al*. A Hierarchical Model for Estimating the Reliability of Complex Systems. In: Bernardo J M, *et al*., editors. Bayesian Statistics 7. Oxford: Oxford University Press. 2003

21. Kelly D L, Smith C L. Bayesian inference in probabilistic risk assessment – the current state of the art. Reliability Engineering and System Safety. 2009, 94: 628-643

22. Komarek A, Lesaffre E. The Regression Analysis of Correlated Interval-censored Data: Illustration using Accelerated Failure Time Models with Flexible Distributional Assupmtions. Statistical Modelling. 2009, 9(4): 299-319

23. Kuo L, Mallick B. Bayesian Semiparametric Inference for the Accelerated Failure Time Model. The Canadian Journal of Statistics. 1997, 25:457-472

24. Kuo L, Yang T Y. Bayesian Reliability Modeling for Masked System Lifetime Data. Statistics & Probability Letters. 2000, 47 (3):229-241

25. Li G Y, Wu Q G, Zhao Y H. On Bayesian Analysis of Binomial Reliability Growth. The Japan Statistical Society.2002, 32(1):1-14

26. Lin J, Asplund M, Parida A. Bayesian Parametric Analysis for Reliability Study of Locomotive Wheels. Conference Proceedings. The 59th Annual Reliability and Maintainability Symposium (RAMS® 2013). January 28-31, Orlando, FL, USA.

27. Lin J, *et al*. Bayesian Survival Analysis in Reliability for Complex System with a Cure Fraction. International Journal of Performability Engineering. 2011, 7 (2): 109-120

28. Lin J. A two-stage Failure Model for Bayesian Change Point analysis. IEEE Transactions on Reliability. 2008, 57 (2): 388-393

29. Liu Y M, et al. Multiaxial Fatigue Reliability Analysis of Railroad Wheels. Journal of Reliability Engineering and System Safety. 2008. 93:456-467

30. Palo M, et al. Rolling stock condition monitoring using wheel/rail forces. Journal of Insight - Non-Destructive Testing and Condition Monitoring. 2012. 54 (8): 451-455

31. Palo M. Condition Monitoring of Railway Vehicles: A Study on Wheel Condition for Heavy Haul Rolling Stock. Licentiate Thesis. Luleå University of Technology, Sweden. 2012;37.

32. Pombo J, Ambrosio J, Pereira M. A Railway Wheel Wear Prediction Tool based on A Multibody Software. Journal of Theoretical and Applied Mechanics. 2010. 48, 3:751-770

33. Robinson D G. A Hierarchical Bayes Approach to System Reliability Analysis. Sandia National Laboratories. 2001. Sandia 2001-3513

34. Sahu S K, Dey D K, Aslanidou H, Sinaha D. A Weibull regression model with gamma frailties for multivariate survival data. *Lifetime Data Analysis*. 1997, 3:123-137

35. Schueller G I, Pradlwarter H J. Benchmark study on Reliability Estimation in Higher Dimensions of Structural System- an Overview. Structural Safety. 2007, 29 (3): 167-182

36. Skarlatos D, Karakasis K, Trochidis A. Railway Wheel Fault Diagnosis Using A Fuzzy-logic Method. Journal of Applied Acoustics. 2004. 65:951-966

37. Spiegelhalter D J, et al. Bayesian measures of model complexity and fit. Journal of Royal Statist. Society Series B.2002, 64(3):583-639

38. Stratman B, Liu Y, Mahadevan S. Structural Health Monitoring of Railroad Wheels Using Wheel Impact Load Detectors. Journal of Failure Analysis and Prevention. 2007. 7(3):218-225

39. Tamura Y, Takehara H, Yamada S. Component-Oriented Reliability Analysis Based on Hierarchical Bayesian Model for an Open Source Software. American Journal of Operations Research. 2011, 1: 25-32

40. Tao J Y, Ming ZM, Chen X. The Bayesian Reliability Growth Models based on a New Dirichlet Prior Distribution. In: Soares CG, *et al.*, editors. Reliability, Risk and Safety. Chapter 237.CRC Press. 2009

41. Tassini N, et al. A Numerical Model of Twin Disc Test Arrangement for the Evaluation of Railway Wheel Wear Prediction Methods. Journal of Wear. 2010. 268: 660-667

42. Tont G, *et al*. Hierarchical Bayesian Reliability Analysis of Complex Dynamical Systems. Proceeding of the 9th WSEAS International Conference on Applications of Electrical Engineering. 2010. 181-186

43. Walker S, Mallick B K. A Bayesian Semi-parametric Accelerated Failure Time Model. Biometrics.1999, 55: 477-483

44. Wilson A G. Hierarchical Markov Chain Monte Carlo (MCMC) for Bayesian System Reliability. Encyclopedia of Statistics in Quality and Reliability. 2008

45. Yang C, Letourneau S. Learning to Predict Train Wheel Failures. Conference Proceedings. The 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005). Chicago, Illinois, USA.

10 Appendix

Paper I

- Lin J. *An Integrated Procedure for Bayesian Reliability Inference using Markov Chain Monte Carlo Methods*. Submitted to Journal.

Paper II

- Lin J, Asplund M, Parida A. *Reliability Analysis for Degradation of Locomotive Wheels using Parametric Bayesian Approach.* Accepted by Quality and Reliability Engineering International. Will be published in 2013. DOI: 10.1002/qre.1518

Paper III

- Lin J, Asplund M. *Bayesian Semi-parametric Analysis for Locomotive Wheel Degradation using Gamma Frailties*. Submitted to Journal.

Paper IV

- Lin J, Asplund M. *A Comparison Study for Locomotive Wheels' Reliability Assessment using the Weibull Frailty Model*. Revised Manuscript has been submitted to Journal of Rail and Rapid and Transit, in April, 2013.

Paper V

- Lin J, Asplund M, Parida A. *Bayesian Parametric Analysis for Reliability Study of Locomotive Wheels. Conference Proceedings.* The 59th Annual Reliability and Maintainability Symposium (RAMS® 2013). January 28-31, Orlando, FL, USA.

Paper I

An Integrated Procedure for Bayesian Reliability Inference using Markov Chain Monte Carlo Methods

- Lin Jing.
- Submitted to Journal.
- Under Review.
- 2013, Feb

An Integrated Procedure for Bayesian Reliability Inference using Markov Chain Monte Carlo Methods

Jing Lin
Division of Operation, and Maintenance Engineering
Luleå University of Technology, 97187, Luleå, Sweden

Abstract:
The recent proliferation of Markov Chain Monte Carlo (MCMC) approaches has led to the use of the Bayesian inference in a wide variety of fields. To facilitate MCMC applications, this paper proposes an integrated procedure for Bayesian inference using MCMC methods, from a reliability perspective. The goal is to build a framework for related academic research and engineering applications to implement modern computational-based Bayesian approaches, especially for reliability inferences. The procedure developed here is a continuous improvement process with four stages (Plan, Do, Study, and Action) and 11 steps, including: 1) data preparation; 2) prior inspection and integration; 3) prior selection; 4) model selection; 5) posterior sampling; 6) MCMC convergence diagnostic; 7) Monte Carlo error diagnostic; 8) model improvement; 9) model comparison; 10) inference making; 11) data updating and inference improvement. The paper illustrates the proposed procedure using a case study.

Keywords:
Bayesian statistics, reliability analysis, Markov Chain Monte Carlo (MCMC), PDSA

1 Introduction

The recent proliferation of Markov Chain Monte Carlo (MCMC) approaches has led to the use of the Bayesian inference in a wide variety of fields, including behavioural science, finance, human health, process control, ecological risk assessment, and risk assessment of engineered systems (Kelly & Smith, 2009, and the references therein). Discussions of MCMC related methodologies and their applications in Bayesian Statistics now appear throughout the literature (Congdon, 2001; 2003). For the most part, studies in reliability analysis focus on the following topics and their cross-applications: 1) hierarchical reliability models (Robinson, 2001; Johnson, *et al.*, 2003; Wilson, 2008; Lin, *et al.*, 2013a); 2) complex system reliability analysis (Tont, *et al.*, 2010; Lin, 2008; Lin, *et al.* 2011); 3) faulty tree analysis (Hamada, *et al.*, 2004; Graves, *et al.*, 2007); 4) accelerated failure models (Kuo & Mallick, 1997; Walker & Mallick, 1999; Hanson & Johnson, 2004; Ghosh & Ghosal, 2006; Komarek & Lesaffre, 2009); 5) reliability growth models (Li, *et al.*, 2002; Tao, *et al.*, 2004); 6) Masked system reliability (Kuo & Yang, 2000); 7) software reliability engineering (Ilkka, 2006; Tamura, *et al*. 2011); 8) Reliability benchmark problems (Schueller & Pradlwarter, 2007; Au, *et al*., 2007). However, most of the literature emphasizes the model's development; no studies offer a full framework to accommodate academic research and engineering applications seeking to implement modern computational-based Bayesian approaches, especially in the area of reliability.

To fill the gap and to facilitate MCMC applications from a reliability perspective, this paper proposes an integrated procedure for the Bayesian inference. The remainder of the paper is organized as follows. Section 2 outlines the integrated procedure; this comprises a continuous improvement process including four stages and 11 sequential steps. Sections 3 to 8 discuss the procedure, focusing on: 1) prior elicitation; 2) model construction; 3) posterior sampling; 4) MCMC convergence diagnostic; 5) Monte Carlo error diagnostic; 6) model comparison. Section 9 gives examples and discusses how to use the procedure. Finally, Section 10 offers conclusions.

2 Description of Procedure

The proposed procedure uses the Bayesian reliability inference to determine system (or unit) reliability and failure distribution, and to support the optimisation of maintenance strategies, *etc*.

The general procedure begins with the collection of reliability data (see Fig.1.). These are the observed values of a physical process, such as various "lifetime data". The data may be subject to uncertainties, such as imprecise measurement, censoring, truncated information, and interpretation errors. Reliability data are found in the "current data set"; they contain original data and include the evaluation, manipulation, and/or organisation of data samples. At a higher level in the collection of data, a wide variety of "historical information" can be obtained, including the results of inspecting and integrating this "information", thereby adding to "prior knowledge". The final level is reliability inference, which is the process of making a conclusion based on "posterior results".

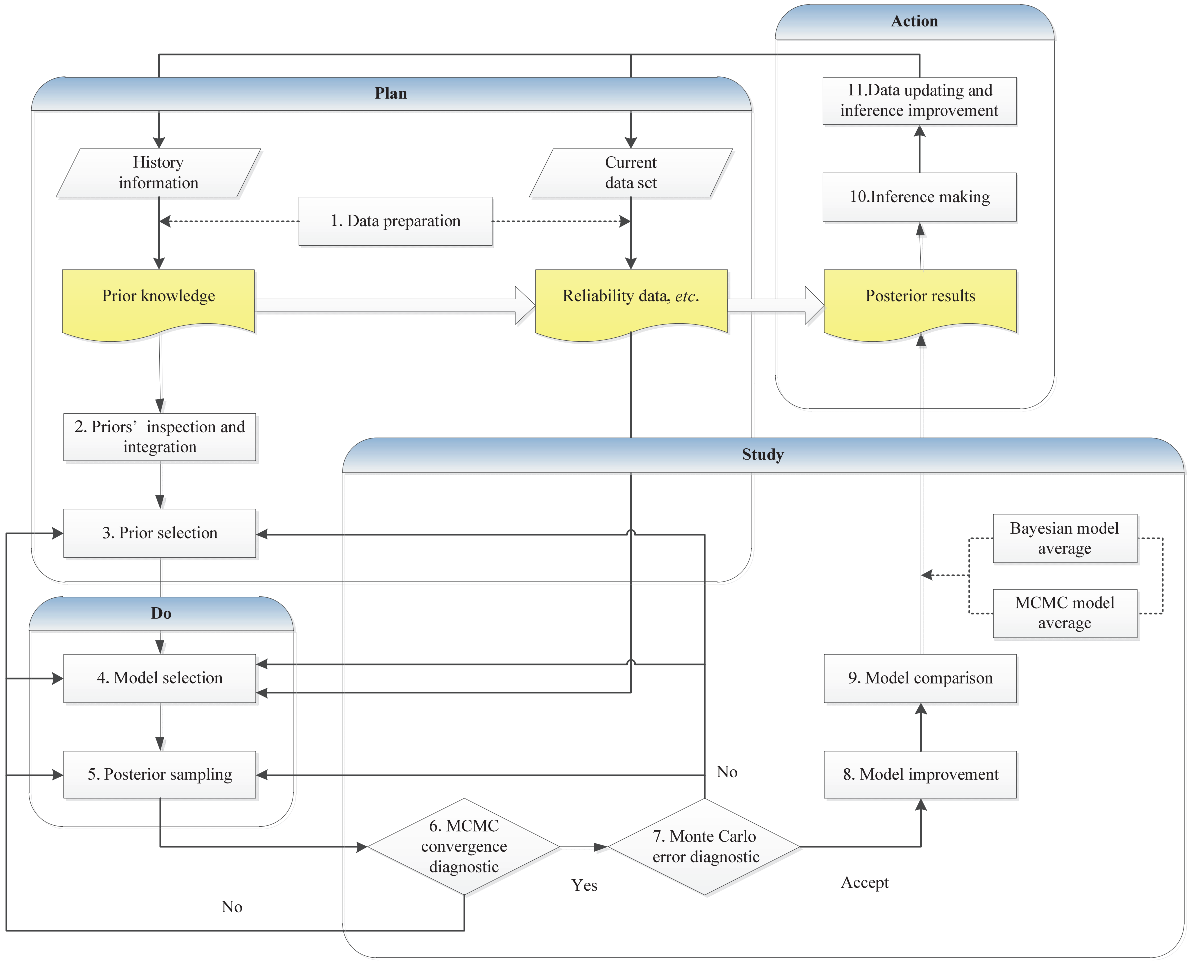


Fig.1. An Integrated Procedure for Bayesian Reliability Inference via MCMC

Using the above definitions, we propose an integrated procedure which constructs a full framework for the standardized process of Bayesian reliability inference. As shown in Fig.1, the procedure is composed of a continuous improvement process including four stages (Plan, Do, Study, Action) which will be discussed later in this section and 11 sequential steps: 1) data preparation; 2) prior inspection and integration; 3) prior selection; 4) model selection; 5) posterior sampling; 6) MCMC convergence diagnostic; 7) Monte Carlo error diagnostic; 8) model improvement; 9) model comparison; 10) inference making; 11) data updating and inference improvement.

- Step 1: Data preparation. The original data sets for “history information” and “current data” related to reliability studies need to be acquired, evaluated, and merged. In this way, “history information” can be transferred to “prior knowledge”, and “current data” can become “reliability data” in later steps.
- Step 2: Prior inspection and integration. During this step, “prior knowledge” receives a second and more extensive treatment, including a reliability consistency check, a credence test, and a multi-source integration. This step improves prior reliability data.
- Step 3: Prior selection. This step uses the results achieved in Step 2 to determine the model’s form and parameters, for instance, selecting informative or non-informative priors, or unknown parameters and their distributed forms.
- Step 4: Model selection. This step determines a reliability model (parametric or non-parametric), selecting from n candidates for the studied system/units. It considers both “reliability data” and the inspection, integration, and selection of priors to implement the i^{th} ($i = 1, \cdots i+1, \cdots n$).
- Step 5: Posterior sampling. In this step, we determine a sampling method (for instance, Gibbs sampling, Metropolis-Hastings sampling, *etc.*) to implement MCMC simulation for the model’s posterior calculations.
- Step 6: MCMC convergence diagnostic. In this step, we check whether the Markov chains have reached convergence. If they have, we move on to the next step; if they have not, we return to Step 5 and re-determine the iteration times of posterior sampling or re-choose the sampling methods; if the results still cannot be satisfied, we return to Steps 3 and 4 and re-determine the prior selection and model selection.
- Step 7: Monte Carlo error diagnostic. We need to decide if the Monte Carlo error is small enough to be accepted in this step. As discussed in Step 6, if it is accepted, we go on to the next step; if it is not, we return to Step 5 and re-decide the iteration times of the posterior sampling or re-choose the sampling methods; if the results still cannot be accepted, we go back to Steps 3 and 4 and recalculate the prior selection and model selection.
- Step 8: Model improvement. Here, we choose the $i+1^{th}$ candidate model, and restart from Step 4.
- Step 9: Model comparison. After implementing n candidate models, we need to: 1) compare the posterior results to determine the most suitable model; or 2) adopt the average posterior estimations (using the Bayesian model average or the MCMC model average) as the final results.
- Step 10: Inference making. After achieving the posterior results in Step 9, we can perform Bayesian reliability inference to determine system (or unit) reliability, find the failure distribution, and optimise maintenance strategies, *etc.*;
- Step 11: Data updating and inference improvement. Along with the passage of time, new “current data” can be obtained, relegating “previous” inference results to “historical data”. By updating “reliability data” and “prior knowledge”, and restarting at Step 1, we can improve the reliability inference.

In summary, by using this step-by-step method, we can create a continuous improvement process for the Bayesian reliability inference.

Note that Steps 1, 2, and 3 are assigned to the "Plan" stage when data for MCMC implementation are prepared. In addition, a part of Steps 1, 2, and 3 refer to the elicitation of prior knowledge. Steps 4 and 5 are both assigned to the "Do" stage, where the MCMC sampling is carried out. Steps 6 to 9 are treated as the "Study" stage; in these steps, the sampling results are checked and compared; in addition, knowledge is accumulated and improved upon by implementing various candidate reliability models. The "Action" stage consists of Steps 10 and 11; at this point, a continuously improved loop can be obtained. In other words, by implementing the step-by-step procedure, we can accumulate and gradually update prior knowledge. Equally, posterior results will be improved upon and become increasingly robust, thereby improving the accuracy of the inference results.

Also note that this paper will focus on six steps and their relationship to MCMC inference implementation: 1) prior elicitation; 2) model construction; 3) posterior sampling; 4) MCMC convergence diagnostic; 5) Monte Carlo error diagnostic; 6) model comparison.

3 Elicitation of Prior Knowledge

In the proposed procedure, the elicitation of prior knowledge plays a crucial role. It is related to Steps 1, 2, and 3 and is part of the Plan Stage, as shown in Fig.1.

In practice, prior information is derived from a variety of data sources and is also considered "historical data" (or "experience data"). Those data take various forms are require various processing methods. Although in the first step, "historical information" can be transferred to "prior knowledge", this is not enough. Credible prior information and proper forms of these data are necessary to compute the model's posterior probabilities, especially in the case of a small sample set. Meanwhile, either non-credible or improper prior data may cause instability in the estimation of the model's probabilities or lead to convergence problems in MCMC implementation. This section will discuss some relevant prior elicitation problems in Bayesian reliability inference, namely, including acquiring priors, performing a consistency check and credence test, fusing multi-source priors, and selecting which priors to use in MCMC implementation.

3.1 Acquisition of Prior Knowledge

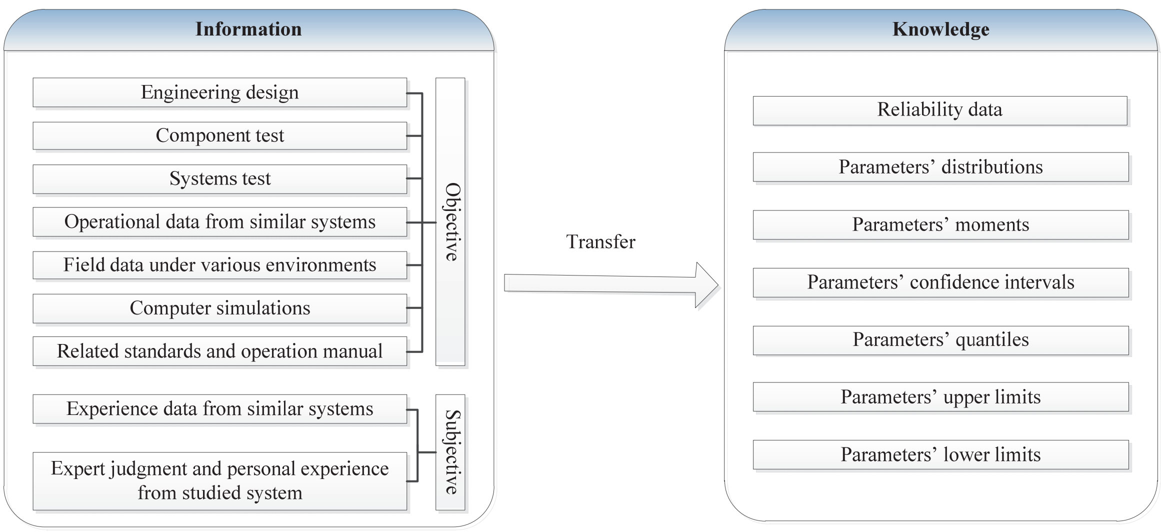


Fig.2. Data Source: Transfer from Historical to Prior Knowledge

In Bayesian reliability analysis, prior knowledge comes from a wide range of historical information. As shown in Fig.2, data sources include: 1) engineering design data; 2) component test data; 3) system test data; 4) operational data from similar systems; 5) field data in various environments; 6) computer simulations; 7) related standards and operation manuals; 8) experience data from similar systems; 9) expert judgment and personal experience. Of these, the first seven yield objective prior data, and the last two provide subjective prior data.

Prior data also take a variety of forms, including reliability data, the distribution of reliability parameters, moments, confidence intervals, quantiles, upper and lower limits, *etc*.

3.2 Inspection of Priors

In Bayesian analysis, different prior information will lead to similar results when the data sample is sufficiently large. While the selection of priors and their form has little influence on posterior inferences, in practice, particularly with a small data sample, we know that some prior information is associated with the current observed reliability data. However, we are not sure whether the prior distributions are the same as the posterior distributions. In other words, we cannot confirm that all posterior distributions converge and are consistent (a consistency check issue). Even if they pass a consistency check, we can only say that they are consistent under a certain specified confidence interval. Therefore, an important prerequisite for applying any prior information is to confirm its credibility by performing a consistency check and credence test.

As noted by Li (1999), the consistency check of prior and posterior distributions was first studied from a statistical viewpoint by Walker (1969). Under specified conditions, posterior distributions are not only consistent with those of the priors, but they have an asymptotic normality which could simplify their calculation. Li (1999) also notes that studies on the consistency check of priors have focused on checking the moments and confidence intervals of reliability parameters, as well as historical data. A number of checking methodologies have been developed, including Robustness analysis, Significance test, Smirov-test, Rank-Sum test, Mood test, *etc*. More studies have been reviewed by Ghosal (2000) and Choi and Ramamoorthi (2008).

The credibility of prior information can be viewed as the probability that it and the collected data come from the same population. Li (1999) lists the following methods to perform a credence test: Frequency method, Bootstrap method, and Rank-Sum text, *etc*. However, in the case of a small sample or simulated data, the above methods are not suitable because even if data pass the credence test, selecting different priors will lead to different results. We therefore suggest a comprehensive use of the above methods to ensure the superiority of Bayesian inference.

3.3 Fusion of Prior Information

Due to the complexity of the system, not to mention the diversification of test equipment and methodologies, prior information can come from many sources. As all priors can pass a consistency check, an integrated fusion estimation based on the smoothness of credibility is sometimes necessary. In such situations, a common choice is parallel fusion estimation or serial fusion estimation, achieved by determining the reliability parameters' weighted credibility (Li, 1999). However, as the credibility computation can be difficult, other methods to fuse the priors may be called for. In the area of reliability, related research studies include the following: Alok *et al*. (1994) develop Bayes estimators for the true binomial survival probability when there are multiple sources of prior information; Ren *et al*. (2002) adopt Kullback information as the distance measure between different prior information and fusion distributions, minimizing the sum to get the combined prior distribution; looking at aerospace propulsion as a case study, Liu *et al*. (2004a) discuss a similar fusion problem in a complex system and suggest (Liu *et al*.; 2004b) a fusion approach based on expert experience with the Analytic Hierarchy Process (AHP); Fang (2006) proposes using multi-source information fusion techniques with Fuzzy-Bayesian for reliability assessment; Zhou *et al*. (2012) propose a Bayes fusion approach for assessment of spaceflight products, integrating degradation data and field lifetime data with Fisher information and the Weiner process. In general, the most important thing for multi-source integration is to determine the weights of the different priors.

3.4 Selection of priors based on MCMC

In Bayesian reliability inference, two kinds of priors are very useful: the conjugate prior and the non-informative prior. To apply MCMC methods, however, the “log-concave prior” is recommended.

The conjugate prior family is very popular, because it is convenient for mathematical calculation. The concept, along with the term "conjugate prior", was introduced by Howard and Robert (1961) in their work on Bayesian decision theory. If the posterior distributions are in the same family as the prior distributions, the prior and posterior distributions are called conjugate distributions, and the prior is called a conjugate prior. For instance, the Gaussian family is a conjugate of itself (or a self-conjugate) with respect to a Gaussian likelihood function: if the likelihood function is Gaussian, choosing a Gaussian prior distribution over the mean distribution will ensure that the posterior distribution is also Gaussian. This means that the Gaussian distribution is a conjugate prior for the likelihood function which is also Gaussian. Other examples include the following: the conjugate distribution of a Normal distribution is a Normal or inverse-Normal distribution; the Poisson and the Exponential distribution’s conjugate both have a Gamma distribution, while the Gamma distribution is a self-conjugate; the Binomial and the negative Binomial distribution’s conjugate both have a Beta distribution; the Polynomial distribution’s conjugate is a Dirichlet distribution, *etc*.

Non-informative prior refers to a prior for which we only know certain parameters’ value ranges or their importance; for example, there may be a uniform distribution. A non-informative prior can also be called a vague prior, flat prior, diffuse prior, or ignorance prior, *etc*. There are many different ways to determine the distribution of a non-informative prior, including Bayes hypothesis, Jeffrey’s rule, reference prior, inverse reference prior, probability-matching prior, Maximum entropy prior, relative likelihood approach, cumulative distribution function, Monte Carlo method, bootstrap method, random weighting simulation method, Harr invariant measurement, Laplace prior, Lindley rule, generalized maximum entropy principle, and the use of marginal distributions. From another perspective, the types of prior distribution also include informative prior, hierarchical prior, Power prior and non-parameter prior processes.

At this point, there are no set rules for selecting prior distributions. Regardless of the manner used to determine a prior’s distribution, the selected prior should be both reasonable and convenient for calculation. Of the above, the conjugate prior is a common choice. To facilitate the calculation of MCMC, especially for adaptive rejection sampling and Gibbs sampling, a popular choice is log-concave prior distribution. Log-concave prior distribution refers to a prior distribution in which the natural logarithm is concave, *i.e.* the second derivative is non-positive. Common logarithmic concavity prior distributions include the normal distribution family, logistic distribution, student's t distribution, the exponential distribution family, the uniform distribution on a finite interval, greater than the gamma distribution with a shape parameter greater than 1, Beta distribution with a value interval (0 , 1), *etc*. As logarithmic concavity prior distributions are very flexible, this paper recommends their use in reliability studies.

4 Model Construction

To apply MCMC methods, we divide the reliability models into four categories: parametric, semi-parametric, frailty, and other non-traditional reliability models.

Parametric modelling offers straightforward modelling and analysis techniques. Common choices include Bayesian Exponential Model, Bayesian Weibull Model, Bayesian Extreme Value Model, Bayesian Log-Normal Model, and Gamma Model. Lin *et al.* (2013b) present a reliability study using the Bayesian parametric framework to explore the impact of a railway train wheel's installed position on its service lifetime and to predict its reliability characteristics. They apply a MCMC computational scheme to obtain the parameters' posterior distributions. Besides the hierarchical reliability models mentioned above, other parametric models include Bayesian Accelerated Failure Models (AFM), Bayesian Reliability Growth models, Bayesian Faulty Tree Analysis (FTA), *etc*.

Semi-parametric reliability models have become quite common and are well accepted in practice, since they offer a more general modelling strategy with fewer assumptions. In these models, the failure rate is described in a semi-parametric form, or the priors are developed by a stochastic process. With respect to the semi-parametric failure rate, Lin *et al.* (2013c) adopt the piecewise constant hazard model to analyze the distribution of the locomotive wheels' lifetime. The applied hazard function is sometimes called a piecewise exponential model; it is convenient because it can accommodate various shapes of the baseline hazard over a number of intervals. In a study of the processes of priors, Ibrahim *et al.* (2001) examine the gamma process, beta process, correlated prior prior processes, and the Dirichlete process separately, using an MCMC computational scheme. By introducing the gamma process of the prior's increment, Lin *et al.* (2007) propose its reliability when applied to the Gibbs sampling scheme.

In reliability inference, most studies are implemented under the assumption that individual lifetimes are independent identified distributed (i.i.d). However, Cox proportional hazard (CPH) models can sometimes not be used because of the dependence of data within a group. Take train wheels as an example; because they have the same operating conditions, the wheels mounted on a particular locomotive may be dependent. In a different context, some data may come from multiple records of wheels installed in the same position but on other locomotives. Modelling dependence has received considerable attention, especially in cases where the datasets may come from related subjects of the same group (Sahu *et al*., 1998; Aslanidou *et al*., 1978). A key development in modelling such data is to build frailty models in which the data are conditionally independent. When frailties are considered, the dependence within subgroups can be considered an unknown and unobservable risk factor (or explanatory variable) of the hazard function. In a recent reliability study, Lin *et al.* (2013c) consider a gamma shared frailty to explore the unobserved covariates' influence on the wheels on the same locomotive.

Some non-traditional Bayesian reliability models are both interesting and helpful. For instance, Lin *et al.* (2011) point out that in traditional methods of reliability analysis, a complex system is often considered to be composed of several subsystems in series. The failure of any subsystem is usually considered to lead to the failure of the entire system. However, some subsystems' lifetimes are long enough or they never fail during the life cycle of the entire system. In addition, such subsystems' lifetimes will not be influenced equally under different circumstances. For example, the lifetimes of some screws will far exceed the lifetime of the compressor in which they are placed. However, the failure of the compressor's gears may directly lead to its failure. In practice, such interferences will affect the model's accuracy, but are seldom considered in traditional analysis. To address these shortcomings, Lin *et al.* (2011) present a new approach to reliability analysis for complex systems, in which a certain fraction of subsystems is defined as a "cure fraction" based on the consideration that such subsystems' lifetimes are long enough or they never fail during the life cycle of the entire system; this is called the cure rate model.

5 Posterior Sampling

To implement MCMC calculations, Markov chains require a stationary distribution. There are many ways to construct these chains. During the last decade, the following Monte Carlo (MC) based sampling methods for evaluating high-dimensional posterior integrals have been developed: MC importance sampling, Metropolis-Hastings sampling, Gibbs sampling, and other hybrid algorithms. In this section, we introduce two common samplings: 1) Metropolis-Hastings sampling, the best known MCMC sampling algorithm, and 2) Gibbs sampling, the most popular MCMC sampling algorithm in the Bayesian computation literature, which is actually a special case of Metropolis-Hastings sampling.

5.1 Metropolis-Hastings sampling

Metropolis-Hastings sampling is a well-known MCMC sampling algorithm, which comes from importance sampling. It was first developed by Metropolis *et al.* (1953) and later generalized by Hastings (1970). Tierney (1994) gives a comprehensive theoretical exposition of the algorithm; Chib and Greenberg (1995) provide an excellent tutorial on it.

Suppose we need to create a sample using the probability density function $p(\theta)$. Let K be a regular constant; this is a complicated calculation (e.g., a regular factor in Bayesian analysis) and is normally an unknown parameter. Then, let $p(\theta) = h(\theta)/K$. Metropolis sampling from $p(\theta)$ can be described as follows:

- Step 1. Choose an arbitrary starting point θ_0, and set $h(\theta_0) > 0$;
- Step 2. Generate a proposal distribution $q(\theta_1, \theta_2)$, which represents the probability for θ_2 to be the next transfer value as the current value is θ_1. The distribution $q(\theta_1, \theta_2)$ is named as the candidate generating distribution. This candidate generating distribution is symmetric, which means $q(\theta_1, \theta_2) = q(\theta_2, \theta_1)$. Now based on the current θ, generate a candidate point θ^* from $q(\theta_1, \theta_2)$;
- Step 3. For the specified candidate point θ^*, calculate the density ratio α with θ^* and the current value $\theta_{t\text{-}1}$ as follows:

  $$\alpha(\theta_{t-1}, \theta^*) = \frac{p(\theta^*)}{p(\theta_{t-1})} = \frac{h(\theta^*)}{h(\theta_{t-1})}$$

  The ratio α refers to the probability to accept θ^*, where the constant K can be neglected;
- Step 4. If θ^* increases the probability density, so that $\alpha > 1$, then accept θ^* and let $\theta_t = \theta^*$. Otherwise, if $\alpha < 1$, then let $\theta_t = \theta_{t-1}$ and go to Step 2.

The acceptance probability α can be written as:

$$\alpha(\theta_{t-1},\theta^*) = \min\left(1, \quad \frac{h(\theta^*)}{h(\theta_{t-1})}\right)$$

Following the above steps, generate a Markov chain with the sampling points $\theta_0,\theta_1,\cdots,\theta_k,\cdots$. The transfer probability from θ_t to θ_{t+1} is related to θ_t, but not related to $\theta_0,\theta_1,\cdots,\theta_{t-1}$. After experiencing a sufficiently long burn-in period, the Markov chain reaches a steady state and the sampling points $\theta_{t+1},\cdots,\theta_{t+n}$ from $p(\theta)$ are obtained.

Metropolis-Hastings sampling was promoted by Hastings (1970). The candidate generating distribution can adopt any form and does not need to be symmetric. In Metropolis-Hastings sampling, the acceptance probability α can be written as:

$$\alpha(\theta_{t-1},\theta^*) = \min\left(1, \quad \frac{h(\theta^*)q(\theta^*,\theta_{t-1})}{h(\theta_{t-1})q(\theta_{t-1},\theta^*)}\right)$$

In the above equation, as $q(\theta^*,\theta_{t-1}) = q(\theta_{t-1},\theta^*)$, and Metropolis-Hastings sampling becomes Metropolis sampling. The Markov transfer probability function can therefore be:

$$p(\theta,\theta^*) = \begin{cases} q(\theta_{t-1},\theta^*) & h(\theta^*) > h(\theta_{t-1}) \\ q(\theta_{t-1},\theta^*)\dfrac{h(\theta^*)}{h(\theta_{t-1})} & h(\theta^*) < h(\theta_{t-1}) \end{cases}$$

From another perspective, when using Metropolis-Hastings sampling, say we need to generate the candidate point θ^*. In this case, generate an arbitrary μ from a uniform distribution $U(0,1)$. Set $\theta_t = \theta^*$ if $\mu \le \alpha(\theta_{t-1},\theta^*)$ and $\theta_t = \theta_{t-1}$ otherwise.

5.2 Gibbs sampling

Metropolis-Hastings sampling is convenient for lower-dimensional numerical computation. However, if θ has a higher dimension, it is not easy to choose an appropriate candidate generating distribution. By using Gibbs sampling, we only need to know the full conditional distribution. Therefore, it is more advantageous in high-dimensional numerical computation. Gibbs sampling is essentially a special case of Metropolis-Hastings sampling, as the acceptance probability equals to one. It is currently the most popular MCMC sampling algorithm in the Bayesian reliability inference literature. Gibbs sampling is based on the ideas of Grenader (1983), but the formal term comes from Geman & Geman (1984) to analyze lattice in image processing. A landmark work for Gibbs sampling in problems of Bayesian inference is Gelfand & Smith (1990). Gibbs sampling is also called heat bath algorithm in statistical physics. A similar idea, data augmentation, is introduced by Tanner & Wong (1987).

Gibbs sampling belongs to the Markov update mechanism and adopts the ideology of "divide and conquer". It supposes all other parameters are fixed and known, inferring a set of parameters. Let θ_i be a random parameter or several parameters in the same group. For the j *th* group, the conditional distribution is $f(\theta_j)$. To carry out Gibbs sampling, the basic scheme is as follows:

- Step1. Choose an arbitrary starting point $\theta^{(0)} = (\theta_1^{(0)},\cdots,\theta_k^{(0)})$;
- Step2. Generate $\theta_1^{(1)}$ from the conditional distribution $f(\theta_1 \mid \theta_2^{(0)},\cdots,\theta_k^{(0)})$ and generate $\theta_2^{(1)}$ from the conditional distribution $f(\theta_2 \mid \theta_1^{(1)},\theta_3^{(0)},\cdots,\theta_k^{(0)})$ ……
- Step3. Generate $\theta_j^{(1)}$ from $f(\theta_j \mid \theta_1^{(1)},\cdots, \theta_{j-1}^{(1)},\theta_{j+1}^{(0)},\cdots,\theta_k^{(0)})$ ……

- Step4. Generate $\theta_k^{(1)}$ from $f(\theta_k \mid \theta_1^{(1)}, \theta_2^{(1)}, \cdots, \theta_{k-1}^{(1)})$;
  As shown above, one step transitions from $\theta^{(0)}$ to $\theta^{(1)} = (\theta_1^{(1)}, \cdots, \theta_k^{(1)})$, where $\theta^{(1)}$ can be viewed as a one-time accomplishment of the Markov chain.
- Step5. Go back to Step2.

After t iterations, $\theta^{(t)} = (\theta_1^{(t)}, \cdots, \theta_k^{(t)})$ can be obtained, and each component $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \cdots$ will be achieved. The Markov transition probability function is:

$$p(\theta, \theta^*) = f(\theta_1 \mid \theta_2, \cdots, \theta_k) f(\theta_2 \mid \theta_1^*, \theta_3 \cdots, \theta_k) \cdots f(\theta_k \mid \theta_1^*, \cdots, \theta_{k-1}^*)$$

Starting from different $\theta^{(0)}$, as $t \to \infty$, the marginal distribution of $\theta^{(t)}$ can be viewed as a stationary distribution based on the theory of the ergodic average. In this case, the Markov chain is seen as converging and the sampling points are seen as observations of the sample.

For both methods, it is not necessary to choose the candidate generating distribution, but it is necessary to do sampling from the conditional distribution. There are many other way to do sampling from the conditional distribution, including sample-importance re-sampling, rejection sampling, Adaptive Rejection Sampling (ARS), *etc*.

6 Markov Chain Monte Carlo Convergence Diagnostic

Because of the Markov chain's ergodic property, all statistics inferences are implemented under the assumption that the Markov chain converges. Therefore, the Markov Chain Monte Carlo convergence diagnostic is very important. When applying MCMC for reliability inference, if the iteration times are too small, the Markov chain will not "forget" the influence of the initial values; if the iteration times are simply increased to a large number, there will be insufficient scientific evidence to support the results, causing a waste of resources.

Markov Chain Monte Carlo convergence diagnostic is a hot topic for Bayesian researchers. Efforts to determine MCMC convergence have been concentrated in two areas. The first is theoretical, and the second is practical. For the former, the Markov transition kernel of the chain is analyzed in an attempt to predetermine a number of iterations that will insure convergence in a total variation distance within a specified tolerance of the true stationary distribution. Related studies include Polson (1996), Roberts & Rosenthal (1997), Mengerson & Tweedie (1996), *etc*. While approaches like these are promising, they typically involve sophisticated mathematics, as well as laborious calculations that must be repeated for every model under consideration. As for practical studies, most research is focused on developing diagnostic tools, including Gelman & Rubin (1992), Raftery & Lewis (1992), Garren & Smith (1993), Roberts & Rosenthal (1998), *etc*. When these tools are used, sometimes the bounds obtained are quite loose; even so, they are considered both reasonable and feasible.

At this point, more than 15 MCMC convergence diagnostic tools have been developed. In addition to a basic tool which provides intuitive judgments by tracing a time series plot of the Markov chain, other examples include tools developed by Gelman & Rubin (1992), Raftery & Lewis (1992), Garren & Smith (1993), Brooks & Gelman (1998), Geweke (1992), Johnson (1994), Mykland, *et al.* (1995), Ritter & Tanner (1992), Roberts (1992), Yu (1994), Yu & Mykland (1995), Zellner & Min (1995), Heidelberger & Welch (1983), Liu *et al.* (1992), *etc*.

We can divide diagnostic tools into categories depending on the following: 1) if the target distribution needs to be monitored; 2) if the target distribution needs to calculate a single Markov chain or multiple chains; 3) if the diagnostic results depend on the output of the Monte Carlo algorithm, not on other information from the target distribution.

It is not wise to rely on one convergence diagnostic technique, and researchers suggest a more comprehensive use of different diagnostic techniques. Some suggestions include: 1) simultaneously diagnosing the convergence of a set of parallel Markov chains; 2) monitoring the parameters' auto-correlations and cross-correlations; 3) changing the parameters of the model or the sampling methods; 4) using different diagnostic methods and choosing the largest pre-iteration times as the formal iteration times; 5) combining the results obtained from the diagnostic indicators' graphs.

Six tools have been achieved by computer programs. The most widely used are the convergence diagnostic tools proposed by Gelman & Rubin (1992) and Raftery & Lewis (1992); the latter is an extension of the former. Both are based on the theory of Analysis of Variance (ANOVA); both use multiple Markov chains and both use the output to perform the diagnostic.

6.1 Gelman-Rubin Diagnostic

In traditional literature on iterative simulations, many researchers suggest that to guarantee the Markov chain's diagnostic ability, multiple parallel chains should be used simultaneously to do the iterative simulation for one target distribution. In this case, after running the simulation for a certain period, it is necessary to determine whether the chains have "forgotten" the initial value and if they converge. Gelman & Rubin (1992) point out that lack of convergence can sometimes be determined from multiple independent sequences but cannot be diagnosed using simulation output from any single sequence. They also find that more information can be obtained during one single Markov chain's replication process than in multiple chains' iterations. Moreover, if the initial values are more dispersed, the status of non-convergence is more easily found. Therefore, they propose a method using multiple replications of the chain to decide whether it becomes stationary in the second half of each sample path. The idea behind this is an implicit assumption that convergence will be achieved within the first half of each sample path; the validity of this assumption is tested by the Gelman-Rubin diagnostic or the variance ratio method.

Based on normal theory approximations to exact Bayesian posterior inference, Gelman-Rubin diagnostic involves two steps. In step one, for a target scalar summary φ, select an over-dispersed starting point from its target distribution $\pi(\varphi)$. Then, generate m Markov chains for $\pi(\varphi)$, where each chain has $2n$ times iterations. Delete the first n times iterations and use the second n times iterations for analysis. In Step 2, apply the between-sequence variance *B/n* and the within-sequence variance W to compare the Corrected Scale Reduction Factor (CSRF). CSFR is calculated by m Markov chains and the formed mn values in the sequences which stem from $\pi(\varphi)$. By comparing the CSFR value, the convergence diagnostic can be implemented. In addition,

$$\frac{B}{n} = \frac{1}{m-1}\sum_{j=1}^{m}(\overline{\varphi}_{j.} - \overline{\varphi}_{..})^2 \ , \ \overline{\varphi}_{j.} = \frac{1}{n}\sum_{t=1}^{n}\varphi_{jt} \ , \ \overline{\varphi}_{..} = \frac{1}{m}\sum_{j=1}^{m}\overline{\varphi}_{j.} \ , W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{t=1}^{n}(\varphi_{jt} - \overline{\varphi}_{j.})^2$$

where φ_{jt} denotes the t th of the n iterations of φ in chain j, and $j = 1,\cdots,m$, $t = 1,\cdots,n$. Let φ be a random variable of $\pi(\varphi)$, with mean μ and variance σ^2 under the target distribution. Suppose that $\widehat{\mu}$ has some unbiased estimator $\hat{V}$. To explain the variability between $\widehat{\mu}$ and $\hat{V}$, Gelman and Rubin construct a student-t distribution with a mean $\widehat{\mu}$ and variance $\hat{V}$ as follows:

$$\hat{\mu} = \overline{\varphi}_{..} \quad \hat{V} = \frac{n-1}{n}W + (1+\frac{1}{m})\frac{B}{n}$$
,

where $\hat{V}$ is a weighted average of W and B. The above estimation will be unbiased if the starting points of the sequences are drawn from the target distribution. However, it will be over-estimated if the starting distribution is over-dispersed. Therefore, $\hat{V}$ is also called a "conservative" estimate. Meanwhile, because the iteration number n is limited, the within-sequence variance W can be too low, leading to falsely diagnosing convergence. As $n \to \infty$, both $\hat{V}$ and W should be close to σ^2. In other words, the scale of current distribution of φ should be reducing as n is increasing.

Denote $\sqrt{\hat{R}_S} = \sqrt{\hat{V}/\sigma^2}$ as the scale reduction factor (SRF). By applying W, $\sqrt{\hat{R}_p} = \sqrt{\hat{V}/W}$ becomes the potential scale reduction factor (PSRF). By applying a correct factor $df/(df-2)$ for PSRF, a correct scale reduction factor (CSRF) can be obtained as follows:

$$\sqrt{\hat{R}_c} = \sqrt{(\frac{\hat{V}}{W})\frac{df}{df-2}} = \sqrt{(\frac{n-1}{n} + \frac{m+1}{mn}\frac{B}{W})\frac{df}{df-2}}$$

where df represents the degree of freedom in student- t distribution. Following Fisher (Lawless, 1982), $df \approx 2\hat{V}/Var(\hat{V})$. The diagnostic based on CSRF can be implemented as follows: if $\hat{R}_c > 1$, it indicates that the iteration number n is too small. When n is increasing, $\hat{V}$ will become smaller and W will become larger. If $\hat{R}_c$ is close to 1 (e.g. $\hat{R}_c < 1.2$), we can conclude that each of the m sets of n simulated observations is close to the target distribution, and the Markov chain can be viewed as converging.

6.2 Brooks-Gelman Diagnostic

Although the Gelman-Rubin diagnostic is very popular, its theory has several defects. Therefore, Brooks & Gelman (1998) extend the method in the following way.

First, in the above equation, $df/(df-2)$ represents the ratio of the variance between the student- t distribution and the normal distribution. Brooks & Gelman (1998) point out some obvious defects in the equation. For instance, if the convergence speed is slow, or $df < 2$, CSRF could be infinite and may even be negative. Therefore, they set up a new and correct factor for PSRF; the new CSRF becomes:

$$\sqrt{\hat{R}_c^*} = \sqrt{(\frac{n-1}{n} + \frac{m+1}{mn}\frac{B}{W})\frac{df+3}{df+1}}$$

Second, Brooks & Gelman (1998) propose a new and more easily implemented way to calculate PSRF. The first step is similar to Gelman-Rubin's diagnostic. Using m chains' second n iterations, obtain an empirical interval $100(1-\alpha\%)$ after each chain's second n iteration. Then, m empirical intervals can be achieved within a sequence, denoted by l. In the second step, determine the total empirical intervals for sequences from mn estimates of m chains, denoted by L. Finally, calculate the PSRF following

$$\sqrt{\hat{R}_p^*} = \sqrt{\frac{l}{L}}$$

The basic theory behind the Gelman-Rubin and Brooks-Gelman diagnostics is the same. The difference is that we compare the variance in the former and the interval length in the latter.

Third, Brooks & Gelman (1998) point out that the value of CSRF being close to one is a necessary but not sufficient condition for MCMC convergence. Additional condition is that both W and $\hat{V}$ should stabilize as a function of n. That is to say, if W and $\hat{V}$ have not reached the steady state, CSRF could still be one. In other words, before convergence, $W < \hat{V}$ and both should be close to one. Therefore, as an alternative, Brooks & Gelman (1998) propose a graphical approach to monitoring convergence. Divide the m sequences into batches of length b. Then calculate $\hat{V}(k)$ 、 $W(k)$ and $\hat{R}_c(k)$ based upon the latter half of the observations of a sequence of length $2kb$, for $k = 1,\cdots,n/b$. Plot $\sqrt{\hat{V}(k)}$, $\sqrt{W(k)}$ and $\hat{R}_c(k)$ as a function of k on the same plot. Approximate convergence is attained if $\hat{R}_c(k)$ is close to one and at the same time, both $\hat{V}(k)$ and $W(k)$ stabilize at one.

Fourth and finally, Brooks & Gelman (1998) discuss the multivariate situation. Let φ denote the parameter vector and calculate the following:

$$\frac{B}{n} = \frac{1}{m-1}\sum_{j=1}^{m}(\overline{\varphi}_{j.} - \overline{\varphi}_{..})(\overline{\varphi}_{j.} - \overline{\varphi}_{..})'$$

$$W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{t=1}^{n}(\varphi_{jt} - \overline{\varphi}_{j.})(\varphi_{jt} - \overline{\varphi}_{j.})'$$

Let λ_1 be Maximum Characteristic Root of $W^{-1}B/n$; then, the PSRF can be expressed as:

$$\sqrt{R_p} = \sqrt{\frac{n-1}{n} + \frac{m+1}{m}\lambda_1}$$

7 Monte Carlo Error Diagnostic

When implementing the MCMC method, besides determining the Markov chains' convergence diagnostic, we must check two uncertainties related to the Monte Carlo point estimation: statistical uncertainty and Monte Carlo uncertainty.

Statistical uncertainty is determined by the sample data and the adopted model. Once the data are given and the models are selected, the statistical uncertainty is fixed. For Maximum Likelihood Estimation (MLE), statistical uncertainty can be calculated by the inverse square root of the Fisher information. For Bayesian inference, statistical uncertainty is measured by the parameters' posterior standard deviation.

Monte Carlo uncertainty stems from the approximation of the model's characteristics, which can be measured by a suitable Standard Error (SE). Monte Carlo standard error of the mean, also called Monte Carlo error (MC error), is a well-known diagnostic tool. In this case, define MC error as the ratio of the sample's standard deviation and the square root of the sample volume, which can be written as:

$$SE[\hat{I}(y)] = \frac{SD[\hat{I}(y)]}{\sqrt{n}} = \left[\frac{1}{n}\left(\frac{1}{n-1}\sum_{i=1}^{n}\left(h(y|x_i)\right) - \hat{I}(y)\right)^2\right]^{1/2}$$

Obviously, as n becomes larger, MC error will be smaller. Meanwhile, the average of the sample data will be closer to the average of the population.

As in the MCMC algorithm, we cannot guarantee that all the sampled points are independent identified distributed (i.i.d.), we must correct the sequence's correlation. To this end, we introduce the auto-correlation function and Sample Size Inflation Factor (SSIF).

Following the sampling methods introduced in section 5, define a sampling sequence $\theta_1, \cdots, \theta_n$ with length n. Suppose there are auto-correlations which exist among the adjacent sampling points; this means $\rho(\theta_i, \theta_{i+1}) \neq 0$. Furthermore, $\rho(\theta_i, \theta_{i+k}) \neq 0$. Then, the auto-correlation coefficient ρ_k can be calculated by:

$$\hat{\rho}_k = \frac{Cov(\theta_t, \theta_{t+k})}{Var(\theta_t)} = \frac{\sum_{t=1}^{n-k} (\theta_t - \bar{\theta})(\theta_{t-k} - \bar{\theta})}{\sum_{t=1}^{n-k} (\theta_t - \bar{\theta})^2}$$

where $\bar{\theta} = \frac{1}{n}\sum_{t=1}^{n} \theta_t$. Following the above discussion, the MC error with consideration of auto-correlations in MCMC implementation can be written as:

$$SE(\bar{\theta}) = \frac{SD}{\sqrt{n}} \sqrt{\frac{1+\rho}{1-\rho}}$$

In the above equation, $SD/\sqrt{n}$ represents the MC error shown in $SE[\hat{I}(y)]$. Meanwhile, $\sqrt{(1+\rho)/(1-\rho)}$ represents the SSIF. $SD/\sqrt{n}$ is helpful to determine whether the sample volume n is sufficient, and SSIF reveals the influence of the auto-correlations on the sample data's standard deviation. Therefore, by following each parameter's MC error, we can evaluate the accuracy of its posterior.

The core idea of the Monte Carlo method is to view the integration of some function $f(x)$ as an expectation of the random variable; therefore, the sampling methods implemented on the random variable are very important. If the sampling distribution is closer to $f(x)$, the MC error will be smaller. In other words, by increasing the sample volume or improving the adopted models, the statistical uncertainty could be reduced; the improved sampling methods could also reduce the Monte Carlo uncertainty.

8 Model Comparison

In Step 8, we might have several candidate models which could pass the MCMC convergence diagnostic and the MC error diagnostic. Thus, model comparison is a crucial part of reliability inference. Broadly speaking, discussions of the comparison of Bayesian models focus on: Bayes factors, model diagnostic statistics, and measure of fit, *etc*. In a more narrow sense, the concept of model comparison refers to selecting a better model after comparing several candidate models. The purpose of doing model comparison is not to determine the model's correctness. Rather, it is to find out why some models are better than others (e.g., which parametric model or non-parametric model; which prior distribution; which covariates; which family of parameters for application, *etc*.), or to obtain an average estimation based on the weighted estimate of the model parameters and stemming from the posterior probability of model comparison (e.g., model average).

In Bayesian reliability inference, the model comparison methods can be divided into three categories:

1) Separate estimation (Gelman *et al.* 2004; Kass & Raftery, 1995; Gelfand & Dey, 1994; Volinsky & Raftery, 2000; Kass & Wasserman, 1995; Spiegelhalter *et al.*2002; Geisser & Eddy, 1979; Gelfand *et al.*1992; Newton & Raftery, 1994; Lenk & Desarbo, 2000; Meng & Wong, 1996; Chib 1995; Chib & Jeliazhov, 2001; Lewis & Raftery, 1997): posterior predictive distribution, Bayes Factors (BF) and its approximate estimation-Bayesian Information Criterion (BIC), Deviance Information Criterion (DIC), Pseuco Bayes Factor (PFB), Conditional Predictive Ordinate (CPO). It also includes some estimations based on the likelihood theory, using the same as BF but offering more flexibility, for instance,

harmonic mean estimator, importance sampling, reciprocal importance estimator, bridge sampling, candidate estimator, Laplace-Metropolis estimator, data augmentation estimator, *etc*.;

2) Comparative estimation, including different distance measures (Sahu & Cheng, 2003; Mengersen & Robert, 1996; Ibrahim & Laud, 1994; Gelfand & Ghosh, 1998; Chen *et al*. 2004): entropy distance, Kullback-Leibler Divergence (KLD), L-measure, and weighted L-measure;
3) Simultaneous estimation (Green, 1995; Robert & Casella, 2004; Stephens, 2000; Tatiana *et al*. 2003; Tanner & Wong, 1987; Gregory, 2012): Reversible Jump MCMC (RJMCMC), Birth and Death MCMC (BDMCMC), Fusion MCMC (FMCMC).

Related reference reviews are given by Kass & Raftery (1995), Tatiana, *et al*.(2003), and Chen & Huang (2005) .

This section introduces three popular comparison methods used in Bayesian reliability studies: BF, BIC and DIC. BF is the most traditional method, BIC is BF's approximate estimation, and DIC improves BIC by dealing with the problem of the parameters' degree of freedom.

8.1 Bayes Factors (BF)

Suppose M represents k models which need to be compared. The data set D stems from $M_i (i = 1, \cdots, k)$, and $M_1, \cdots M_k$ are called competing models. Let $f(D|\theta_i, M_i) = L(\theta_i|D, M_i)$ denote the distribution of D , with consideration of the i th model and its unknown parameter vector θ of dimension p_i , also called the likelihood function of D with a specified model. Under prior distribution $\pi(\theta_i|M_i)$ and $\sum_{i=1}^{k} \pi(\theta_i|M_i) = 1$, the marginal distributions of D are found by integrating out the parameters as:

$$p(D|M_i) = \int_{\Theta_i} f(D|\theta_i, M_i)\pi(\theta_i|M_i)d\theta_i$$

where Θ_i represents the parameter data set for the i th mode. As in the data likelihood function, the quantity $p(D|M_i) = L(D|M_i)$ is called model likelihood. Suppose we have some preliminary knowledge about model probabilities $\pi(M_i)$; after considering the given observed data set D , the posterior probability of i th model being the best model is determined as:

$$\begin{aligned} p(M_i|D) &= \frac{p(D|M_i)\pi(M_i)}{p(D)} \\ &= \frac{\pi(M_i)\int f(D|\theta_i, M_i)\pi(\theta_i|M_i)d\theta_i}{\sum_{j=1}^{k}[\pi(M_j)\int f(D|\theta_j, M_j)\pi(\theta_j|M_j)d\theta_j]} \end{aligned}$$

The integration part of the above equation is also called the prior predictive density or marginal likelihood, where $p(D)$ is a non-conditional marginal likelihood of D .

Suppose there are two models, M_1 and M_2 . Let BF_{12} denote the Bayes factors, equal to the ratio of the posterior odds of the models to the prior odds:

$$BF_{12} = \frac{p(M_1|D)}{p(M_2|D)} \Big/ \frac{\pi(M_1)}{\pi(M_2)} = \frac{p(D|M_1)}{p(D|M_2)}$$

The above equation shows that, BF_{12} equals to the ratio of the model likelihoods for the two models. Thus, it can be written as:

$$\frac{p(M_1|D)}{p(M_2|D)} = \frac{p(D|M_1)}{p(D|M_2)} \times \frac{\pi(M_1)}{\pi(M_2)}$$

We can also say that BF_{12} shows the ratio of posterior odds of the model M_1 and the prior odds of M_1 .In this way, the collection of model likelihoods $p(D|M_i)$ is equivalent to the model probabilities themselves (since the prior probabilities $\pi(M_i)$ are known in advance) and, hence, could be considered as the key quantities needed for Bayesian model choice.

Jeffreys (1961) recommends a scale of evidence for interpreting Bayes factors. Kass & Raftery (1995) provide a similar scale, along with a complete review of Bayes factors, including their interpretation, computation or approximation, robustness to the model-specific prior distributions and applications in a variety of scientific problems.

8.2 Bayesian Information Criterion (BIC)

Under some situations, it is difficult to calculate BF, especially for those models which consider different random effects, or adopt diffusion priors or a large number of unknown and informative priors. Therefore, we need to calculate BF's approximate estimation. The Bayesian Information Criterion (BIC) is also called the Schwarz information criterion (SIC), and is the most important method to get BF's approximate estimation. The key point of BIC is to obtain the approximate estimation of $p(D|M_i)$. It is proved by Volinsky & Raftery (2000) that

$$\ln p(D|M_i) \approx \ln f(D|\hat{\theta}_i, M_i) - \frac{p_i}{2}\ln(n)$$

Then, we get SIC as follows:

$$\begin{aligned}\ln BF_{12} &= \ln p(D|M_1) - \ln p(D|M_2) \\ &= \ln f(D|\hat{\theta}_1, M_1) - \ln f(D|\hat{\theta}_2, M_2) - \frac{p_1 - p_2}{2}\ln(n)\end{aligned}$$

As discussed above, considering two models, M_1 and M_2 , BIC_{12} represents the likelihood ratio test statistic with model sample size n and the model's complexity as penalty. It can be written as

$$\begin{aligned}BIC_{12} &= -2\ln BF_{12} \\ &= -2\ln\left(\frac{f(D|\hat{\theta}_1, M_1)}{f(D|\hat{\theta}_2, M_2)}\right) + (p_1 - p_2)\ln(n)\end{aligned}$$

where p_i is proportional to the model's sample size and complexity.

Kass & Raftery (1995) discuss BIC's calculation program. Kass & Wasserman (1995) show how to decide n . Volinsky & Raftery (2000) discuss the way to choose n if the data are censored. If n is large enough, BF's approximate estimation can be written as

$$BF_{12} \approx \exp(-0.5 BIC_{12})$$

Obviously, if the BIC is smaller, we should consider model M_1 ; otherwise, M_2 should be considered.

8.3 Deviance Information Criterion (DIC)

Traditional methods for model comparison consider two main aspects: the model's measure of fit and the model's complexity. Normally, the model's measure of fit can be increased by increasing the model's complexity. For this reason, most model comparison methods are committed balancing both two points. To utilize BIC, the number p of free parameters of the model must be calculated. However, for complex Bayesian hierarchical models, it is very difficult to get p 's exact number. Therefore, Spiegelhalter *et al.* (2002) propose the Deviance Information Criterion (DIC) to compare Bayesian models. Celeux *et al.* (2006) discuss DIC issues for a censored data set; this paper and other researchers' discussion of it are representative literature in the DIC field in recent years.

DIC utilizes deviance to evaluate the model's measure of fit, and it utilizes the number of parameters to evaluate its complexity. Note that it is consistent with the Akaike Infromation Criterion (AIC), which is used to compare classical models (Akaike, 1971).

Let $D(\theta)$ denote the Bayesian deviance, and

$$D(\theta) = -2\log(p(D|\theta))$$

Let p_d denote the model's effective number of parameters, and

$$\begin{aligned} p_d &= \overline{D}(\theta) - D(\overline{\theta}) \\ &= -\int 2\ln(p(D|\theta))d\theta - (-2\ln(p(D|\overline{\theta}))) \end{aligned}$$

Then,

$$\begin{aligned} DIC &= D(\overline{\theta}) + 2p_d \\ &= \overline{D}(\theta) + p_d \end{aligned}$$

Select the model with a lower DIC value. As DIC < 5 , the difference between models can be ignored.

9 Discussions with a case study

In this section, we discuss a case study for a locomotive wheel's degradation data to illustrate the proposed procedure. The case was first discussed by Lin *et al*. (2013b).

To explore the impact of a locomotive wheel's installed position on its service lifetime and to predict its reliability characteristics, the Bayesian Exponential Regression Model, Bayesian Weibull Regression Model and Bayesian Log-normal Regression Model are used to establish the wheel's lifetime using degradation data and taking into account the position of the wheel. The position is described by three different discrete covariates: the bogie, the axle and the side of the locomotive where the wheel is mounted. The goal is to determine reliability, failure distribution, and optimal maintenance strategies for the wheel.

During the Plan Stage, we first collect the "current data," including the diameter measurements of the locomotive wheel, total distances corresponding to the "time to maintenance" and the wheel's bill of material (BOM) data. Then, we note the installed position and transfer the diameter into degradation data, which becomes "reliability data" during the "data preparation" process. We consider the non-informative prior for the constructed models and select the vague prior with log-concave form, which has been determined to be a suitable choice as a non-informative prior selection.

In the Do Stage, we set up the three models noted above for the degradation analysis, Bayesian Exponential Regression Model, Bayesian Weibull Regression Model and Bayesian Log-normal Regression Model. For our calculations, we implement Gibbs sampling.

After checking the MCMC convergence diagnostic and accepting the Monte Carlo error diagnostic for all three models, in the Study Stage, we compare the model with the three DIC values. After comparing the DIC values, we select the Bayesian Log-normal Regression Model as the most suitable.

With the chosen model's posterior results, in the Action Stage, we make our maintenance predictions and apply them to the proposed maintenance inspection level. This, in turn, allows us to evaluate and optimise the wheel's replacement and maintenance strategies (including the re-profiling interval, inspection interval, lubrication interval, and so on).

As more data are collected in the future, the old "current data set" will be replaced by new "current data"; meanwhile, the results achieved in this case will become "history information", which will be transferred to be "prior knowledge" and a new cycle will start. With this step-by-step method, we can create a continuous improvement process for the locomotive wheel's reliability inference.

10 Conclusions

This paper has proposed an integrated procedure for Bayesian reliability inference using Markov Chain Monte Carlo Methods. The goal is to build a full framework for related academic research and engineering applications to implement modern computational-based Bayesian approaches, especially for reliability inference. The suggested procedure is a continuous improvement process with four stages (Plan, Do, Study, and Action) and 11 sequential steps including: 1) data preparation; 2) priors' inspection and integration; 3) prior selection; 4) model selection; 5) posterior sampling; 6) MCMC convergence diagnostic; 7) Monte Carlo error diagnostic; 8) model improvement; 9) model comparison; 10) inference making; 11) data updating and inference improvement. The paper illustrates the proposed procedure using a case study. It concludes that the procedure can be used to perform Bayesian reliability inference to determine system (or unit) reliability, failure distribution, and to support maintenance strategies optimization, *etc*.

Acknowledgements

The author would like to thank Luleå Railway Research Centre (Järnvägstekniskt Centrum, Sweden) for initiating the research study and Swedish Transport Administration (Trafikverket) for providing financial support.

References

1. Akaike H. Information theory and an extension of the maximum likelihood principle.2nd Internat. International Symposium on Information Theory. Tsahkadsor. 1971
2. Alok P, Ashok S, William J Z. Bayes Reliability Estimation Using Multiple Source of Prior Information: Bionmial Sampling. IEEE Transactions on Reliability. 1994, 43(1): 138-144
3. Aslanidou H, Dey D K, Sinha D. Bayesian analysis of multivariate survival data using Monte Carlo methods. Canadian Journal of Statistics. 1998, 26: 33-48
4. Au S K, Ching J, Beck J L. Application of Subset Simulation methods to reliability benchmark problems. Structural Safety.2007, 29(3): 183-193
5. Brooks S P, Gelman A. Alternative methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics. 1998, 7:434-455

6. Celeux G, Forbes F, Robert C P, Titterington D M. Deviance Information Criteria for Missing Data Models. Bayesian Analysis.2006, 1(4):651--674
7. Chen M H, Dey D K, Ibrahim J G. Bayesian Criterion Based Model Assessment for Categorical Data. Biometrika.2004,91: 45-63
8. Chen M H, Huang S Q. A Comparative Study on Model Selection and Multiple Model Fusion. 7th International Conference on Information Fusion (FUSION). 2005, Conference proceedings: 820-826
9. Chib S, Greenberg E. Understanding the Metropolis-Hastings algorithm. Journal of American Statistician. 1995, 49: 327-335
10. Chib S, Jeliazhov I. Marginal likelihood from the Metropolis-Hastings output. Journal of American Statistical Association.2001. 96: 270–281.
11. Chib S. Marginal likelihood from the Gibbs output. Journal of American Statistical Association.1995, 90: 1313–1321
12. Choi T, and Ramamoorthi R V. Bertrand Clarke and Subhashis Ghosal, eds., Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh (Beachwood, Ohio, USA: Institute of Mathematical Statistics, 2008), 170-186.2008
13. Congdon P. Applied Bayesian Modeling. England: John Wiley and Sons. 2003
14. Congdon P. Bayesian Statistical Modeling. England: John Wiley and Sons. 2001
15. Fang G H. Research on the Multi-source Information Fusion Techniques in the Process of Reliability Assessment. Doctor Thesis. Hefei University of Technology. China. 2006
16. Garren S T, Smith R L. Convergence Diagnostics for Markov Chain Samplers. Department of Statistics, University of North Carolina.1993
17. Geisser S, Eddy W. A predictive approach to model selection. Journal of American Statistical Association. 1979,74:153–60
18. Gelfand A E, Ghosh S K. Model choice: A Minimum Posterior Predictive Loss Approach. Biometrika. 1998, 85: 1-13
19. Gelfand A E, Smith A F M. Sampling-based approaches to calculation marginal densities. Journal of American Asscioation.1990, 85: 398-409
20. Gelfand A, Dey D. Bayesian model choice: Asymptotic and exact calculations. Journal of Royal Statistical Society, Series B.1994, 56:501–514
21. Gelfand, A E, Dey D K, and Chang H. Model Determinating Using Predictive Distributions with Implementation via Sampling-based Methods (with Discussion).In Bayesian Statistics 4. Oxford: Oxford University Press.1992
22. Gelman A, Carlin J B, Stern H S, Rubin D B. Bayesian Data Analysis. New York: Chapman and Hall/CRC.2004
23. Gelman A, Rubin D B. Inference from Iterative Simulation Using Multiple Sequences (with discussion). Statistical Science. 1992, 7: 457-511
24. Geman S, Geman D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. Transactions on Pattern Analysis and Machine Intelligence. 1984, 6: 721-741
25. Geweke J. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In Bayesian Statistics 4 (eds. J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith),Oxford: Oxford University Press, 1992:169-193
26. Ghosal S. A review of consistency and convergence of posterior distribution. Technical report, In Proceedings of National Conference in Bayesian Analysis, Benaras Hindu University, Varanashi, India. 2000
27. Ghosh S K, Ghosal S. Semiparametric Accelerated Failure Time Models for Censored Data. In: Upadhyay S K, *et al.*, editors. Bayesian Statistics and Its Applications. New Delhi: Anamaya Publishers. 2006
28. Graves T L, *et al.* A fully Bayesian Approach for Combining Multi-level Information in Multi-state Fault Tree Quantification. Reliability Engineering and System Safety. 2007, 92: 1476-1483
29. Green P. Reversible Jump Markov Chain Monte Carlo computation and Bayesian model determination. Biometrika. 1995, 82:711–732
30. Gregory P C. Extra-solar Planets via Bayesian Fusion MCMC. Chapter 7 in Astrostatistical Challenges for the New Astronomy. Springer Series in Astrostatistics, Hilbe, J.M (ed), New York: Springer. 2012
31. Grenander U. Tutorial in pattern theory. Providence R I.: Division of Applied Mathematics, Brow University. 1983
32. Hamada M, *et al.* A Fully Bayesian Approach for Combining Multilevel Failure Information in Fault Tree Quantification and Optimal Follow-on Resource Allocation. Reliability Engineering and System Safety. 2004, 86: 297-305

33. Hanson T, Johnson W O. A Bayesian Semiparametric AFT Model for Interval-Censored Data. Journal of Computational and Graphical Statistics. 2004, 13(2): 341-361
34. Hastings W K. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. Biometika. 1970, 57:97-109
35. Heidelberger P, Welch P D. Simulation Run Length Control in the Presence of an Initial Transient Research. 1983, 31: 1109-1144.
36. Howard R, Robert S. Applied Statistical Decision Theory. Division of Research, Graduate School of Business Administration, Harvard University, 1961.
37. Ibrahim J G, Chen M H, Sinha D. Bayesian Survival Analysis. New York: Berlin Heidelberg, 2001
38. Ibrahim J G, Laud P W. A Predictive Approach to the Analysis of Designed Experiments. Journal of the American Statistical Association.1994, 89:309-319
39. Ilkka K. Methods and problems of software reliability estimation. VTT Technical Research Centre of Finland. 2006. VTT Working Papers 1459-7683
40. Jeffreys, H. Theory of Probability. Oxford: Oxford University Press. 1961
41. Johnson V E, *et al*. A Hierarchical Model for Estimating the Reliability of Complex Systems. In: Bernardo J M, *et al.*, editors. Bayesian Statistics 7. Oxford: Oxford University Press. 2003
42. Johnson V E. Studying Convergence of Markov Chain Monte Carlo Algorithms Using Coupled Sampling Paths. Technical Report. Institute for Statistics and Decision Sciences, Duke University. 1994
43. Kass R A, Raftery A E. Bayes factor. Journal of American Statistical Association. 1995, 90: 773-795
44. Kass R E, Wasserman L. A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. Journal of American Statistical Association.1995, 90:928-934
45. Kelly D L, Smith C L. Bayesian inference in probabilistic risk assessment – the current state of the art. Reliability Engineering and System Safety. 2009, 94: 628-643
46. Komarek A, Lesaffre E. The Regression Analysis of Correlated Interval-censored Data: Illustration using Accelerated Failure Time Models with Flexible Distributional Assumptions. Statistical Modelling. 2009, 9(4): 299-319
47. Kuo L, Mallick B. Bayesian Semiparametric Inference for the Accelerated Failure Time Model. The Canadian Journal of Statistics. 1997, 25:457-472
48. Kuo L, Yang T Y. Bayesian Reliability Modeling for Masked System Lifetime Data. Statistics & Probability Letters. 2000, 47 (3):229-241
49. Lawless J F. Statistical Models and Method for Lifetime Data. New York: John Wiley & Sons, 1982.
50. Lenk P, Desarbo W. Bayesian inference for finite mixtures of generalized linear models with random effects. Psychometrika. 2000, 65(1):93-119
51. Lewis S M, Raftery A E. Estimating Bayes factors via posterior simulation with the Laplace-Metropolis estimator. Journal of American Statistical Association.1997, 92: 648-655
52. Li G Y, Wu Q G, Zhao Y H. On Bayesian Analysis of Binomial Reliability Growth. The Japan Statistical Society.2002, 32(1):1-14
53. Li R. Bayes Reliability studies for Complex system. National University of Defense Technology. Doctor Thesis. 1999. In Chinese
54. Lin J, Asplund M, Parida A. (a) Bayesian Parametric Analysis for Reliability Study of Locomotive Wheels. Conference Proceedings. The 59th Annual Reliability and Maintainability Symposium (RAMS® 2013). January 28-31, Orlando, FL, USA.
55. Lin J, Asplund M, Parida A. (b) Reliability Analysis for Degradation of Locomotive Wheels using Parametric Bayesian Approach. Quality and Reliability Engineering International. Accepted in March, 2013
56. Lin J, *et al*. Bayesian Survival Analysis in Reliability for Complex System with a Cure Fraction. International Journal of Performability Engineering. 2011, 7 (2): 109-120
57. Lin J, Han Y Q, Zhu H M. A Semi-parametric Proportional Hazards Regression Model Based on Gamma Process Priors and its Application in Reliability. Chinese Journal of system simulation. 2007, 22: 5099-5102
58. Lin J. A two-stage Failure Model for Bayesian Change Point analysis. IEEE Transactions on Reliability. 2008, 57 (2): 388-393
59. Liu C, Liu J, Rubin D B. A Variational Control Variable for Assessing the Convergence of the Gibbs Sampler. In Proceedings of the American Statistical Association, Statistical Computing Section, 1992:74-78.
60. Liu Q, Feng J, Zhou J L. Application of Similar System Reliability Information to Complex System Bayesian Reliability Evaluation. Journal of Aerospace Power. 2004,19(1):7-10
61. Liu Q, Feng J, Zhou J L. The Fusion Method for Prior Distribution Based on Expert's Information. Journal of Chinese Space Science and Technology. 2004, 24(3): 68-71

62. Meng X, Wong W. Simulating ratios of normalizing constants via a simple identity. Statistical Sinica. 1996, 6: 831–860.
63. Mengersen K, Robert C P. Testing for mixtures: a Bayesian entropic approach. In Bayesian statistics, 5. 25 Oxford: Oxford University Press. 1996
64. Mengerson K, Tweedie R. Rates of convergence of the Hastings and Metropolis algorithms. The Annals of Statistics. 1996, 24: 101-121
65. Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H, Teller E. Equations of state calculations by fast computing machines. Journal of Chemical Physics. 1953, 21:1087-1092
66. Mykland P, Tierney L, Yu B. Regeneration in Markov Chain Samplers. Journal of the American Statistical Association. 1995,(90):233-241
67. Newton M, Raftery A. Approximate Bayesian inference by the weighted likelihood bootstrap. Journal of Royal Statistical Society, Series B.1994, 56:1–48
68. Polson N G. Convergence of Markov chain Conte Carlo algorithms In Bayesian Statistics 5.Oxford: Oxford University Press.1996
69. Raftery A E, Lewis S. How Many Iterations in the Gibbs Sampler in Bayesian Statistics 4. Oxford: Oxford University Press. 1992: 763-773
70. Ren K J, Wu M D, Liu Q. The Combination of Prior Distributions based on Kullback Information. Journal of the Academy of Equipment of Command & Technology. 2002, 13(4):90-92
71. Ritter C, Tanner M A. Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy Gibbs sampler. Journal of the American Statistical Association, 1992:87, 861-868.
72. Robert C, Casella G. Monte Carlo Statistical Methods. New York: Springer-Verlag.2004
73. Roberts G O, Rosenthal J S. Markov chain Monte Carlo: Some practical implications of theoretical results. Canadian Journal of Statistics. 1998, 26:5–31
74. Roberts G O，Rosenthal J S. Markov chain Monte Carlo: some practical implications of theoretical results. Canadian Journal of Statistics. 1997, 26 (1):4-31
75. Roberts G O. Convergence Diagnostics of the Gibbs Sampler In Bayesian Statistics 4.Oxford: Oxford University Press, 1992: 775-782.
76. Robinson D G. A Hierarchical Bayes Approach to System Reliability Analysis. Sandia National Laboratories. 2001. Sandia 2001-3513
77. Sahu S K, Dey D K, Aslanidou H, Sinaha D. A Weibull regression model with gamma frailties for multivariate survival data. Lifetime Data Analysis. 1997, 3:123-137
78. Sahu S, Cheng R. A fast distance based approach for determining the number of components in mixtures. Canadian Journal of Statistics.2003, (31): 3-22.
79. Schueller G I, Pradlwarter H J. Benchmark study on Reliability Estimation in Higher Dimensions of Structural System- an Overview. Structural Safety. 2007, 29 (3): 167-182
80. Spiegelhalter D J, Best N G, Carlin B P, Linde A. Bayesian measures of model complexity and fit. Journal of Royal Statist. Society Series B.2002, 64(3):583-639
81. Stephens M. Bayesian analysis of mixture models with an unknown number of components-an alternative to reversible jump methods. Annual Statistican.2000, 28: 40-74
82. Tamura Y, Takehara H, Yamada S. Component-Oriented Reliability Analysis Based on Hierarchical Bayesian Model for Open Source Software. American Journal of Operations Research. 2011, 1: 25-32
83. Tanner M A, Wong W H. The Calculation of Posterior Distributions by Data Augmentation (with discussion). Journal of American Statistical Association.1987,82:805-811
84. Tao J Y, Ming ZM, Chen X. The Bayesian Reliability Growth Models based on a New Dirichlet Prior Distribution. In: Soares CG, *et al.*, editors. Reliability, Risk and Safety. Chapter 237.CRC Press. 2009
85. Tatiana M, Sylvia F S, Georg D. A comparison of Bayesian model selection based on MCMC with an application to GARCH-type models.[R].Vienna University of Economics and Business Administration. 2003
86. Tierney L. Markov Chains for Exploring Posterior Distributions. The Annals of Statistics. 1994, 22(4):1701-1762
87. Tont G, *et al*. Hierarchical Bayesian Reliability Analysis of Complex Dynamical Systems. Proceeding of the 9th WSEAS International Conference on Applications of Electrical Engineering. 2010. 181-186
88. Volinsky C T, Raftery A E. Bayesian information criterion for censored survival models. Biometrics. 2000, 56:256-262
89. Walker A M. On the asymptotic behaviour of posterior distributions. Journal of Royal Statistics Society B. 1969, 31(1): 80-88
90. Walker S, Mallick B K. A Bayesian Semi-parametric Accelerated Failure Time Model. Biometrics.1999, 55: 477-483
91. Wilson A G. Hierarchical Markov Chain Monte Carlo (MCMC) for Bayesian System Reliability. Encyclopaedia of Statistics in Quality and Reliability. 2008

92. Yu B, Mykland P. Looking at Markov Samplers through Cusum Path Plots: A Simple Diagnostic Idea. Department of Statistics, University of California at Berkeley.1994
93. Yu B. Monitoring the Convergence of Markov Samplers Based on Estimated L1 Error. Department of Statistics, University of California at Berkeley.1994
94. Zellner A, Min C K. Gibbs Sampler Convergence Criteria. Journal of the American Statistical Association. 1995, 90: 921-927
95. Zhou Z B, Li H T, Liu X M, Jin G, Zhou J L. A Bayes information fusion approach for reliability modeling and assessment of spaceflight long life product. Journal of Systems Engineering—Theory & Practice. 2012, 32(11):2517-2522

Paper II

Reliability Analysis for Degradation of Locomotive Wheels using Parametric Bayesian Approach

- Lin Jing, Asplund Matthias, Parida Aditya.
- Accepted by the Journal of Quality and Reliability Engineering International.
- DOI: 10.1002/qre.1518
- Will be published in 2013

Reliability Analysis for Degradation of Locomotive Wheels using Parametric Bayesian Approach

Jing Lin, Matthias Asplund, Aditya Parida
Division of Operation and Maintenance Engineering
Luleå University of Technology, 97187, Luleå, Sweden

Abstract:
This paper undertakes a reliability study using a Bayesian survival analysis framework to explore the impact of a locomotive wheel's installed position on its service lifetime and to predict its reliability characteristics. The Bayesian Exponential Regression Model, Bayesian Weibull Regression Model and Bayesian Log-normal Regression Model are used to analyze the lifetime of locomotive wheels using degradation data and taking into account the position of the wheel. This position is described by three different discrete covariates: the bogie, the axle and the side of the locomotive where the wheel is mounted. The goal is to determine reliability, failure distribution, and optimal maintenance strategies for the wheel. The results show that: 1) under specified assumptions and a given topography, the position of the locomotive wheel could influence its reliability and lifetime; 2) the Bayesian Lognormal Regression Model is a useful tool.

Keywords:

Reliability; Bayesian survival analysis; Locomotive wheels; Deviance Information Criterion (DIC); Markov Chain Monte Carlo (MCMC).

1 Introduction

The service life of a railroad wheel can be significantly reduced due to failure or damage, leading to excessive cost and accelerated deterioration. Damage data show that a major proportion of wheel damage stems from degradation.

In order to monitor the performance of wheels and make replacements before adverse effects occur, the railway industry uses both preventive and predictive maintenance. By predicting the wear of train wheels (Johansson & Andersson [1]; Braghin *et al* [2]; Tassini *et al* [3]), fatigue (Bernasconi *et al* [4]; Liu, *et al* [5]), tribological aspects (Clayton [6]), and failures (Yang & Letourneau [7]), the railway industry can design strategies for different types of preventative maintenance (re-profiled, lubrication, *etc.*) for various time periods (days, months, seasons, running distance, *etc*.). Software dedicated to predicting wear rate has also been studied recently (Pombo *et a* [8]). In addition, condition monitoring data have been studied to increase the wheels' lifetime (Skarlatos D, Karakasis K & Trochidis A [9]; Donato P *et al* [10]; Stratman *et al* [11]; Palo [12,13]). A large number of related studies examining both experimental and numerical aspects have been published in the last decade (see above references).

In one common preventive maintenance policy in the Swedish railway company studied, a wheel's diameter is measured after running a certain distance. If it is reduced to a pre-specified height, the wheel is replaced. Otherwise, it is re-profiled or other maintenance strategies are adopted. To optimize maintenance strategies for railway wheels, some researchers have used degradation data to determine reliability and failure distribution (Freitas *et al* [14,15]; and the reference therein). However, these studies cannot solve the combined problem of small data samples and incomplete data sets while simultaneously considering the influence of several covariates. For example, to avoid the potential influence of the different locations of wheels, the researchers only consider those on the left side of axle number 1 and on certain specified cars.

Other researchers have noted that the wheel's position on the locomotive could influence degradation. For example, researchers from Canada (Yang & Letourneau [7]) suggest that certain attributes, including a wheel's installed position (right or left), might influence its wear rate, but they do not provide case studies. Freitas and colleagues[14] point out that "the degradation of a given wheel might be associated with its position on a given car"; Palo[12,13] conclude that "different wheel positions in a bogie show significantly different force signatures". In a recent seminar in Sweden (Kiruna, April 2012), experts from Norway illustrated their new findings that in a given topography the wheels installed on the right and the left sides experience different force. Unfortunately, they only illustrated the results with signal charts derived from condition monitoring tools. Nor did they consider the influence of wheel position on degradation.

To address the above issues, this paper undertakes a reliability study using a Bayesian survival analysis framework (Ibrahim[16]; Congdon[17,18]; Lin[19]) to explore the impact of the wheel's installed position on its service lifetime and to predict its reliability characteristics. The Bayesian Exponential Regression Model, Bayesian Weibull Regression Model and Bayesian Log-normal Regression Model are used to analyze the lifetime of locomotive wheels using degradation data and taking into account the position of the wheel. This position is described by three different discrete covariates: the bogie, the axle and the side of the locomotive where the wheel is mounted. In particular, by introducing the covariate $\mathbf{x}_i$'s linear function $\mathbf{x}_i^{'}\boldsymbol{\beta}$, these three parameter models are constructed depending on the failure rate λ_i in the exponential model, the log of the rate parameter $\ln(\gamma_i)$ in the Weibull model and the logarithmic mean μ_i in the log-normal models. The contribution of this work is to propose Bayesian survival models, which can solve the combined problem of small data samples and incomplete data sets while simultaneously considering the influence of several covariates. The goal is to determine reliability, failure distribution, and optimal maintenance strategies for the wheel.

The organization of this paper is as follows. The introductory section defines the problem. Section 2 describes the data. Section 3 presents three Bayesian survival models. In those models, some parameters depend on the above-mentioned covariates: the bogie, the axle and the side of the locomotive where the wheel is mounted. Section 4 provides the results for a real data set. This section adopts vague priors and a Markov Chain Monte Carlo (MCMC) computational scheme to obtain the parameters' posterior distributions. Section 5 compares the proposed models with Deviance Information Criterion (DIC), MTTF predictions and discusses the effect on the results of setting maintenance inspection levels. Finally, Section 6 offers conclusions and comments. We also note our ongoing study in the JVTC (Järnvägstekniskt Centrum, Sweden) program.

2 Data Description

This paper focuses on the wheels of the locomotive of a cargo train. While two types of locomotives with the same type of wheels are used in cargo trains, we consider only one.

2.1 Locomotive Wheels' Degradation Data

As shown in Fig.1, there are two bogies for each locomotive and three axels for each bogie. The installed position of the wheels on a particular locomotive is specified by a bogie number (I, II-number of bogies on the locomotive), an axel number (1, 2, 3-number of axels for each bogie) and the side of the wheel on the axle (right or left) where each wheel is mounted.

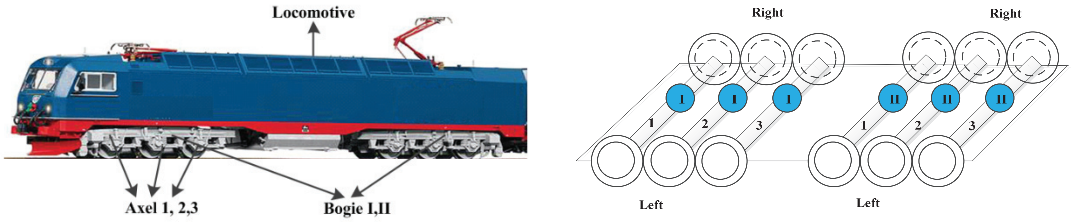


Fig. 1 Wheel positions specified in this study

The diameter of a new locomotive wheel in the studied railway company is 1250 mm. In the company's current maintenance strategy, a wheel's diameter is measured after running a certain distance. If it is reduced to 1150 mm, the wheel is replaced by a new one. Otherwise, it is re-profiled or other maintenance strategies are implemented. A threshold level for failure, denoted as H_1 in this paper, is defined as 100 mm (H_1= 1250 mm -1150 mm). The wheel's failure condition is assumed to be reached if the diameter reaches H_1 .

The company's complete database also includes the diameters of all locomotive wheels at a given observation time, the total running distances corresponding to their "time to be maintained (be re-profiled or replaced)", and the wheels' bill of material (BOM) data, from which we can determine their positions.

2.2 Locomotive Wheels' Lifetime Data

In reliability analyses using degradation data, Freitas et al [14,15] set up a threshold level as defined in section 2.1. The researchers used the degradation data as the wheels' lifetime data. The censored lifetime data were defined if the degradation measurements had reached the threshold level when they were observed. However, in our study, the way we obtain the wheels' lifetime data and how we define censored data differ from theirs.

We make the following assumptions. First, the wheel's degradation follows a linear path (assumption 1). Second, all maintenance activities are assumed to be effective (assumption 2). A specified special maintenance inspection level is denoted as H_2 and $0 \le H_2 \le H_1$. If effective maintenance activities are implemented before the degradation height reaches H_2 , the wheel's degradation speed will be lower; if effective maintenance activities are implemented when the degradation height exceeds H_2 , the degradation speed will remain unchanged.

With respect to assumption (1), Freitas et al [14,15] show that the linear degradation path is reasonable by plotting the historical records of wheel degradation. In our studies, we calculate the squares of their correlation coefficient for a linear path, which are all larger than 0.9 and indicate that the linear degradation path is also a reasonable choice in our case. Assumption (2) is also logical, as the railway company's maintenance activities are intended to prolong the service life of the wheels. However, if maintenance activities are implemented too late (for example, after the degradation height exceeds H_2), the improvement effects are not significant; at most, they will prevent the degradation from speeding up. In addition, any maintenance activity could affect the wheel's diameter, especially if it is re-profiled.

Based on the above assumptions, and as shown in Fig.2 we take the following steps:

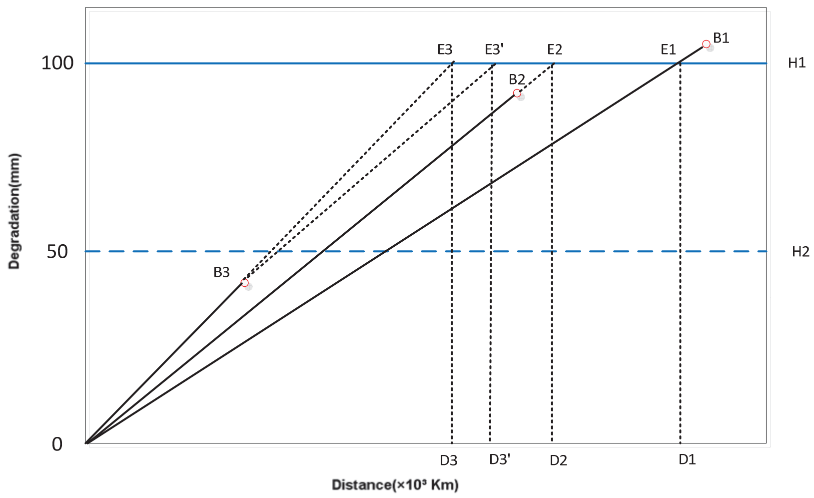


Fig.2 Plot of the wheel degradation data: one example

- Step 1: Establish threshold level H_1. As defined in section 2.1, in Fig.2, we use H_1=100 mm. In addition, according to assumption (2), we establish the specified maintenance inspection level H_2, where $0 \le H_2 \le H_1$. In Fig.2, we use H_2=50 mm.

- Step 2: Transfer the diameters of locomotive wheels at observation time t to degradation data; this equals to 1250 mm minus the corresponding observed diameter. B_1, B_2 and B_3 are three examples of degradation data shown in Fig.2;

- Step 3: According to assumption (1), we assume a liner degradation path and construct a degradation line using the origin point and the degradation data.

- Step 4: If the degradation data are not less than H_1=100 mm (for example: B_1), the degradation line will intersect with H_1. Based on the point of intersection (for example, E_1) and the wheel's failure conditions (see section 2.1), the wheel's lifetime can be determined (for example, D_1).

- Step 5: If the degradation data are less than H_1=100 mm but more than H_2=50 mm (for example, B_2), the extended degradation line will also intersect with H_1 (for example, E_2). Based on the intersection point and according to assumption (2) the wheel's lifetime can be obtained (for example, D_2).

- Step 6: If the degradation data are less than H_2, which equals to 50 mm in Fig.2 (for example, B_3), the intersection point and the lifetime can be derived (for example, E_3 and D_3) as discussed above. However, according to assumption (2), the slope of the extended degradation line could be lower and intersection point could be changed (for example, $E_3^{'}$). Therefore, the lifetime could be longer than the original prediction (for example, $D_3^{'}$). In this case, the lifetime D_3 is defined as right-censored.

By following the above steps, we can obtain the locomotive wheels' lifetime data and the right-censored data.

We consider the wheels of only one locomotive because for the same locomotive: 1) the wheels' maintenance strategies are similar; 2) the axle load and running speed can be supposed to have no obvious difference; 3) the operational environments including routes and climates are expected to be the same. Given these expectations, the installed positions become covariates. Ultimately, we can predict a wheel's lifetime based on its positioning and other important characteristics, including mean time to failure (MTTF).

3 Bayesian Survival Models

3.1 Likelihood construction for right-censored data

In reliability analysis, the lifetime data are usually incomplete, and only a portion of the individual lifetimes are known. Take the locomotive wheels' lifetime data for example. As discussed in section 2.2, if the degradation data B_3 is less than the specified maintenance inspection level H_2, the predicted lifetime D_3 is viewed as right-censored under assumption (2). Therefore, we believe maintenance activities will diminish degradation and the real lifetime $D_3^{'}$ will exceed the predicted lifetime (see Fig.2).

Right-censored data are often called Type I censoring in the literature; the corresponding likelihood construction problem has been extensively studied (Klein & Moeschberger [20]; Lawless [21]). Suppose there are n individuals whose lifetimes and censoring times are independent. The i^{th} individual has life time T_i and censoring time L_i. The T_i s are assumed to have probability density function $f(t)$ and reliability function $R(t)$. The exact lifetime T_i of an individual will be observed only if $T_i \le L_i$. The lifetime data involving right censoring can be conveniently represented by n pairs of random variables (t_i, υ_i), where $t_i = \min(T_i, L_i)$ and $\upsilon_i = 1$ if $T_i \le L_i$, and $\upsilon_i = 0$ if $T_i > L_i$. That is, υ_i indicates whether the lifetime T_i is censored or not. The likelihood function is deduced (Klein & Moeschberger [20]; Lawless [21]) as

$$L(t) = \prod_{i=1}^{n} [f(t_i)]^{\upsilon_i} R(t_i)^{1-\upsilon_i} \tag{1}$$

3.2 Bayesian Exponential regression model

Suppose the lifetimes $\mathbf{t} = (t_1, \cdots t_n)^{'}$ for n individuals are independent identically distributed (i.i.d.), and each corresponds to an exponential distribution with failure rate λ, where $\lambda > 0$. Therefore, the probability density function (p.d.f.) is $f(t_i|\lambda) = \lambda \exp(-\lambda t_i)$. Correspondingly, the cumulative distribution function (c.d.f.) $F(t_i|\lambda)$ and the reliability function $R(t_i|\lambda)$ are $F(t_i|\lambda) = 1 - \exp(-\lambda t_i)$ $= 1 - R(t_i|\lambda)$. Let $\boldsymbol{\upsilon} = (\upsilon_1, \upsilon_2, \cdots, \upsilon_n)^{'}$ indicate whether the lifetime is censored or not, and let the observed data set for current study be denoted as D_0, where $D_0 = (n, \mathbf{t}, \boldsymbol{\upsilon})$; following equation (1), the likelihood function related to λ is given by

$$L(\lambda|D_0) = \prod_{i=1}^{n} [\lambda \exp(-\lambda t_i)]^{\upsilon_i} [\exp(-\lambda t_i)]^{1-\upsilon_i} \quad . \tag{2}$$

Suppose $\mathbf{x}_i = (x_{1i}, \cdots x_{pi})^{'}$ denotes the i^{th} individual of the $p \times 1$ vector of covariates; $\mathbf{X}$ is the $n \times p$ vector of covariates studied in reliability analysis, where p denotes the quantity of the considered covariates. Suppose $\boldsymbol{\beta}$ is a $p \times 1$ vector of regression coefficients, representing the degree of the covariates' influence. Let $\lambda_i = \exp(\mathbf{x}_i^{'}\boldsymbol{\beta})$, and the data set for the current study be denoted by D, where $D = (n, \mathbf{t}, \mathbf{X}, \boldsymbol{\upsilon})$. Following equation (2), the joint likelihood function for the exponential regression model is given by

$$
\begin{aligned}
L(\boldsymbol{\beta}|D) &= \prod_{i=1}^{n}[\exp(\mathbf{x_i}'\boldsymbol{\beta})\exp(-\exp(\mathbf{x_i}'\boldsymbol{\beta})t_i)]^{\upsilon_i}[\exp(-\exp(\mathbf{x_i}'\boldsymbol{\beta})t_i)]^{1-\upsilon_i} \\
&= \exp\left[\sum_{i=1}^{n}\upsilon_i\mathbf{x_i}'\boldsymbol{\beta}\right]\exp\left[-\sum_{i=1}^{n}\exp(\mathbf{x_i}'\boldsymbol{\beta})t_i\right]
\end{aligned} \qquad (3)
$$

The prior distributions should be realistic and computationally feasible. There are two common choices for $\boldsymbol{\beta}$'s prior distribution. One is uniform improper prior distribution, for example, $\pi(\boldsymbol{\beta}) \propto 1$; the other is normal distribution. As proved by Ibrahim *et al* [16], the latter is a log-concave prior distribution and is convenient for the computation of the posterior distribution. In this paper, we assume a multinormal prior distribution $\boldsymbol{\beta} \sim N_p(\boldsymbol{\mu_0}, \boldsymbol{\Sigma_0})$, with mean $\boldsymbol{\mu_0}$ and covariance matrix $\boldsymbol{\Sigma_0}$. Let $\pi(\cdot)$ denote the prior or posterior distributions for the parameters; then, the joint posterior distribution $\pi(\boldsymbol{\beta}|D)$ can be written as

$$
\begin{aligned}
\pi(\boldsymbol{\beta}|D) &\propto L(\boldsymbol{\beta}|D)\times\pi(\boldsymbol{\beta}|\boldsymbol{\mu_0},\boldsymbol{\Sigma_0}) \\
&\propto \exp\left[\sum_{i=1}^{n}\upsilon_i\mathbf{x_i}'\boldsymbol{\beta}\right]\exp\left[-\sum_{i=1}^{n}\exp(\mathbf{x_i}'\boldsymbol{\beta})t_i\right]\times\exp\left[-\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu_0})'\boldsymbol{\Sigma_0^{-1}}(\boldsymbol{\beta}-\boldsymbol{\mu_0})\right] \\
&= \exp\left[\sum_{i=1}^{n}\upsilon_i\mathbf{x_i}'\boldsymbol{\beta}-\sum_{i=1}^{n}\exp(\mathbf{x_i}'\boldsymbol{\beta})t_i-\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu_0})'\boldsymbol{\Sigma_0^{-1}}(\boldsymbol{\beta}-\boldsymbol{\mu_0})\right]
\end{aligned} \qquad (4)
$$

Obviously, it is not easy to get the exact integration results for $\pi(\boldsymbol{\beta}|D)$ due to its complexity. Therefore, we select the MCMC method with the Gibbs sampler, which has been widely applied to Bayesian statistics since the 1990s, to carry out the posterior inference. Let $(-j)$ denote some vector without the j^{th} component. The j^{th} full conditional distribution can be written as

$$
\pi(\beta_j|D,\boldsymbol{\beta}^{(-j)}) \propto L(\beta_j,\boldsymbol{\beta}^{(-j)}|D)\times\pi(\beta_j,\boldsymbol{\beta}^{(-j)}|\boldsymbol{\mu_0},\boldsymbol{\Sigma_0}) . \qquad (5)
$$

3.3 Bayesian Weibull regression model

Suppose the lifetimes $\mathbf{t} = (t_1, \cdots t_n)'$ for n individuals are i.i.d., and each corresponds to a 2-parameter Weibull distribution $W(\alpha, \gamma)$, where $\alpha > 0$ and $\gamma > 0$. The p.d.f. is $f(t_i|\alpha,\gamma) = \alpha\gamma t_i^{\alpha-1}\exp(-\gamma t_i^{\alpha})$ while the c.d.f. $F(t_i|\alpha,\gamma)$ and the reliability function $R(t_i|\alpha,\gamma)$ are $F(t_i|\alpha,\gamma) = 1-\exp(-\gamma t_i^{\alpha}) = 1-R(t_i|\alpha,\gamma)$. To facilitate the analysis, let $\xi = \ln(\gamma)$ (note: it also can be viewed as an extreme value distribution). Then the reliability function becomes

$$
f(t_i|\alpha,\xi) = \alpha t_i^{\alpha-1}\exp(\xi - \exp(\xi)t_i^{\alpha}) . \qquad (6)
$$

Similarly, we can get $F(t_i|\alpha,\xi)$ and $R(t_i|\alpha,\xi)$.

As discussed in section 3.2, the censoring indicators are denoted as $\boldsymbol{\upsilon} = (\upsilon_1, \upsilon_2, \cdots, \upsilon_n)'$ and the observed data set is $D_0 = (n, \mathbf{t}, \boldsymbol{\upsilon})$, following equation (2); therefore, the likelihood function for α and ξ is

$$
L(\alpha,\xi|D_0) = \alpha^{\sum_{i=1}^{n}\upsilon_i}\exp\left\{\sum_{i=1}^{n}\upsilon_i\xi + \sum_{i=1}^{n}\upsilon_i\left[(\alpha-1)\ln(t_i) - \exp(\xi)t_i^{\alpha}\right]\right\} . \qquad (7)
$$

To construct the Weibull Regression Model, we introduce the covariates through ξ. Denoting $\xi_i = \mathbf{x_i}'\boldsymbol{\beta}$ and following other definitions in section 3.2, the joint likelihood function for the Weibull regression model is given by

$$L(\alpha,\boldsymbol{\beta}|D)=\alpha^{\sum_{i=1}^{n}\upsilon_i}\exp\left\{\sum_{i=1}^{n}\upsilon_i\left[\mathbf{x_i'}\boldsymbol{\beta}+\upsilon_i(\alpha-1)\ln(t_i)-\exp(\mathbf{x_i'}\boldsymbol{\beta})t_i^{\alpha}\right]\right\} . \quad (8)$$

In this paper, we take α and ξ to be independent. Furthermore, we assume α to be a gamma distribution denoted by $G(a_0,b_0)$ as its prior distribution, written as $\pi(\alpha|a_0,b_0)$, which means

$$\pi(\alpha|a_0,b_0)\propto\alpha^{a_0-1}\exp(-b_0\alpha) . \quad (9)$$

Assume $\boldsymbol{\beta}$ has a multinormal prior distribution $\pi(\boldsymbol{\beta}|\boldsymbol{\mu_0},\boldsymbol{\Sigma_0})$ with p vector, denoted by $N_p(\boldsymbol{\mu_0},\boldsymbol{\Sigma_0})$. Therefore, the joint posterior distribution can be obtained as

$$\begin{aligned}\pi(\alpha,\boldsymbol{\beta}|D)&\propto L(\alpha,\boldsymbol{\beta}|D)\times\pi(\alpha|a_0,b_0)\times\pi(\boldsymbol{\beta}|\boldsymbol{\mu_0},\boldsymbol{\Sigma_0})\\&\propto\alpha^{a_0-1+\sum_{i=1}^{n}\upsilon_i}\exp\left\{\sum_{i=1}^{n}\left[\upsilon_i\mathbf{x_i'}\boldsymbol{\beta}+\upsilon_i(\alpha-1)\ln(t_i)-\exp(\mathbf{x_i'}\boldsymbol{\beta})t_i^{\alpha}\right]-b_0\alpha-\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu_0})'\boldsymbol{\Sigma_0^{-1}}(\boldsymbol{\beta}-\boldsymbol{\mu_0})\right\}\end{aligned} \quad (10)$$

Then, the parameters' full conditional distribution with Gibbs sampling can be written as

$$\pi(\alpha_j\mid\alpha^{(-j)},\boldsymbol{\beta},D)\propto L(\alpha,\boldsymbol{\beta}\mid D)\times\alpha^{a_0-1}\exp(-b_0\alpha) ; \quad (11)$$

$$\pi(\beta_j\mid\alpha,\boldsymbol{\beta}^{(-j)},D)\propto L(\alpha,\boldsymbol{\beta}\mid D)\times\exp\left[\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu_0})'\boldsymbol{\Sigma_0^{-1}}(\boldsymbol{\beta}-\boldsymbol{\mu_0})\right] . \quad (12)$$

3.4 Bayesian Log-normal regression model

Suppose the lifetimes $\mathbf{t}=(t_1,\cdots t_n)'$ for n individuals are i.i.d. and each $\ln(t)$ corresponds to a normal distribution $N(\mu,\sigma^2)$. We can get t_i 's log-normal distribution with parameters μ and σ^2, denoted by $LN(\mu,\sigma^2)$. Then the p.d.f. and reliability function are given by

$$f(t_i|\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma t_i}\exp\left\{-\frac{1}{2\sigma^2}\left[\ln(t_i)-\mu\right]^2\right\} ; \quad (13)$$

$$R(t_i|\mu,\sigma^2)=1-\Phi\left[\frac{\ln(t_i)-\mu}{\sigma}\right]. \quad (14)$$

The likelihood function related to μ and σ, considering the censoring indicators $\boldsymbol{\upsilon}=(\upsilon_1,\upsilon_2,\cdots,\upsilon_n)'$ and the observed data $D_0=(n,\mathbf{t},\boldsymbol{\upsilon})$, becomes

$$L(\mu,\sigma|D_0)=(2\pi\sigma^2)^{-\frac{1}{2}\sum_{i=1}^{n}\upsilon_i}\exp\left\{-\frac{1}{2\sigma^2}\left[\ln(t_i)-\mu\right]^2\right\}\times\prod_{i=1}^{n}t_i^{-\upsilon_i}\left\{1-\Phi\left[\frac{\log(t_i)-\mu}{\sigma}\right]\right\}^{1-\upsilon_i} . \quad (15)$$

To construct a Log-Normal Regression Model, the covariates through μ are introduced with $\mu_i=\mathbf{x_i'}\boldsymbol{\beta}$. Letting $\tau=1/\sigma^2$, the joint likelihood function is given by

$$L(\boldsymbol{\beta},\tau|D)=(2\pi\tau^{-1})^{-\frac{1}{2}\sum_{i=1}^{n}\upsilon_i}\exp\left\{-\frac{\tau}{2}\sum_{i=1}^{n}\upsilon_i\left[(\ln(t_i)-\mathbf{x_i'}\boldsymbol{\beta})\right]^2\right\}\times\prod_{i=1}^{n}t_i^{-\upsilon_i}\left\{1-\Phi\left[\frac{\ln(t_i)-\mathbf{x_i'}\boldsymbol{\beta}}{\tau^{-1/2}}\right]\right\}^{1-\upsilon_i} . \quad (16)$$

As both μ and τ are assumed unknown, a typical choice for τ is a gamma prior distribution. In this paper, we suppose $\tau\sim G(a_0/2,b_0/2)$. Meanwhile, as $\mu_i=\mathbf{x_i'}\boldsymbol{\beta}$, we also suppose $\boldsymbol{\beta}$ has a multinormal prior distribution with p vector, denoted by $N_p(\boldsymbol{\mu_0}, \tau^{-1}\boldsymbol{\Sigma_0})$. The joint posterior distribution for τ and $\boldsymbol{\beta}$ can be obtained as

$$\begin{aligned}
\pi(\boldsymbol{\beta},\tau|D) &\propto L(\boldsymbol{\beta},\tau|D)\times\pi(\tau|a_0/2,b_0/2)\times\pi(\boldsymbol{\beta}|\boldsymbol{\mu_0},\tau^{-1}\boldsymbol{\Sigma_0}) \\
&\propto \tau^{\frac{a_0+\sum_{i=1}^{n}\upsilon_i}{2}-1}\exp\left\{-\frac{\tau}{2}\sum_{i=1}^{n}\upsilon_i\left[\ln(t_i)-\mathbf{x_i}'\boldsymbol{\beta}\right]^2+(\boldsymbol{\beta}-\boldsymbol{\mu_0})'\boldsymbol{\Sigma_0^{-1}}(\boldsymbol{\beta}-\boldsymbol{\mu_0})+b_0\right\} \\
&\times\prod_{i=1}^{n}t_i^{-\upsilon_i}\left\{1-\Phi\left[\frac{\ln(t_i)-\mathbf{x_i}'\boldsymbol{\beta}}{\tau^{-1/2}}\right]\right\}^{1-\upsilon_i}
\end{aligned} \quad . \tag{17}$$

Therefore, the parameters' full conditional distribution with Gibbs sampling can be written as

$$\pi(\tau_j\mid\tau^{(-j)},\boldsymbol{\beta},D)\propto L(\boldsymbol{\beta},\tau\mid D)\times\tau^{(a_0/2)-1}\exp(-b_0\tau/2)\ ; \tag{18}$$

$$\pi(\beta_j\mid\boldsymbol{\beta}^{(-j)},\tau,D)\propto L(\boldsymbol{\beta},\tau\mid D)\times\exp\left\{\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu_0})'(\tau\boldsymbol{\Sigma}_0)^{-1}(\boldsymbol{\beta}-\boldsymbol{\mu_0})\right\}. \tag{19}$$

4 Case Study

In this section, we present a case study to illustrate our models for locomotive wheels' degradation analysis. The adopted data have been collected from a Swedish railway company's cargo locomotives. The studied locomotive is relatively new compared to others owned by the same company. The degradation data are reported from November 2010 to January 2012. There are 46 records (n =46); we obtained the locomotive wheels' "lifetime" data in the manner described in section 2.2 and shown in Fig.3. In this study, we define H_2 =20 mm; therefore, 12 records are denoted as censored data.

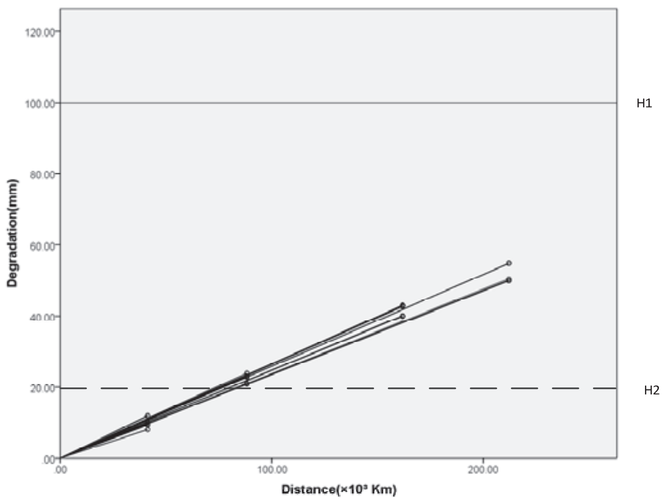


Fig.3 Plot of the wheel degradation data

For each reported datum, a wheel's installation position is documented, and as mentioned above, positioning is used in this study as a covariate. As discussed in section 3, the wheel's position (bogie, axel, and side) or covariate $\mathbf{X}$ is denoted by x_1 (bogie I: x_1=1, bogie II: x_1=2), x_2 (axel 1: x_2=1, axel 2: x_2=2, axel 3, x_2=3) and x_3 (right: x_3=1, left: x_3=2). Correspondingly, the covariates' coefficients are represented by β_1, β_2, and β_3. In addition, β_0 is defined as a constant intercept. Other statistics on the wheel's position and the data structure appear in Table 1.

The calculations are implemented with the software WinBUGS, version 1.4 (Spiegelhalter *et al*[22]). A three-chain is constructed for each MCMC simulation. A burn-in of 10,001 samples is used, with an additional 10,000 Gibbs samples for each Markov chain. Vague prior distributions are adopted here as the following:

- In Bayesian Exponential Regression model:

$\beta_0 \sim N(0,0.0001)$, $\beta_1 \sim N(0,0.0001)$, $\beta_2 \sim N(0,0.0001)$, $\beta_3 \sim N(0,0.0001)$.

- In Bayesian Weibull Regression model:

$\alpha \sim G(0.2,0.2)$, $\beta_0 \sim N(0,0.0001)$, $\beta_1 \sim N(0,0.0001)$, $\beta_2 \sim N(0,0.0001)$, $\beta_3 \sim N(0,0.0001)$.

- In Bayesian Log-normal Regression model:

$\tau \sim G(1,0.01)$, $\beta_0 \sim N(0,0.0001)$, $\beta_1 \sim N(0,0.0001)$, $\beta_2 \sim N(0,0.0001)$, $\beta_3 \sim N(0,0.0001)$.

Table.1 Statistics on quantity and data structure

<table>
<tr><th rowspan="2"></th><th colspan="2">x_1 : Bogie</th><th colspan="2">x_2 : Axel</th><th colspan="2">x_3 : Side</th></tr>
<tr><th>position</th><th>quantities</th><th>position</th><th>quantities</th><th>position</th><th>quantities</th></tr>
<tr><td rowspan="12">n =46</td><td rowspan="6">I (1)*</td><td rowspan="6">24</td><td rowspan="2">1 (1)</td><td rowspan="2">8</td><td>Right (1)</td><td>4</td></tr>
<tr><td>Left (2)</td><td>4</td></tr>
<tr><td rowspan="2">2 (2)</td><td rowspan="2">8</td><td>Right (1)</td><td>4</td></tr>
<tr><td>Left (2)</td><td>4</td></tr>
<tr><td rowspan="2">3 (3)</td><td rowspan="2">8</td><td>Right (1)</td><td>4</td></tr>
<tr><td>Left (2)</td><td>4</td></tr>
<tr><td rowspan="6">II (2)</td><td rowspan="6">22</td><td rowspan="2">1 (1)</td><td rowspan="2">8</td><td>Right (1)</td><td>4</td></tr>
<tr><td>Left (2)</td><td>4</td></tr>
<tr><td rowspan="2">2 (2)</td><td rowspan="2">8</td><td>Right (1)</td><td>4</td></tr>
<tr><td>Left (2)</td><td>4</td></tr>
<tr><td rowspan="2">3 (3)</td><td rowspan="2">6</td><td>Right(1)</td><td>3</td></tr>
<tr><td>Left (2)</td><td>3</td></tr>
</table>

* The number in ( ) denotes the covariate's indicator value as it was used in our models

Following the convergence diagnostics (i.e., checking dynamic traces in Markov chains, time series, and Gelman-Rubin-Statistics, and comparing the MC error with Standard Deviation (SD)), we consider the following posterior distribution summaries (shown in Tables 2, 3 and 4), for our models (Bayesian Exponential Regression Model, Bayesian Weibull Regression Model, and Bayesian Log-normal Regression Model), including the parameters' posterior distribution mean, standard deviation, Monte Carlo error, and 95% HPD (highest posterior distribution density) interval.

Table.2 Posterior Distribution Summaries for Exponential Regression Model

Parameter	Mean	SD	MC error	95 % HPD Interval
β_0	-5.862	0.7355	0.02299	(-7.366,-4.452)
β_1	-0.07207	0.3005	0.007269	(-0.6672,0.5104)
β_2	-0.03219	0.1858	0.003797	(-0.3889,0.3325)
β_3	-0.0124	0.2973	0.00726	(-0.5954,0.5787)

Table.3 Posterior Distribution Summaries for Weibull Regression Model

Parameter	Mean	SD	MC error	95 % HPD Interval
α	10.08	0.9674	0.05559	(8.234,11.76)
β_0	-60.47	5.977	0.3434	(-71.01,-49.16)
β_1	-0.07775	0.306	0.008339	(-0.6845,0.5156)

β_2	-0.146	0.2231	0.005801	(-0.5878,0.2856)
β_3	-0.05026	0.2982	0.007143	(-0.6356,0.5324)

Table.4 Posterior Distribution Summaries for Log-normal Regression Model

Parameter	Mean	SD	MC error	95 % HPD Interval
β_0	5.864	0.05341	0.001622	(5.76,5.97)
β_1	0.06733	0.02174	5.042E-4	(0.02492,0.1103)
β_2	0.02077	0.01373	2.765E-4	(-0.006291,0.04781)
β_3	0.001102	0.02175	5.007E-4	(-0.0412,0.04444)
τ	187.5	39.84	0.3067	(118.3,273.5)

Accordingly, the locomotive wheels' reliability functions can be written as:

- Bayesian Exponential Regression Model:

$$R(t_i \mid \mathbf{X}) = \exp\left[-\exp(-5.862 - 0.072x_1 - 0.032x_2 - 0.012x_3) \times t_i\right] \tag{20}$$

- Bayesian Weibull Regression Model:

$$R(t_i \mid \mathbf{X}) = \exp\left[-\exp(-60.47 - 0.078x_1 - 0.146x_2 - 0.050x_3) \times t_i^{10.08}\right] \tag{21}$$

- Bayesian Log-normal Regression Model:

$$R(t_i \mid \mathbf{X}) = 1 - \Phi\left[\frac{\ln(t_i) - (5.864 + 0.067x_1 + 0.02x_2 + 0.001x_3)}{(187.5)^{-1/2}}\right] \tag{22}$$

Obviously, other reliability characteristics of lifetime distribution, including MTTF, can also be determined.

5 Discussion

5.1 Model Comparison

Traditional technologies for model comparison consider two main aspects: the model's measure of fit and its complexity. Usually, improving the model's complexity can simultaneously improve its fit. For instance, by considering more unknown parameters, the SD and MC error of the model's posterior could be reduced and the model's measure of fit could be improved. However, the complexity of the model will be increased simultaneously. Therefore, most model comparison studies focus on the balance between them. When comparing Bayesian models, both Bayesian Factor (BF) and Bayesian Information Criterion (BIC) can be used. However, for complex Bayesian hierarchical models, it becomes more difficult. Spiegelhalter *et al*[23] have proposed Deviance Information Criterion (DIC), which utilizes the model's deviance to evaluate its measure of fit, and the effective number of parameters to evaluate its complexity.

Define a Bayesian model's Bayesian deviance, denoted as $D(\theta)$, as:

$$D(\theta) = -2\log(p(D|\theta)) . \tag{23}$$

Define the effective number of parameters, denoted as p_d, as:

$$p_d = \overline{D}(\theta) - D(\overline{\theta}) = -\int 2\ln(p(D|\theta))d\theta - (-2\ln(p(D|\overline{\theta}))) \quad . \tag{24}$$

Then,

$$DIC = D(\overline{\theta}) + 2p_d = \overline{D}(\theta) + p_d \ . \tag{25}$$

We calculate the DIC values for the above three Bayesian parametric models separately, as shown in Table 5.

Table.5 DIC Summaries

Model	$\overline{D}(\theta)$	$D(\overline{\theta})$	p_d	DIC
Exponential Regression	648.98	645.03	3.95	652.93
Weibull Regression	472.22	467.39	4.83	477.05
Log-normal Regression	442.03	436.87	5.16	447.19

Based on Celeux et al [24] and related discussions of their paper, we choose the model with the lowest DIC value. When DIC<5, the difference among models can be neglected. Our results show that the DIC for Log-normal Regression Model is the lowest (447.19), and following the arguments above, it is more suitable than the other two. In addition, we analyse other locomotives' wheels with the same model, which are running under similar situations. The results show that similar conclusions can be achieved. However, comparing the DIC values for Weibull Regression Model and for Exponential Regression model, which is 477.05 and 652.93, separately, they indicate that the performance of Weibull Regression Model is close to the Log-normal Regression Model, which might also be a suitable choice under specified situations.

5.2 Maintenance Predictions

Although there is a little difference among the different Bayesian parameter models, all results achieve consistent common conclusion: the installation positions influence the wheels' lifetimes. In addition, considering the character of the covariates' coefficients in our case study, we find the following: 1) the lifetime of the wheel installed in the second bogie is longer than that of the wheel installed in the first one; 2) the wheel installed in the third axel has a longer lifetime than that installed in the second axel, and the wheel in the second axel has a longer lifetime than the one in the first axel; 3) the right side wheel's lifetime is shorter than the left side. (Researchers from Norwegian National Rail Administration cited previously concur with this. Using condition monitoring methods on train wheels operating on the same route, they found that the wheel forces on the right and the left sides can be different, even for wheels in the same axel.). Possible causes include the influence of the earth's rotation, topographical complexity, and the position of the locomotive's centre of gravity.

The three Bayesian parametric regression models presented here are all effective according to Markov chain convergence and other diagnostic tools; see, for example, Spiegelhalter *et al* [23] who compare the computation process, including checking Markov chains' dynamic traces, time series and Gelman-Rubin-Statistics, and comparing the MC error with Standard Deviation (SD). However, we prefer Bayesian Lognormal regression model because of its DIC values. The prediction of the locomotive wheels MTTF, following Bayesian Lognormal regression model, appears in Table.6.

Table.6 MTTF statistics based on Bayesian Lognormal Regression Model

Bogie	Axel	Side	μ_i	MTTF($\times 10^3$ km)
I (x_1 =1)	1 (x_2 =1)	Right (x_3 =1)	5.9532	387.03

		Left (x_3 =2)	5.9543	387.46
	2 (x_2 =2)	Right (x_3 =1)	5.9740	395.16
		Left (x_3 =2)	5.9751	395.60
	3 (x_2 =3)	Right (x_3 =1)	5.9947	403.43
		Left (x_3 =2)	5.9958	403.87
II (x_1 =2)	1 (x_2 =1)	Right (x_3 =1)	6.0205	413.97
		Left (x_3 =2)	6.0216	414.43
	2 (x_2 =2)	Right (x_3 =1)	6.0413	422.67
		Left (x_3 =2)	6.0424	423.14
	3 (x_2 =3)	Right(x_3 =1)	6.0621	431.56
		Left (x_3 =2)	6.0632	432.03

It should be pointed out that the 95% HPD interval in Bayesian Lognormal regression model for β_2 and β_3 includes 0 (Table.4). This means that, although the positioning does have an influence, in some instances, the impact on the wheel's service lifetime is not significantly strong. In our case, the bogies have more impact on service lifetime than axels or sides. Given this conclusion, we can deal with such covariates better in our future research.

5.3 Maintenance Inspection Level

According to the assumptions in section 2.2, the maintenance inspection level H_2 (where $0 \le H_2 \le H_1$) determines how many lifetime data are "right-censored". Obviously, the higher the maintenance inspection level, the more data are considered "right-censored" and vice versa. For instance, in Fig.4, we show a higher maintenance inspection level (80 mm) and a lower one (20 mm). We denote the area between H_1 and H_2 as Zone I, and the area between H_2 and zero degradation level as Zone II. Therefore, based on the likelihood functions discussed in section 3, the MTTF statistics which are achieved from the higher H_2 (the left picture in Fig.4, where H_2 = 80 mm) will be higher than those obtained from the lower H_2 (the right picture, where H_2 =20 mm), because fewer degradation data are considered right-censored. In other words, the results achieved from the former are more "optimistic", and the results obtained from the latter are more "pessimistic". An extreme condition is to suppose H_2 =0 mm.

For this reason, we can get an interval prediction between "optimistic" and "pessimistic" with different maintenance inspection levels, which actually reflect the different risk confidence levels. This will be studied in another research paper.

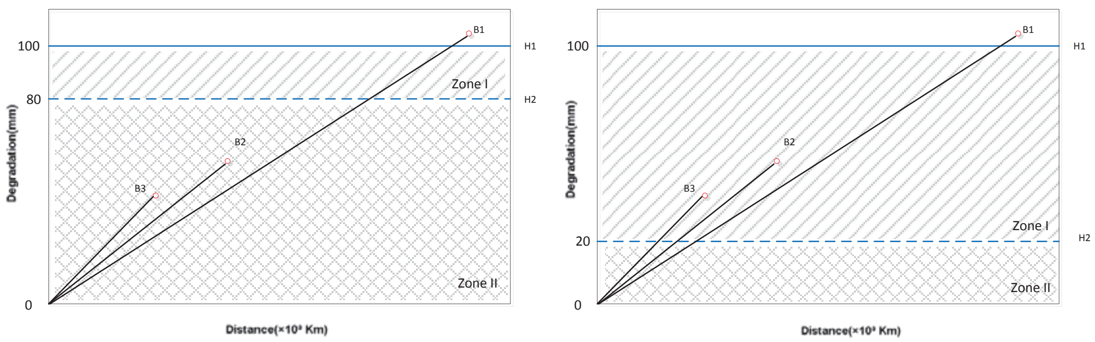


Fig.4 Maintenance Inspection Level with Zone I and Zone II

6 Conclusions

This paper proposes three parametric Bayesian models for locomotive wheels' reliability analysis using degradation data: Bayesian Exponential Regression Model, Bayesian Weibull Regression Model (as discussed in section 3.3.1, it can easily be transferred to an Extreme-Value Regression Model), and Log-normal Regression Model. The Bayesian survival models can deal with small and incomplete data sets and simultaneously consider the influence of several covariates.

The case study's results suggest that the wheels' lifetimes differ according to where they are installed on the locomotive. The wheel installed in the second bogie has a longer lifetime than the one installed in the first bogie; the one installed in the "back" axel has a longer lifetime than the "front" one; the right and left side wheels also differ. The differences between the latter two could be influenced by many aspects, for instance, the locomotives' heterogeneities, the real running situation (e.g. topography, temperature, moisture, applied loading, train speed, *etc.*) and the locomotive's centre of gravity. But the bogies have the strongest influence on wheel lifetime. We can determine the wheel's MTTF using the prediction results obtained from equation (20) ~ (22); this, in turn, allows us to evaluate and optimise the wheel's replacement and maintenance strategies (including the re-profiling interval, inspection interval, lubrication interval, and so on). In addition, by defining different maintenance inspection levels, we can obtain an interval prediction between "optimistic" and "pessimistic" with different risk confidence levels.

Finally, the approach discussed in this paper can be applied to cargo train wheels or to other technical problems (e.g. other industries, other components).

The study suggests the following additional research:

- The assumed liner degradation path is a simple one. For more complex path models, more degradation paths should be studied.
- The covariates considered in this paper are limited to locomotive wheels' installed positions; more covariates must be considered.
- We have chosen vague prior distributions for the case study. Other prior distributions, including both informative and non-information prior distributions, should be studied.

Acknowledgements

The authors would like to thank Luleå Railway Research Centre (Järnvägstekniskt Centrum, Sweden) for initiating the research study and Trafiverket for providing financial support. Also, the authors would like to thank Thomas Nordmark, Ove Salmonsson and Hans-Erik Fredriksson at LKAB for support and discussions about the locomotive wheels. And we would like to thank the editor and anonymous referees for their constructive comments.

References

1. Johansson A, Andersson C. Out-of-round Railway Wheels- a Study of Wheel Polygonalization through Simulation of Three-dimensional Wheel-Rail Interaction and Wear. *Journal of Vehicle System Dynamics*. 2005, 43(8):539-559. DOI: 10.1080/00423110500184649
2. Braghin F, et al. A Mathematical Model to Predict Railway Wheel Profile Evolution Due to Wear. *Journal of Wear.* 2006. 261: 1253-1264. DOI: 10.1016/j.wear.2006.03.025

3. Tassini N, et al. A Numerical Model of Twin Disc Test Arrangement for the Evaluation of Railway Wheel Wear Prediction Methods. *Journal of Wear*. 2010. 268: 660-667. DOI: 10.1016/j.wear.2009.11.003

4. Bernasconi A, et al. An Integrated Approach to Rolling Contact Sub-surface Fatigue assessment of Railway Wheels. *Journal of Wear.* 2005. 258: 973-980. DOI: 10.1016/j.wear.2004.03.044

5. Liu Y M, et al. Multiaxial Fatigue Reliability Analysis of Railroad Wheels. *Journal of Reliability Engineering and System Safety*. 2008. 93:456-467. DOI: 10.1016/j.ress.2006.12.021

6. Clayton P. Tribological Aspects of Wheel-Rail Contact: A Review of Recent Experimental Research. *Journal of Wear*. 1996. 191: 170-183. DOI: 10.1016/0043-1648(95)06651-9

7. Yang C, Letourneau S. Learning to Predict Train Wheel Failures. *Conference Proceedings*. The 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005). Chicago, Illinois, USA.

8. Pombo J, Ambrosio J, Pereira M. A Railway Wheel Wear Prediction Tool based on A Multibody Software. *Journal of Theoretical and Applied Mechanics*. 2010. 48, 3:751-770

9. Skarlatos D, Karakasis K, Trochidis A. Railway Wheel Fault Diagnosis Using A Fuzzy-logic Method. *Journal of Applied Acoustics*. 2004. 65:951-966. DOI: 10.1016/j.apacoust.2004.04.003

10. Donato P, et al. Design and Signal Processing of A Magnetic Sensor Array for Train Wheel Detection. *Journal of Senors and Actuators A.* 2006. 132: 516-525. DOI: 10.1016/j.sna.2006.02.043

11. Stratman B, Liu Y, Mahadevan S. Structural Health Monitoring of Railroad Wheels Using Wheel Impact Load Detectors. *Journal of Failure Analysis and Prevention*. 2007. 7(3):218-225. DOI: 10.1007/s11668-007-9043-3

12. Palo M, et al. Rolling stock condition monitoring using wheel/rail forces. *Journal of Insight - Non-Destructive Testing and Condition Monitoring*. 2012. 54 (8): 451-455. DOI: 10.1784/insi.2012.54.8.451

13. Palo M. Condition Monitoring of Railway Vehicles: A Study on Wheel Condition for Heavy Haul Rolling Stock. Licentiate Thesis. Luleå University of Technology, Sweden. 2012;37.

14. Freitas M A, et al. Using Degradation Data to Assess Reliability: A Case Study on Train Wheel Degradation. *Journal of Quality and Reliability Engineering International.* 2009, 25: 607-629. DOI: 10.1002/qre.995

15. Freitas M A, et al. Reliability assessment using degradation models: Bayesian and classical approaches. *Pesquisa Operacional.* 2010,30 (1):195-219

16. Ibrahim J G, Chen M H, Sinha D. Bayesian Survival Analysis. New York: Berlin Heidelberg, 2001

17. Congdon P. Applied Bayesian Modelling. England: John Wiley and Sons.2003

18. Congdon P. Bayesian Statistical Modelling. England: John Wiley and Sons.2001

19. Lin J. Two-Stage Failure Model for Bayesian Change Point Analysis. *Journal of IEEE Transactions on Reliability*. 2008, 57(2): 388-393. DOI: 10.1109/TR.2008.923484

20. Klein J P, Moeschberger M L. Survival Analysis: Techniques for Censored and Truncated Data. Springer-Verlag: New York, Inc.1997; 55.

21. Lawless. Statistical Models and Methods for Lifetime Data. John Wiley and Sons.1982;31.

22. Spiegelhalter D, et al. WinBUGS User Manual (Version 1.4). January, 2003. http:// www.mrc-bsu.cam.ac.uk/bugs [11 September 2012]

23. Spiegelhalter D J, et al. Bayesian measures of model complexity and fit. *Journal of Royal Statist. Society Series B*.2002, 64(3):583-639. DOI: 10.1111/1467-9868.00353

24. Celeux G, et al. Deviance Information Criteria for Missing Data Models. *Journal of Bayesian Analysis.* 2006, 1(4):651--674

Paper III

Bayesian Semi-parametric Analysis for Locomotive Wheel Degradation using Gamma Frailties

- Lin Jing, Asplund Matthias.
- Submitted to Journal.
- Under Review.
- 2013, March.

Bayesian Semi-parametric Analysis for Locomotive Wheel Degradation using Gamma Frailties

Jing Lin, Matthias Asplund
Division of Operation and Maintenance Engineering, Luleå University of Technology, 97187, Luleå, Sweden

Abstract: This paper undertakes a reliability study using a Bayesian semi-parametric framework to explore the impact of a locomotive wheel's position on its service lifetime and to predict its other reliability characteristics. A piecewise constant hazard regression model is used to establish the lifetime of locomotive wheels using degradation data and taking into account the wheel's bogie. The gamma frailties are included in this study to explore unobserved covariates within the same group. The goal is to flexibly determine reliability for the wheel. The case study is performed using Markov Chain Monte Carlo (MCMC) methods; the results show that: 1) a polynomial degradation path is a better choice for the studied locomotive wheels; 2) under given operation conditions, the position of the locomotive wheel, i.e., in which bogie it is mounted, could influence its reliability; 3) the piecewise constant hazard regression model is a useful tool since it contains fewer assumptions; 4) considering gamma frailties is helpful for exploring unobserved covariates' influence and for improving the model's precision; 5) some change points exist after the wheels run a certain distance, a finding which could be applied maintenance review and optimisation.

Keywords: reliability; Bayesian survival analysis; locomotive wheels; frailty; piecewise constant hazard rate; Markov Chain Monte Carlo (MCMC).

1 Introduction

The service life of a train wheel can be significantly reduced due to failure or damage, leading to excessive cost and accelerated deterioration, a point which has received considerable attention in recent literature. In order to monitor the performance of wheels and make replacements in a timely fashion, the railway industry uses both preventive and predictive maintenance. By predicting the wear of train wheels (Johansson & Andersson, 2005; Braghin *et al.*, 2006; Tassini *et al.*, 2010), fatigue (Bernasconi *et al.,* 2005; Liu, *et al.,* 2008), tribological aspects (Clayton, 1996), and failures (Yang & Letourneau, 2005), the industry can design strategies for different types of preventative maintenance (re-profiling, lubrication, *etc.*) for various periods (days, months, seasons, running distance, *etc.*). Software dedicated to predicting wear rate has also been proposed (Pombo *et al.,* 2010). Finally, condition monitoring data have been studied with a view to increasing the wheels' lifetime (Skarlatos *et al.*, 2004; Donato *et al.,* 2006; Stratman *et Al.*, 2007; Palo, 2012).

One common preventive maintenance strategy (used in the case study) is re-profiling wheels after they run a certain distance. Re-profiling affects the wheel's diameter; once the diameter is reduced to a pre-specified length, the wheel is replaced by a new one. Seeking to optimise this maintenance strategy, researchers have examined wheel degradation data to determine wheel reliability and failure distribution (Freitas *et al.,* 2009, 2010; and the references therein). However, these studies cannot solve the combined problem of small data samples and incomplete datasets while simultaneously considering the influence of several covariates. For example, to avoid the potential influence of wheel location, Freitas *et al.* (2009, 2010) only consider those on the left side of specified axle and on certain specified cars, but point out that "the degradation of a given wheel might be associated with its position on a given car". Yang and Letourneau (2005) suggest that certain attributes, including a wheel's installed position (right or left), might influence its wear rate, but they do not provide case studies. Palo *et al.* (2012) concludes that "different wheel positions in a bogie show significantly different force signatures". In a recent seminar in Sweden (Kiruna, April 2012), experts from Norway illustrated their new findings that in a given topography, the wheels installed on the right and the left sides of a locomotive experience different force. Unfortunately, they did not illustrate the results with signal charts derived from condition monitoring tools, nor did they consider the influence of wheel position on degradation. To address the above issues, Lin *et al.* (2013) have explored the influence of locomotive wheels' positioning on reliability. Their results indicate that the particular bogie in which the wheel is mounted has more influence on its lifetime than does the axel or which side it is on. Therefore, in this paper, we only use the bogie as a covariate.

Most reliability studies are implemented under the assumption that individual lifetimes are independent identified distributed (*i.i.d*). However, sometimes Cox proportional hazard (CPH) models cannot be used because of the dependence of data within a group. For instance, because they have the same operating conditions, the wheels mounted on a particular locomotive may be dependent. In a different context, some data may come from multiple records which actually belong to the wheels installed in the same position but on another locomotive. Modelling dependence in multivariate survival data has received considerable attention, in cases where the datasets may come from subjects of the same group which are related to each other (Sahu *et al.,* 1997; Aslanidou *et al.,* 1998). A key development in modelling such data is to consider frailty models, in which the data are conditionally independent. When frailties are considered, the dependence within subgroups can be considered an unknown and unobservable risk factor (or explanatory variable) of the hazard function. In this paper, we consider a gamma shared frailty, first discussed by Clayton (1978) and Oakes (1982) and later developed by Sahu *et al.* (1997), to explore the unobserved covariates' influence on the wheels on the same locomotive.

In addition, since semi-parametric Bayesian methods offer a more general modelling strategy that contains fewer assumptions (Ibrahim *et al.,* 2001), we adopt the piecewise constant hazard model to establish the distribution of the locomotive wheels' lifetime. The applied hazard function is sometimes referred to as a piecewise exponential model; it is convenient because it can accommodate various shapes of the baseline hazard over the intervals.

This paper explores the impact of a locomotive wheel's installed position on its service lifetime and predicts its other reliability characteristics by using a Bayesian semi-parametric framework. The remainder of the paper is organised as follows. Section 2 presents the piecewise constant hazard regression model with gamma frailties. In the proposed model, a discrete-time martingale process is considered as a prior process for the baseline hazard rate. Section 3 describes a real case study using the dataset for the wheels of two locomotives in a heavy haul cargo train. Using polynomial degradation, it considers the bogies as covariates and uses a gamma frailty for each locomotive. It adopts a Markov Chain Monte Carlo (MCMC) computational scheme (Congdon 2001, 2003) and discusses maintenance strategies for optimisation. Section 4 offers conclusions and comments for future study.

2 Models

In this section, we propose a Bayesian semi-parametric framework, incorporating the piecewise constant hazard regression model, a gamma shared frailty model, the discrete-dime martingale process for the baseline hazard rate, and a MCMC computation scheme.

2.1 Piecewise Constant Hazard Regression Model

The piecewise constant hazard model is one of the most convenient and popular semi-parametric models in survival analysis. Begin by denoting the j^{th} individual in the i^{th} group as having lifetime t_{ij}, where $i = 1,\cdots,n$ and $j = 1,\cdots,m_i$. Divide the time axis into intervals $0 < s_1 < s_2 < \cdots < s_k < \infty$, where $s_k > t_{ij}$, thereby obtaining k intervals $(0, s_1]$, $(s_1, s_2]$, $\cdots$ $(s_{k-1}, s_k]$. Suppose the j^{th} individual in the i^{th} group has a constant baseline hazard $h_0(t_{ij}) = \lambda_k$ as in the k^{th} interval, where $t_{ij} \in I_k = (s_{k-1}, s_k]$. Then, the hazard rate function for the piecewise constant hazard model can be written as

$$h_0(t_{ij}) = \lambda_k, \qquad t_{ij} \in I_k \tag{1}$$

Equation (1) is sometimes referred to as a piecewise exponential model; it can accommodate various shapes of the baseline hazard over the intervals.

Studies of how to divide the time axis into k intervals include the following. Kalbfleisch & Prentice (1973) suggest that the selection of intervals should be made independently of the data; this has been adopted in the construction the traditional lifetime table. Breslow (1974) suggests using distinct failure times as end points of each interval. Sahu *et al.* (1997), Aslanidou *et al.* (1998), and Ibrahim *et al.* (2001) discuss the robustness of choosing different k separately. In this paper, we discuss the choice of k in the case study.

Suppose $\mathbf{x}_i = (x_{1i}, \cdots x_{pi})'$ denotes the covariate vector for the individuals in the i^{th} group, and $\boldsymbol{\beta}$ is the regression parameter. Therefore, the regression model with the piecewise constant hazard rate can be written as

$$h(t_{ij}) = \begin{cases} \lambda_1 \exp(\mathbf{x}_{\mathbf{ij}}'\boldsymbol{\beta}) & 0 < t_{ij} \le s_1 \\ \lambda_2 \exp(\mathbf{x}_{\mathbf{ij}}'\boldsymbol{\beta}) & s_1 < t_{ij} \le s_2 \\ \vdots & \vdots \\ \lambda_k \exp(\mathbf{x}_{\mathbf{ij}}'\boldsymbol{\beta}) & s_{k-1} < t_{ij} \le s_k \end{cases} \tag{2}$$

Correspondingly, its probability density function $f(t_{ij})$, cumulative distribution function $F(t_{ij})$, reliability function $R(t_{ij})$, together with the cumulative hazard rate $\Lambda(t_{ij})$ can be achieved (Ibrahim *et al.,* 2001).

2.2 Gamma Shared Frailty Model

Frailty models are first considered by Clayton (1978) to handle multivariate survival data. In their models, the event times are conditionally independent according to a given frailty factor, which is an individual random effect. As discussed by Sahu *et al.* (1997), the models formulate different variabilities and come from two distinct sources. The first source is natural variability, which is explained by the hazard function; the second is variability common to individuals of the same group or variability common to several events of an individual, which is explained by the frailty.

Assume the hazard function for the j^{th} individual in the i^{th} group is

$$h_{ij}(t) = h_0(t)\exp(\mu_i + \mathbf{x}_{\mathbf{ij}}'\boldsymbol{\beta}) \tag{3}$$

In equation (3), μ_i represents the frailty parameter for the i^{th} group. By denoting $\omega_i = \exp(\mu_i)$, the equation can be written as

$$h_{ij}(t) = h_0(t)\omega_i \exp(\mathbf{x}_{\mathbf{ij}}'\boldsymbol{\beta}) \tag{4}$$

Equation (3) is an additive frailty model, and equation (4) is a multiplicative frailty model. In both equations, μ_i and ω_i are shared by the individuals in the same group, and they are thus referred to as shared-frailty models and actually are extensions of the CPH model.

To this point, discussions of frailty models have focused on the choices of: 1) the form of the baseline hazard function; 2) the form of the frailty's distribution. Representative studies related to the former include the gamma process for the accumulated hazard function (Clayton, 1991; Sinha, 1993), Weibull baseline hazard rate (Sahu *et al.,* 1995), and the piecewise constant hazard rate (Aslanidou *et al.,* 1998) which is adopted in this paper due to its flexibility. Some researchers have examined finite mean frailty distributions, including gamma distribution (Clayton *et al.,* 1978; Clayton & Cuzick, 1985), lognormal distribution (McGilchrist, 1991), and the like; others have studied non-parameter methods, including the inverse Gaussion frailty distribution (Hougaard, 1986), the power variance function for frailty (Crowder, 1989), the positive stable frailty distribution (Hougaard, 1995; Qiou *et al.,*1999), the Dirichlet process frailty model (Pennell & Dunson, 2006) and the Levy process frailty model (Hakon *et al.,* 2003). In this paper, we consider the gamma shared frailty model, the most popular model for frailty.

From equation (4), suppose the frailty parameters ω_i are independent and identically distributed (*i.i.d*) for each group, and follow a gamma distribution, denoted by $Ga(\kappa^{-1},\kappa^{-1})$. Therefore, the probability density function can be written as

$$f(\omega_i)=\frac{(\kappa^{-1})^{\kappa^{-1}}}{\Gamma(\kappa^{-1})}\cdot\omega_i^{\kappa^{-1}-1}\exp(-\kappa^{-1}\omega_i) \tag{5}$$

In equation (5), the mean value of ω_i is one, where κ is the unknown variance of ω_i s. Greater values of κ signify a closer positive relationship between the subjects of the same group as well as greater heterogeneity among groups. Furthermore, as ω_i >1, the failures for the individuals in the corresponding group will appear earlier than if ω_i =1; in other words, as ω_i <1, their predicted lifetimes will be greater than those found in the independent models.

Suppose $\boldsymbol{\omega}=(\omega_1,\omega_2,\cdots,\omega_n)'$; then

$$\pi(\boldsymbol{\omega}|\kappa)\propto\prod_{i=1}^{n}\omega_i^{\kappa^{-1}-1}\exp(-\kappa^{-1}\omega_i) \tag{6}$$

2.3 Discrete-time Martingale Process for Baseline Hazard Rate

Based on the above discussions (equation (2), (4), and (5)), the piecewise constant hazard model with gamma shared frailties can be written as:

$$h(t_{ij})=\begin{cases}\lambda_1\omega_i\exp(\mathbf{x_{ij}'\beta}) & 0<t_{ij}\le s_1\\ \lambda_2\omega_i\exp(\mathbf{x_{ij}'\beta}) & s_1<t_{ij}\le s_2\\ \vdots & \vdots\\ \lambda_k\omega_i\exp(\mathbf{x_{ij}'\beta}) & s_{k-1}<t_{ij}\le s_k\end{cases} \tag{7}$$

In equation (7), $\omega_i\sim Ga(\kappa^{-1},\kappa^{-1})$.

To analysis the baseline hazard rate λ_k , a common choice is to construct an independent incremental process, e.g., the Gamma process, the Beta process, or the Dirichlet process. However, as pointed out by Ibrahim *et al.* (2001), in many applications, prior information is often available on the smoothness of the hazard rather than the actual baseline hazard itself. In addition, given the same covariates, the ratio of marginal hazards at the nearby time-points is approximately equal to the ratio of the baseline hazards at these points. In such situations, correlated prior processes for the baseline hazard can be more suitable. Such models, for instance, the discrete-time martingale process for the baseline hazard rate λ_k , are discussed by Sahu *et al.* (1997) and Aslanidou *et al.* (1998).

Given ($\lambda_1,\lambda_2,\cdots,\lambda_{k-1}$), we specify that

$$\lambda_k|\lambda_1,\lambda_2,\cdots,\lambda_{k-1}\sim Ga(\alpha_k,\frac{\alpha_k}{\lambda_{k-1}}) \tag{8}$$

Let $\lambda_0 = 1$. In equation (8), the parameter α_k represents the smoothness for the prior information. If $\alpha_k = 0$, then λ_k and λ_{k-1} are independent. As $\alpha_k \to \infty$, the baseline hazard is the same in the nearby intervals. In addition, the Martingale λ_k 's expected value at any time point is the same, and

$$E(\lambda_k | \lambda_1, \lambda_2, \cdots, \lambda_{k-1}) = \lambda_{k-1} \tag{9}$$

Equation (9) shows that given specified historical information ($\lambda_1, \lambda_2, \cdots, \lambda_{k-1}$), the expected value of λ_k is fixed.

2.4 Bayesian Semi-parametric Model using MCMC

In reliability analysis, the lifetime data are usually incomplete, and only a portion of the individual lifetimes are known. Right-censored data are often called Type I censoring, and the corresponding likelihood construction problem has been extensively studied in the literature (Lawless, 1982; Klein & Moeschberger, 1997). Suppose the j^{th} individual in the i^{th} group has lifetime T_{ij} and censoring time L_{ij}. The observed lifetime $t_{ij} = \min(T_{ij}, L_{ij})$; therefore, the exact lifetime T_{ij} will be observed only if $T_{ij} \le L_{ij}$. In addition, the lifetime data involving right censoring can be represented by n pairs of random variables (t_{ij}, υ_{ij}), where $\upsilon_{ij} = 1$ if $T_{ij} \le L_{ij}$ and $\upsilon_{ij} = 0$ if $T_{ij} > L_{ij}$. This means that υ_{ij} indicates whether lifetime T_{ij} is censored or not. The likelihood function is deduced as

$$L(t) = \prod_{i=1}^{n} \prod_{j=1}^{m_i} [f(t_{ij})]^{\upsilon_{ij}} R(t_{ij})^{1-\upsilon_{ij}} \tag{10}$$

In the above piecewise constant hazard model, we denote g_{ij} as $t_{ij} \in (s_{g_{ij}}, s_{g_{ij}+1}) = I_{g_{ij}+1}$ and the model's dataset as $D = (\boldsymbol{\omega}, \mathbf{t}, \mathbf{X}, \boldsymbol{\upsilon})$. Following equation (7) ~ (10), the complete likelihood function $L(\boldsymbol{\beta}, \boldsymbol{\lambda} | D)$ for the individuals for the i^{th} group in k intervals can be written as

$$\prod_{i=1}^{n} \prod_{j=1}^{m_i} \left\{ [\prod_{k=1}^{g_{ij}} \exp(-\lambda_k \omega_i \exp(\mathbf{x}_{\mathbf{ij}}^{'} \boldsymbol{\beta})(s_k - s_{k-1})] \times (\lambda_{g_{ij}+1} \omega_i \exp(\mathbf{x}_{\mathbf{ij}}^{'} \boldsymbol{\beta}))^{\upsilon_{ij}} \times \exp[-\lambda_{g_{ij}+1} \omega_i \exp(\mathbf{x}_{\mathbf{ij}}^{'} \boldsymbol{\beta})(t_{ij} - s_{g_{ij}})] \right\} \tag{11}$$

Let $\pi(\cdot)$ denote the prior or posterior distributions for the parameters. Following equation (6) and (11), the joint posterior distribution $\pi(\omega_i | \boldsymbol{\beta}, \boldsymbol{\lambda}, D)$ for gamma frailties ω_i can be written as

$$\begin{aligned} \pi(\omega_i | \boldsymbol{\beta}, \boldsymbol{\lambda}, D) &\propto L(\boldsymbol{\beta}, \lambda | D) \times \pi(\boldsymbol{\omega} | \kappa) \\ &\propto \omega_i^{\kappa^{-1} + \sum_{j=1}^{m_i} \upsilon_{ij} - 1} \exp\left\{ -(\kappa^{-1} + [\sum_{j=1}^{m_i} \exp(\mathbf{x}_{\mathbf{ij}}^{'} \boldsymbol{\beta})]) \times (\sum_{k=1}^{g_{ij}} \lambda_k (s_k - s_{k-1}) + \lambda_{g_{ij}+1}(t_{ij} - s_{g_{ij}})) \right\} \\ &\sim Ga\left\{ \kappa^{-1} + \sum_{j=1}^{m_i} \upsilon_{ij}, \kappa^{-1} + [\sum_{j=1}^{m_i} \exp(\mathbf{x}_{\mathbf{ij}}^{'} \boldsymbol{\beta})](\sum_{k=1}^{g_{ij}} \lambda_k (s_k - s_{k-1}) + \lambda_{g_{ij}+1}(t_{ij} - s_{g_{ij}})) \right\} \end{aligned} \tag{12}$$

Equation (12) shows that the full conditional density of each ω_i is a gamma distribution. Similarly, the full conditional density of κ^{-1} and $\boldsymbol{\beta}$ can be given by

$$\pi(\kappa^{-1} | \boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\lambda}, D) \propto \prod_{i=1}^{n} \omega_i^{\kappa^{-1}-1} (\kappa^{-1})^{-n\kappa^{-1}} \times \frac{\exp\left(-\kappa^{-1} \sum_{i=1}^{n} \omega_i\right)}{[\Gamma(\kappa^{-1})]^n} \cdot \pi(\kappa^{-1}) \tag{13}$$

$$\pi(\boldsymbol{\beta} | \kappa^{-1}, \boldsymbol{\omega}, \boldsymbol{\lambda}, D) \propto \exp\left\{ \sum_{i=1}^{n} \sum_{j=1}^{m_i} \upsilon_{ij} \mathbf{x}_{\mathbf{ij}}^{'} \boldsymbol{\beta} - \sum_{i=1}^{n} \sum_{m=1}^{n_i} \exp(\mathbf{x}_{\mathbf{ij}}^{'} \boldsymbol{\beta}) \omega_i \times \left[\sum_{k=1}^{g_{ij}} \lambda_k (s_k - s_{k-1}) + \lambda_{g_{ij}+1}(t_{ij} - s_{g_{ij}}) \right] \right\} \times \pi(\boldsymbol{\beta}) \tag{14}$$

Let $R_k = \{(i, j); t_{ij} > s_k\}$ denote the risk set at s_k and $D_k = R_{k-1} - R_k$; let d_k denote the failure individuals in the interval I_k . Let $\pi(\lambda_k | \boldsymbol{\lambda}^{(-k)})$ denote the conditional prior distribution for ($\lambda_1, \lambda_2, \cdots, \lambda_J$) without λ_k . We therefore derive $\pi(\lambda_k | \boldsymbol{\beta}, \boldsymbol{\omega}, \kappa^{-1}, D)$ as

$$\lambda_k^{d_k} \exp\left\{ -\lambda_k \omega_i \exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta}) \times [\sum_{(i,j)\in R_k} (s_k - s_{k-1}) + \sum_{(i,j)\in D_k} (t_{ij} - s_{k-1})] \right\} \times \pi(\lambda_k | \boldsymbol{\lambda}^{(-k)}) \tag{15}$$

3 Case Study

In this section, we present a case study to illustrate the use of the proposed Bayesian semi-parametric models for the degradation analysis of locomotive wheels.

3.1 Degradation Data

The data were collected by a Swedish railway company from November 2010 to January 2012. We use the degradation data from two heavy haul cargo trains' locomotives (denoted as locomotive 1 and locomotive 2). Correspondingly, there are two studied groups, and $n = 2$. For each locomotive, see Fig.1, there are two bogies (*incl.*, Bogie I, Bogie II), and for each bogie, there are six wheels. The installed positions of the wheels on a particular locomotive are specified by the bogie number and are defined as covariates $\mathbf{x}$. The covariates' coefficients are represented by $\boldsymbol{\beta}$. More specifically, $x = 1$ represents the wheel mounted in Bogie I, while $x = 2$ represents the wheel mounted in Bogie II. β_1 is the coefficient and β_0 is defined as natural variability.

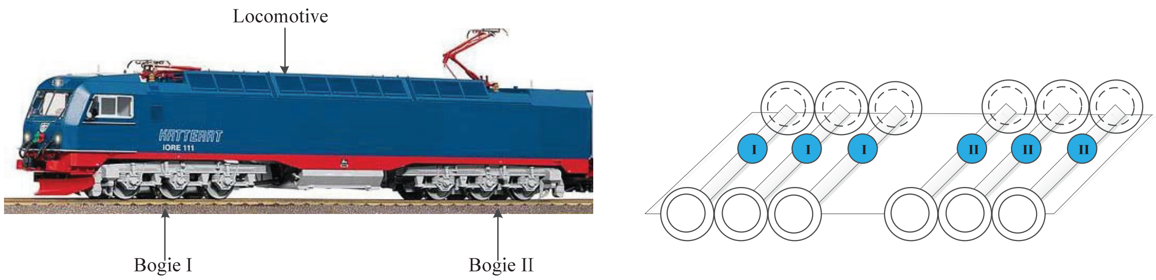


Fig.1 Wheel positions specified in this study

The diameter of a new locomotive wheel is 1250 mm. In the company's current maintenance strategy, a wheel's diameter is measured after running a certain distance. If it is reduced to 1150 mm, the wheel is replaced by a new one. Otherwise, it is re-profiled or other maintenance strategies are implemented. Therefore, a threshold level for failure, denoted as y_0 , is defined as 100 mm (y_0 = 1250 mm -1150 mm). The wheel's failure condition is assumed to be reached if the diameter reaches y_0 . The company's complete dataset includes the diameters of all locomotive wheels at a given inspection time, the total running distances corresponding to their "time to be maintained (re-profiled or replaced)", and the wheels' bill of material (BOM) data, from which we can determine their positions.

3.2 Degradation Path and Lifetime Data

From the dataset, we can obtain 5 to 6 measurements of the diameter of each wheel during its lifetime. By connecting these measurements, we can determine a degradation trend (e.g., in Fig.2, the blue line). In their analyses of train wheels, most studies (Freitas *et al.* 2009, 2010) assume a linear degradation path (see the black dotted line in Fig.2). However, in our study, the results show that a better choice is a polynomial degradation path (see Fig.2, the purple dotted line). We plot the degradation data for one locomotive wheel in Fig.2. The squares of their correlation coefficients (denoted as R^2) are 0.9973 for a polynomial path and 0.9271 for a linear path, indicating that a polynomial degradation path is better than a linear degradation path for the wheels studied.

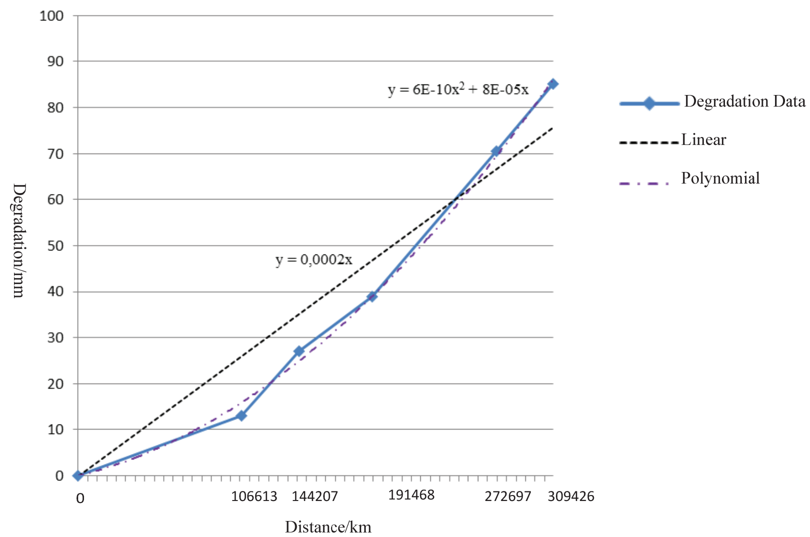


Fig.2 Plot of the wheel degradation data: an example

Take locomotive 1 for example; for all wheels installed on this locomotive, the assumptions of the polynomial function are supported by the statistics shown in Table.1. In particular, the average value for R^2 with a polynomial path is around 0.9943; for the linear path, it is around 0.8922. This supports the assumption of a multinomial degradation path.

Table.1 Statistics on degradation path and lifetime data: locomotive 1

Number	Positions	Polynomial path	R²-polynomial	Linear path	R²-linear	Lifetime (× 1000 km)
1	Bogie I	$y = 6E\text{-}10x^2 + 8E\text{-}05x$	0.9973	y = 0.0002x	0.9271	347
2	Bogie I	$y = 6E\text{-}10x^2 + 8E\text{-}05x$	0.9974	y = 0.0002x	0.9274	347
3	Bogie I	$y = 7E\text{-}10x^2 + 6E\text{-}05x$	0.9981	y = 0.0002x	0.9109	338
4	Bogie I	$y = 7E\text{-}10x^2 + 6E\text{-}05x$	0.9982	y = 0.0002x	0.9102	338
5	Bogie I	$y = 7E\text{-}10x^2 + 7E\text{-}05x$	0.9986	y = 0.0002x	0.9211	354
6	Bogie I	$y = 7E\text{-}10x^2 + 7E\text{-}05x$	0.9986	y = 0.0002x	0.9215	354
7	Bogie II	$y = 1E\text{-}09x^2 + 4E\text{-}06x$	0.9960	y = 0.0002x	0.8485	*314
8	Bogie II	$y = 1E\text{-}09x^2 + 4E\text{-}06x$	0.9960	y = 0.0002x	0.8485	*314
9	Bogie II	$y = 1E\text{-}09x^2 - 4E\text{-}06x$	0.9964	y = 0.0002x	0.8419	*314
10	Bogie II	$y = 1E\text{-}09x^2 - 3E\text{-}06x$	0.9963	y = 0.0002x	0.8430	*315
11	Bogie II	$y = 7E\text{-}10x^2 + 7E\text{-}05x$	0.9792	y = 0.0002x	0.9039	331
12	Bogie II	$y = 7E\text{-}10x^2 + 7E\text{-}05x$	0.9805	y = 0.0003x	0.9027	331
Average		/	0.9943	/	0.8922	/

* Right-censored data

Note: some lifetime data are right-censored (denoted by asterisk in Table.1). However, we know the real lifetimes will exceed the predicted lifetimes.

Table.2 shows the results of the same test for the wheels on locomotive 2. Again, a polynomial degradation path is a better choice.

Following the above discussion, a wheel's failure condition is assumed to be reached if the diameter reaches y_0. We adopt the polynomial path for all wheels and set $y_0 = y$. The lifetimes for these wheels are now easily determined and are shown in the last columns of Table1 and Table 2.

Table.2 Statistics on degradation path and lifetime data: locomotive 2

Number	Positions	Polynomial path	R^2-polynomial	Linear path	R^2-linear	Lifetime (× 1000 km)
1	Bogie I	y = 8E-10x^2 + 0.0002x	0.9807	y = 0.0003x	0.9579	250
2	Bogie I	y = 8E-10x^2 + 0.0002x	0.9817	y = 0.0003x	0.9597	250
3	Bogie I	y = 8E-10x^2 + 0.0002x	0.9805	y = 0.0003x	0.9590	250
4	Bogie I	y = 8E-10x^2 + 0.0002x	0.9798	y = 0.0003x	0.9589	250
5	Bogie I	y = 7E-10x^2 + 0.0002x	0.9790	y = 0.0003x	0.9624	261
6	Bogie I	y = 7E-10x^2 + 0.0002x	0.9792	y = 0.0003x	0.9624	261
7	Bogie II	y = 9E-10x^2 + 0.0002x	0.9751	y = 0.0003x	0.9491	240
8	Bogie II	y = 1E-09x^2 + 0.0002x	0.9750	y = 0.0003x	0.9484	232
9	Bogie II	y = 1E-09x^2 + 0.0002x	0.9709	y = 0.0003x	0.9420	232
10	Bogie II	y = 1E-09x^2 + 0.0002x	0.9689	y = 0.0003x	0.9415	232
11	Bogie II	y = 9E-10x^2 + 0.0002x	0.9727	y = 0.0003x	0.9498	240
12	Bogie II	y = 9E-10x^2 + 0.0002x	0.9726	y = 0.0003x	0.9495	240
Average		/	0.9763	/	0.9534	/

3.3 Parameter Configuration

It is clear that a very small k will make the model nonparametric. However, if k is too small, estimates of the baseline hazard rate will be unstable, and if k is too large, a poor model fit could result (Ibrahim *et al.,* 2001). In our study, determining the degradation path requires us to make 5 to 6 measurements for each locomotive wheel; in other words, the lifetime data are based on the data acquired at 5 to 6 different inspections. Following the reasoning above, we divide the time axis into 6 sections piecewise. In our case study, no predicted lifetime exceeds 360,000 kilometres. Therefore, k =6, and each interval is equal to 60,000km. We get 6 intervals (0, 60 000], (60 000, 120 00]… (300 000, 360 000].

For convenience, we let $\lambda_k = \exp(b_k)$, and vague prior distributions are adopted here as the following:

- Gamma frailty prior: $\omega_i \sim Ga(\kappa^{-1}, \kappa^{-1})$
- Normal prior distribution: $b_k \sim N(b_{k-1}, \kappa)$
- Normal prior distribution: $b_1 \sim N(0, \kappa)$
- Gamma prior distribution: $\kappa \sim Ga$ (0.0001, 0.0001);
- Normal prior distribution: $\beta_0 \sim N$ (0.0, 0.001);
- Normal prior distribution: $\beta_1 \sim N$ (0.0, 0.001).

At this point, the MCMC calculations are implemented with the software WinBUGS (Spiegelhalter *et al.,* 2003). A burn-in of 10,001 samples is used, with an additional 10,000 Gibbs samples.

3.4 Results

Following the convergence diagnostics (*incl.*, checking dynamic traces in Markov chains, time series, and comparing the Monte Carlo (MC) error with Standard Deviation (SD); see Spiegelhalter *et al.,* 2003), we consider the following posterior distribution summaries (Table.3). Statistics summaries include the parameters' posterior distribution mean, SD, MC error, and the 95% highest posterior distribution density (HPD) interval.

In Table.3, $\beta_1 > 0$ means that the wheels mounted in the first bogie (as $x = 1$) have a shorter lifetime than those in the second (as $x = 2$). However, the influence could possibly be reduced as more data are obtained in the future, because the 95% HPD interval includes 0 point. Because $\kappa > 0.5$, there is a positive relationship between the wheels mounted on the same locomotive; in addition, the heterogeneity among the locomotives is significant. Meanwhile, $\omega_1 < 1$ suggests that the predictive lifetimes for those wheels mounted on the first locomotive are longer than if the frailties are not considered; in fact, $\omega_2 > 1$ indicates the opposite conclusion.

Table.3 Posterior Distribution Summaries

Parameter	mean	SD	MC error	95% HPD Interval
β_0	-10.39	2.888	0.2622	(-16.61, -4.79)
β_1	0.3293	0.4927	0.02016	(-0.661, 1.271)
κ	0.563	0.269	0.01038	(0.1879, 1.225)
ω_1	0.1441	0.1374	0.004822	(0.01192, 0.5258)
ω_2	1.866	1.016	0.03628	(0.3846, 4.308)
b_1	0.1361	1.595	0.1037	(-3.196, 3.364)
b_2	0.758	2.182	0.1672	(-3.7, 5.248)
b_3	1.94	2.514	0.2105	(-3.126, 7.342)
b_4	4.447	2.668	0.2389	(-0.5652, 10.48)
b_5	6.342	2.684	0.2415	(1.126, 12.29)
b_6	8.159	2.724	0.2417	(2.843, 14.15)

Baseline hazard rate statistics based on the above results are shown in Table.4 and Fig.3. At the fourth piecewise interval, the wheels' baseline hazard rate increases dramatically.

Table.4 Baseline Hazard Rate Statistics

Piecewise	1	2	3	4	5	6
Intervals($\times$ 1000km)	(0, 60]	(60, 120]	(120, 180]	(180, 240]	(240, 300]	(300, 360]
λ_k	1.15	2.13	6.96	85.37	567.93	3494.69

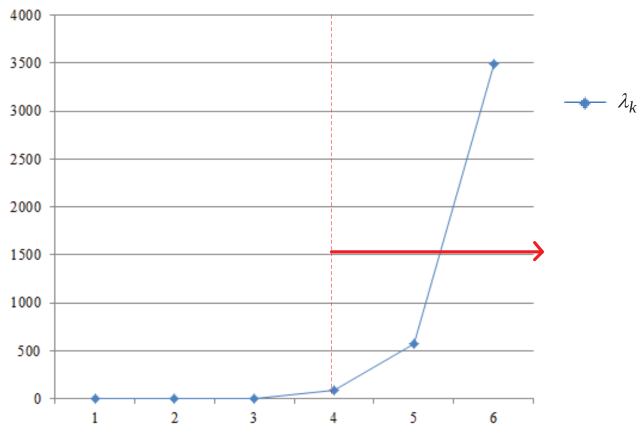


Fig.3 Plot of Baseline Hazard Rate

By considering the random effects resulting from the natural variability (explained by covariates) and from the unobserved random effects within the same group (explained by frailties), we can determine other reliability characteristics of lifetime distribution. The statistics on reliability $R(t)$ and cumulative hazard rate $\Lambda(t)$ for the two wheels mounted in different bogies are listed in Table.4, Fig.4 and Fig.5.

Table.4 Reliability and Cumulative hazard statistics

Distance (1000 km)	Reliability $R(t)$				Cumulative hazard $\Lambda(t)$			
	Locomotive 1		Locomotive 2		Locomotive 1		Locomotive 2	
	Bogie I	Bogie II	Bogie I	Bogie II	Bogie I	Bogie II	Bogie I	Bogie II
60	0.999577	0.999412	0.994534	0.99241	0.000184	0.000256	0.00238	0.003309
120	0.998425	0.997811	0.97979	0.97202	0.000685	0.000952	0.008867	0.012325
180	0.992318	0.989338	0.90496	0.870393	0.003349	0.004655	0.04337	0.060285
240	0.881485	0.839169	0.195241	0.103252	0.054785	0.076151	0.709428	0.986101
300	0.350289	0.232678	1.26E-06	6.31E-09	0.455574	0.633245	5.899379	8.200106
360	0.000433	2.11E-05	2.75E-44	2.82E-61	3.363977	4.67591	43.56128	60.54995

Fig.4 and Fig.5 show frailties between Locomotive 1 and Locomotive 2. In addition, for these locomotives, the wheels mounted in the first bogie ($x = 1$) have lower reliability and a higher cumulative hazard rate than those mounted in the second one ($x = 2$).

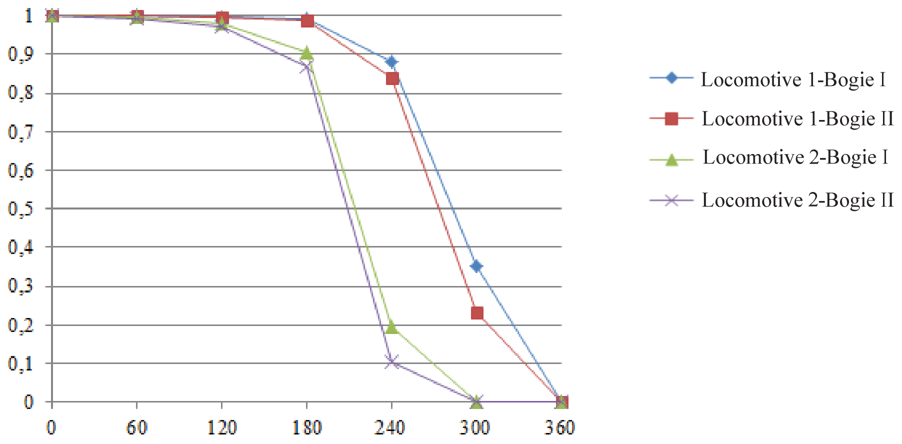


Fig.4 Plot of the reliabilities for Locomotive 1 and Locomotive 2

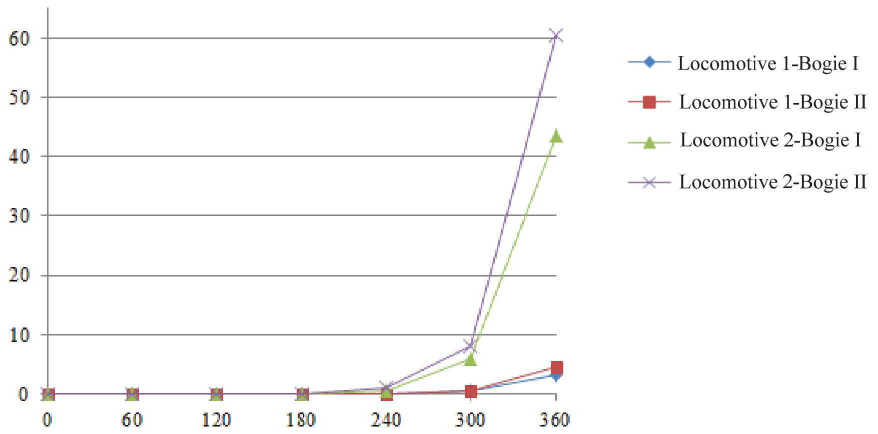


Fig.5 Plot of the Cumulative hazard for Locomotive 1 and Locomotive 2

In addition, Fig.4 and Fig.5 show change points in the wheels. For example, the reliability declines sharply at the fourth piecewise interval, and at the fifth piecewise interval, the cumulative hazard increases dramatically.

3.5 Discussions

The above results can be applied to maintenance optimisation, including lifetime prediction and replacement optimisation, preventative maintenance optimisation, and re-profiling optimisation.

First, determining reliability characteristics distributed over the wheels' lifetime could be used to optimise replacement strategies. The results could also support related predictions for spares inventory.

Second, the change points (Fig.3, Fig.4, Fig.5) appearing in the fourth and fifth piecewise interval (from 180 000 to 300 000 kilometres) indicate that after running about 180 000 kilometres, the locomotive wheel has a high-risk of failure. Rolling contact fatigue (RCF) problems could start at the fifth interval (after 240 000 kilometres). Therefore, special attention should be paid if the wheels have run longer than these change points. In addition, because re-profiling may leave cracks over time and reduce the wheel's lifetime, cracks should be checked after re-profiling to improve the lifetime.

Third, the wheels installed in the first bogie should be given more attention during maintenance. Especially when the wheels are re-profiled, they should be checked starting with the first bogie to avoid duplication of efforts. Note that in the studied company, the wheels' inspecting sequences are random; this means that the first checked wheel could belong in the second bogie. After the second checked wheel is lathed or re-profiled, if the diameter is shorter than predicted, the first checked wheel might need to be lathed or re-profiled again. Therefore, starting with the wheel installed in the first bogie could improve maintenance effectiveness.

Last but not least, the frailties between locomotives could be caused by the different operating environments (e.g., climate, topography, and track geometry), configuration of the suspension, status of the bogies or spring systems, operation speeds and the applied loads. Specific operating conditions should be considered when designing maintenance strategies because even if the locomotives and wheel types are the same, the lifetimes and operating performance could differ.

5 Conclusions

This paper proposes a Bayesian semi-parametric framework to analyse a locomotive wheel's reliability using degradation data. The piecewise constant hazard rate is used to establish the distribution of the wheels' lifetime. The gamma shared frailties $\boldsymbol{\omega}_i$ are used to explore the influence of unobserved covariates within the same locomotive. By introducing covariate $\mathbf{x}_i$'s linear function $\mathbf{x}_i^{'}\boldsymbol{\beta}$, the influence of the bogie in which a wheel is installed can be taken into account. The proposed framework can deal with small and incomplete datasets and simultaneously consider the influence of different covariates. The MCMC technique is used to integrate high-dimensional probability distributions to make inferences and predictions about model parameters.

The results of the case study suggest that a polynomial degradation path for the wheels is better than a liner degradation path. The wheels' lifetimes differ according to where they are installed (in which bogie they are mounted) on the locomotive. The wheel installed in the second bogie has a longer lifetime than the one in the first bogie. The differences could be influenced by the real running situation (e.g. topography), and the locomotive's centre of gravity. The gamma frailties help with exploring the unobserved covariates and thus improve the model's precision. Results also indicate a close positive relationship between the wheels mounted on the same locomotive; the heterogeneity between locomotives is also significant. We can determine the wheel's reliability characteristics, including the baseline hazard rate $\lambda(t)$, reliability $R(t)$, and cumulative hazard rate $\Lambda(t)$, *etc*. The results also indicate the existence of change points. As Fig.3, Fig.4 and Fig.5 show, wheel reliability declines sharply at the fourth piecewise interval, while at the fifth piecewise interval, the cumulative hazard increases dramatically. The results allow us to evaluate and optimise wheel replacement and maintenance strategies (including the re-profiling interval, inspection interval, lubrication interval, depth and optimal sequence of re-profiling, and so on).

Finally, the approach discussed in this paper can be applied to cargo train wheels or to other technical problems (e.g. other industries, other components).

We suggest the following additional research:

- The covariates considered in this paper are limited to locomotive wheels' installed positions; more covariates must be considered. To this end, we will study such factors as operating environment (e.g., climate, topography, and track geometry), configuration of the suspension, status of the bogies and the spring systems, operation speeds and the applied loads, *etc.*

- We have chosen vague prior distributions for the case study. Other prior distributions, including both informative and non-information prior distributions, should be studied.
- In subsequent research, we plan to consider using our results to optimize maintenance strategies and the related LCC (Life Cycle Cost) problem considering maintenance costs, particularly with respect to different maintenance inspection levels and inspection periods (long term, medium term and short term).

Acknowledgements

The authors would like to thank Luleå Railway Research Centre (Järnvägstekniskt Centrum, Sweden) for initiating the research study and Swedish Transport Administration (Trafikverket) for providing financial support.

References

1. Aslanidou H, Dey D K, Sinha D. Bayesian analysis of multivariate survival data using Monte Carlo methods. *Canadian Journal of Statistics*. 1998, 26: 33-48
2. Bernasconi A, *et al.*. An Integrated Approach to Rolling Contact Sub-surface Fatigue assessment of Railway Wheels. *Journal of Wear*. 2005. 258: 973-980
3. Braghin F, *et al.*. A Mathematical Model to Predict Railway Wheel Profile Evolution Due to Wear. *Journal of Wear*. 2006. 261: 1253-1264
4. Breslow N E. Covariance analysis of censored survival data. *Biometrics*.1974,30:80-99
5. Clayton D G, Cuzick J. Multivatiate Generalizations of the Proportional Hazards Model (with Discussion). *Journal of the Royal Statistical Society A*. 1985, 148: 82-117
6. Clayton D G. A Model for Association in Bivariate Life Tales and its Application in Epi-demiological Studies of Familial Tendency in Chronic Disease Incidence. *Biometicka*. 1978, 65: 141-151
7. Clayton D G. A Monte Carlo method for Bayesian inference in frailty models. *Biometrics*. 1991, 47:467-485
8. Clayton P. Tribological Aspects of Wheel-Rail Contact: A Review of Recent Experimental Research. *Journal of Wear*. 1996. 191: 170-183
9. Congdon P. Applied Bayesian Modelling. England: John Wiley and Sons.2003
10. Congdon P. Bayesian Statistical Modelling. England: John Wiley and Sons.2001
11. Crowder M. A multivariate distribution with Weibull connections. *Journal of the Royal Statistical Society B*. 1989, 51: 93-107
12. Donato P, *et al.*. Design and Signal Processing of A Magnetic Sensor Array for Train Wheel Detection. *Journal of Senors and Actuators A*. 2006. 132: 516-525
13. Freitas M A, *et al.*. Reliability assessment using degradation models: Bayesian and classical approaches. *Pesquisa Operacional*. 2010,30 (1):195-219
14. Freitas M A, *et al.*. Using Degradation Data to Assess Reliability: A Case Study on Train Wheel Degradation. *Journal of Quality and Reliability Engineering International*. 2009, 25: 607-629
15. Hakon K G, Odd O A, Nilslid H. Frailty models based on levy processes. *Advances in Applied Probability*.

2003, 35: 532-550

16. Hougaard P. Frailty models for survival data. *Lifetime Data Analysis*.1995, 1:255-273
17. Hougaard P. Survival models for heterogeneous populations derived from stable distributions. *Biometrika*. 1986, 73: 387-396
18. Ibrahim J G, Chen M H, Sinha D. Bayesian Survival Analysis. New York: Berlin Heidelberg, 2001
19. Lin J, Asplund M, Parida A. Bayesian Parametric Analysis for Reliability Study of Locomotive Wheels. *Conference Proceedings*. The 60 th Annual Reliability and Maintainability Symposium (RAMS® 2013). January 28-31, Orlando, FL, USA. Accepted
20. Johansson A, Andersson C. Out-of-round Railway Wheels - a Study of Wheel polygonalization through Simulation of Three-dimensional Wheel-Rail Interaction and Wear. *Journal of Vehicle System Dynamics*. 2005, 43(8):539-559
21. Kalbfleisch J D, Prentice R L. Marginal likelihood's based on Cox's regression and life model. *Biometrika*. 1973, 60:267-278
22. Keiding N, Anderson P, John J. The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. *Statistics in Medicine*, 1997, 16:215-225
23. Klein J P, Moeschberger M L. Survival Analysis: Techniques for Censored and Truncated Data. Springer-Verlag New York, Inc.1997
24. Lawless J F. Statistical Models and Methods for Lifetime Data. John Wiley and Sons.1982
25. Liu Y M, *et al.*. Multiaxial Fatigue Reliability Analysis of Railroad Wheels. *Journal of Reliability Engineering and System Safety*. 2008. 93:456-467
26. McGilchrist C A, Aisbett C W. Regression with Frailty in Survival Analysis. *Biometrics*. 1991, 47:461-466
27. Palo M, *et al.*. Rolling stock condition monitoring using wheel/rail forces. *Journal of Insight-Non-Destructive Testing and Condition Monitoring*. 2012. 54 (8): 451-455
28. Palo M. Condition Monitoring of Railway Vehicles: A Study on Wheel Condition for Heavy Haul Rolling Stock. Licentiate Thesis. Luleå University of Technology, Sweden. 2012
29. Pennell M L, Dunson D B. Bayesian Semi parametric Dynamic Frailty Models for Multiple Event Time Data. *Biometrika*. 2006, 62:1044-1052
30. Pombo J, Ambrosio J, Pereira M. A Railway Wheel Wear Prediction Tool based on A Multibody Software. *Journal of Theoretical and Applied Mechanics*. 2010. 48, 3:751-770
31. Qiou Z, Ravishanker N, Dey D K. Multivariate survival analysis with positive frailties. *Biometrika*. 1999, 55: 637-644
32. Sahu S K, Dey D K, Aslanidou H, Sinaha D. A Weibull regression model with gamma frailties for multivariate survival data. *Lifetime Data Analysis*. 1997, 3:123-137
33. Sinha D. Semiparameteric Bayesian analysis of multiple event time data. *Journal of the American Statistical Association*. 1993, 88: 979-983
34. Skarlatos D, Karakasis K, Trochidis A. Railway Wheel Fault Diagnosis Using A Fuzzy-logic Method. *Journal of Applied Acoustics*. 2004. 65:951-966
35. Spiegelhalter D, *et al.*. WinBUGS User Manual (Version 1.4). January, 2003. http://www.mrc-

bsu.cam.ac.uk/bugs

36. Stratman B, Liu Y, Mahadevan S. Structural Health Monitoring of Railroad Wheels Using Wheel Impact Load Detectors. *Journal of Failure Analysis and Prevention.* 2007. 7(3):218-225
37. Tassini N, *et al.*. A Numerical Model of Twin Disc Test Arrangement for the Evaluation of Railway Wheel Wear Prediction Methods. *Journal of Wear*. 2010. 268: 660-667
38. Yang C, Letourneau S. Learning to Predict Train Wheel Failures. *Conference Proceedings*. The 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005). Chicago, Illinois, USA.

Paper IV

A Comparison Study for Locomotive Wheels' Reliability Assessment using the Weibull Frailty Model

- Lin Jing, Asplund Matthias.
- Revised Manuscript has been submitted to Journal of Rail and Rapid Transit
- 2013, April.

A Comparison Study for Locomotive Wheels' Reliability Assessment using the Weibull Frailty Model

Jing Lin, Matthias Asplund
Division of Operation and Maintenance Engineering, Luleå University of Technology, 97187, Luleå, Sweden

Abstract: The service life of railroad wheels can differ significantly depending on their installed position, operating conditions, re-profiling characteristics, etc. This paper compares the wheels on two selected locomotives on the Iron Ore Line in northern Sweden to explore some of these differences. It proposes integrating reliability assessment data with both degradation data and re-profiling performance data. Its case study compares: 1) degradation analysis using a Weibull frailty model; 2) work orders for re-profiling; 3) the performance of re-profiling parameter; and 4) wear rates. The results show that for the two locomotives: 1) under the specified installation position and operation conditions, the Weibull frailty model is a useful tool to determine wheel reliability; 2) rolling contact fatigue (RCF) is the principal reason for re-profiling work orders; 3) the re-profiling parameters can be applied to monitor both the wear rate and the re-profiling loss; 4) the total wear of the wheels can be investigated by considering natural wear and re-profiling loss separately, but natural wear and re-profiling loss differ depending on the locomotive and the operating conditions; and 5) the bogie in which a wheel is installed influences wheel reliability.

Keywords: reliability analysis; locomotive wheels; frailty; re-profiling; wear; Markov Chain Monte Carlo

1 Introduction

The service life of different railroad wheels can vary greatly. Take a Swedish railway company, for example. For the wheels of its 26 locomotives, statistics show that from 2010 to 2011, the longest mean time between re-profiling was around 59 000 kilometres and the shortest was about 31 000 kilometres. The large difference can be attributed to the non-heterogeneous nature of the wheels; each differs according to its installed position, operating conditions, re-profiling characteristics, *etc.* [1-5].

One common preventive maintenance strategy (used in our study) is re-profiling wheels after they run a certain distance. Re-profiling reduces the wheel's diameter; once the diameter is reduced to a pre-specified length, the wheel is replaced by a new one. Seeking to optimise this maintenance strategy, researchers have examined wheel degradation data to determine wheel reliability and failure distribution [1-3]. However, most studies cannot solve the combined problem of small data samples and incomplete datasets while simultaneously considering the influence of several covariates [5].

In addition, most reliability studies are implemented under the assumption that individual lifetimes are independent and identically distributed (*i.i.d*). In reality, sometimes Cox proportional hazard (CPH) models cannot be used because of the dependence of data within a group. For instance, because they have the same operating conditions, the wheels mounted on a particular bogie may be dependent. Modelling dependence in multivariate survival data has received considerable attention, especially in cases where the datasets comprise inter-related subjects of the same group [6, 7]. A key development in modelling such data is to consider frailty models, in which the data are conditionally independent. When frailties are considered, the dependence within subgroups can be considered an unknown and unobservable risk factor (or explanatory variable) of the hazard function. In this paper, we consider a gamma shared frailty, first discussed by Clayton [8] and Oakes [9] and later developed by Sahu *et al*. [6], to explore the unobserved covariates' influence on the wheels on the same bogie. We also adopt the Weibull hazard model to determine the distribution of the wheels' lifetime; the validity of this model has been established by Lin *et al.* [5].

Besides the degradation analysis, re-profiling information is a key source of data to evaluate the wheels' performance. As Fröhling and Hettasch [10] note, the "loss of material during re-profiling because of hollow or flange wear" is a significant element in the integrated data processing of the wheel-rail interface management. Even so, related studies remain limited.

To fill this gap in the literature, this paper compares the wheels on two selected locomotives on the Iron Ore Line in northern Sweden, taking an integrated data approach to reliability assessment by considering both degradation data and re-profiling data.

The remainder of the paper is organised as follows. Section 2 describes the background of the comparison study, by introducing the Iron Ore Line, as well as the degradation data and re-profiling parameters for the locomotive wheels being studied, along with their operating conditions. Section 3 presents the degradation analysis using a Weibull frailty model; the analysis considers the wheels' location in the bogies and their operating conditions as covariates and uses Markov Chain Monte Carlo (MCMC) methods. Sections 4 to 6 comprise the comparison study; the three sections compare the re-profiling work orders, the specified re-profiling parameters (the wheel diameters, the flange thickness, the radial run-out, and the lateral run-out), and the wear rate of the wheels, respectively. Each section is accompanied by a discussion. Section 7 offers conclusions and makes suggestions for future study.

2 Study Background

This section gives background information on the Iron Ore Line. It also introduces the degradation data and the re-profiling parameters for the locomotive wheels being studied, along with their operating conditions.

2.1 Iron Ore Line (Malmbanan)

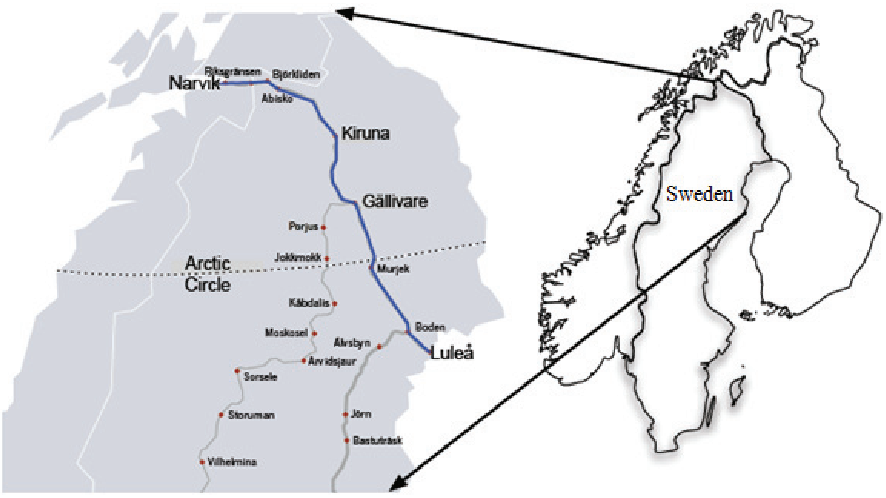


Fig. 1 Geographical location of Iron Ore Line (Malmbanan)

The Iron Ore Line (Malmbanan) is the only existing heavy haul line in Europe; it stretches 473 kilometres and has been in operation since 1903. As Fig. 1 shows, it is mainly used to transport iron ore and pellets from the mines in Kiruna (also Malmberget, close to Kiruna, in Sweden) to Narvik Harbour (Norway) in the northwest and Luleå Harbour (Sweden) in the southeast. The track section on the Swedish side is owned by the Swedish government and managed by Trafikverket (Swedish Transport Administration), while the iron ore freight trains are owned and managed by the freight operator (a Swedish company). Each freight train consists of two IORE locomotives accompanied by 68 wagons with a maximum length of 750 metres and a total train weight of 8500 metric tonnes with axel loads of 30 tonnes. The trains operate in harsh conditions, including snow in the winter and extreme temperatures ranging from - 40 °C to + 25 °C. Because carrying iron ore results in high axle loads and there is a high demand for a constant flow of ore/pellets, the track and wagons must be monitored and maintained on a regular basis. The condition of the locomotive wheel profile is one of the most important aspects to consider.

2.2 Degradation data and re-profiling parameters

We use the degradation data from two selected heavy haul cargo locomotives (denoted as locomotive 1 and locomotive 2), collected from October 2010 to January 2012. The selection criteria are discussed in Section 2.3. Each locomotive is studied separately, and $n=2$. For each locomotive, see Fig.2, there are two bogies (*incl.*, Bogie I, Bogie II); and each bogie contains six wheels. The installed position of a wheel on a particular locomotive is specified by the bogie number (I, II-number of bogies on the locomotive), an axel number (1, 2, 3-number of axels for each bogie) and the position of the axle (right or left) where each wheel is mounted.

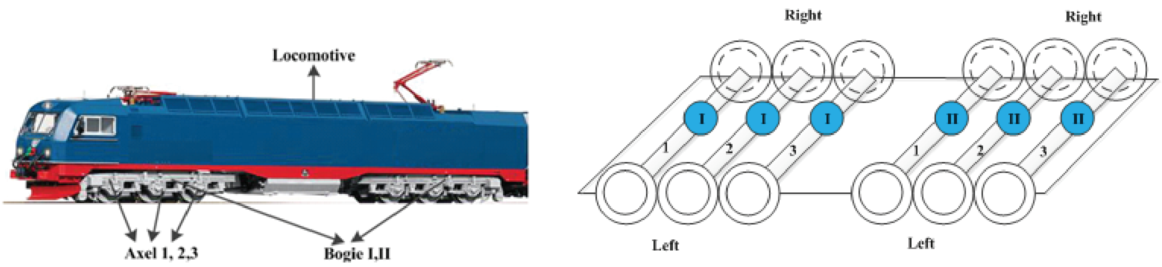
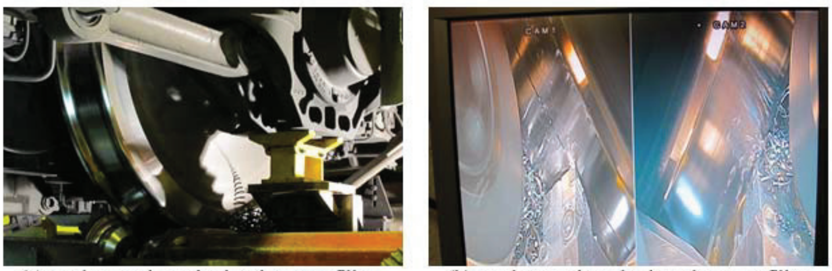


Fig. 2 Wheel positions specified in this study

The diameter of a new locomotive wheel in this study is about 1250 mm. Following the current maintenance strategy, a wheel's diameter is measured after it runs a certain distance. If it is reduced to 1150 mm, the wheel set is replaced by a new one. Otherwise, it is re-profiled (see Fig.3). Therefore, a threshold level for failure, denoted as y_0, is defined as 100 mm (y_0 = 1250 mm -1150 mm). The wheel's failure condition is assumed to be reached if the diameter reaches y_0. The dataset includes the diameters of all locomotive wheels at a given inspection time, the total running distances corresponding to their "mean time between re-profiling", and the wheels' bill of material (BOM) data, from which we can determine their positions.



(a) one locomotive wheel under re-profiling (viewed on site)

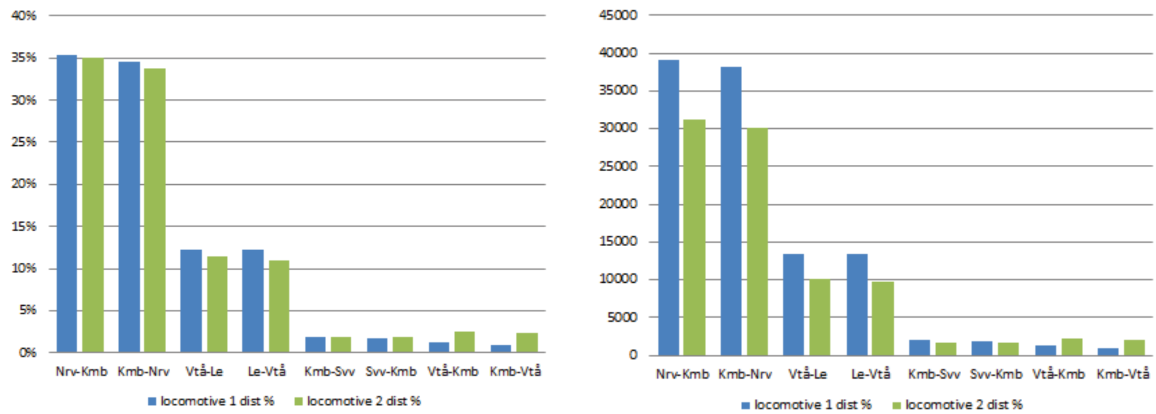
(b) two locomotive wheels under re-profiling (viewed from the monitoring screen)

Fig. 3 Locomotive wheels on-site re-profiling

During the re-profiling process, the re-profiling parameters include but are not limited to: 1) the diameters of the wheels; 2) the flange thickness; 3) the radial run-out; 4) the lateral run-out.

2.3 Comparison of the operating conditions

In this study, both locomotive 1 and locomotive 2 are operating on the Iron Ore Line (Malmbanan). In Fig.4, the horizontal axis represents the different working intervals; “Nrv-Kmb” represents the route from Narvik to Kiruna, while “Kmb-Nrv” represents the route from Kiruna to Narvik. Those intervals make up the whole Iron Ore Line (Malmbanan). The longitudinal axis of Fig. 4 (a) represents the percentage of workloads in each working interval; the longitudinal axis of Fig. 4 (b) represents the running distances in those intervals.



(a) workloads in different working intervals (b) running distances for different working intervals

Fig. 4 Comparison of the operating conditions for locomotive 1 and locomotive 2

As seen in Fig.4, during the period in question (from October 2010 to January 2012), the total running distance for locomotive 1 is 101035 kilometres and for locomotive 2, 81302 kilometres. About 70% of the locomotives’ workload is between Narvik and Kiruna. There is no substantial difference between the running routes, but it seems that locomotive 1 works harder than locomotive 2, because the former runs 24% farther. As there is not a big difference between the topographies, we assume that the only difference in operating conditions is the total running distance.

3 Degradation analyses with the Weibull frailty model

In this section, we propose the Weibull frailty model for analysing the wheels’ degradation data, using a MCMC computation scheme.

Before continuing, it should be pointed that Lin *et al.* [5] have used the Bayesian Exponential Regression Model, Bayesian Weibull Regression Model (easily transferred to an Extreme-Value Regression Model) and Bayesian Lognormal Regression Model, separately, to analyze the lifetime of locomotive wheels using degradation data and taking into account the position of the wheel. Their results show that “the performance of the Weibull Regression Model is close to the Log-normal Regression Model, which could also be a suitable choice under specified situations.” As the Weibull Regression Model is more acceptable to engineers and the differences between the Weibull Regression and Lognormal Regression Models are quite small, we choose the former model in this comparative study.

3.1 Weibull frailty model

Most reliability studies are implemented under the assumption that individual lifetimes are independent and identically distributed (*i.i.d*). In reality, at times, Cox proportional hazard (CPH) models cannot be used because of the dependence of data within a group. For instance, because they have the same operating conditions, the wheels mounted on a particular bogie may be dependent. Modelling dependence in multivariate survival data has received considerable attention, especially in cases where the datasets comprise inter-related subjects of the same group [6, 7]. A key development in modelling such data is to consider frailty models, in which the data are conditionally independent.

Frailty models were first considered by Clayton [8] and Oakes [9] to handle multivariate survival data. In their models, the event times are conditionally independent according to a given frailty factor, which is an individual random effect. As discussed by Sahu *et al.*[6], the models formulate different variabilities and come from two distinct sources. The first source is natural variability, which is explained by the hazard function; the second is variability common to individuals of the same group or variability common to several events of an individual, which is explained by the frailty factor.

Assume the hazard function for the j^{th} individual in the i^{th} group is

$$h_{ij}(t) = h_0(t)\exp(\mu_i + \mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta}) . \tag{1}$$

In equation (1), μ_i represents the frailty parameter for the i^{th} group. If $\omega_i = \exp(\mu_i)$, the equation can also be written as

$$h_{ij}(t) = h_0(t)\omega_i \exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta}) . \tag{2}$$

Equation (1) is an additive frailty model, and equation (2) is a multiplicative frailty model. In both equations, μ_i and ω_i are shared by the individuals in the same group, and they are thus referred to as shared-frailty models and actually are extensions of the CPH model.

To this point, discussions of frailty models have focused on the forms of 1) the baseline hazard function and 2) the frailty's distribution. Representative studies related to the former include the gamma process for the accumulated hazard function [11, 12], Weibull baseline hazard rate [6], and the piecewise constant hazard rate [7] which is adopted in this paper due to its flexibility. Some researchers have examined finite mean frailty distributions, including gamma distribution [8, 13], lognormal distribution [14], and the like; others have studied non-parameter methods, including the inverse Gaussion frailty distribution [15], the power variance function for frailty [16], the positive stable frailty distribution [17, 18], the Dirichlet process frailty model [19] and the Levy process frailty model [20]. In this paper, we consider the gamma shared frailty model, the most popular model for frailty.

From equation (2), suppose the frailty parameters ω_i are independent and identically distributed (*i.i.d*) for each group and follow a gamma distribution, denoted by $Ga(\kappa^{-1}, \kappa^{-1})$. Therefore, the probability density function can be written as

$$f(\omega_i) = \frac{(\kappa^{-1})^{\kappa^{-1}}}{\Gamma(\kappa^{-1})} \cdot \omega_i^{\kappa^{-1}-1} \exp(-\kappa^{-1}\omega_i) \ . \tag{3}$$

In equation (3), the mean value of ω_i is 1, where κ is the unknown variance of ω_i s. Greater values of κ signify a closer positive relationship between the subjects of the same group as well as greater heterogeneity among groups. Furthermore, as ω_i >1, the failures for the individuals in the corresponding group will appear earlier than if ω_i =1; in other words, as ω_i <1, the predicted lifetimes will be greater than those found in the independent models.

Suppose $\boldsymbol{\omega} = (\omega_1, \omega_2, \cdots, \omega_n)'$; then

$$\pi(\boldsymbol{\omega}|\kappa) \propto \prod_{i=1}^{n} \omega_i^{\kappa^{-1}-1} \exp(-\kappa^{-1}\omega_i) \quad . \tag{4}$$

Denote the j^{th} individual in the i^{th} group as having lifetime $\mathbf{t_{ij}} = (t_{11}, t_{12}, \cdots, t_{nm_i})'$, where $i = 1, \cdots, n$ and $j = 1, \cdots, m_i$. Suppose the j^{th} individual in the i^{th} group has a 2-parameter Weibull distribution $W(\alpha, \gamma)$, where $\alpha > 0$ and $\gamma > 0$. Then, the p.d.f. is $f(t_{ij}|\alpha,\gamma) = \alpha\gamma t_{ij}^{\alpha-1}\exp(-\gamma t_{ij}^{\alpha})$, and the c.d.f. $F(t_{ij}|\alpha,\gamma)$ and the reliability function $R(t_{ij}|\alpha,\gamma)$ are $F(t_{ij}|\alpha,\gamma) = 1 - \exp(\gamma t_{ij}^{\alpha}) = 1 - R(t_{ij}|\alpha,\gamma)$. Meanwhile, the hazard rate function can be written as

$$h_0(t_{ij}|\alpha,\gamma) = \gamma\alpha t_{ij}^{\alpha-1} \quad . \tag{5}$$

Based on the above discussions (equation (2), (3), and (5)), the Weibull frailty model with gamma shared frailties can be written as

$$h(t_{ij}|\mathbf{x_{ij}}, \omega_i) = \gamma\alpha\omega_i t_{ij}^{\alpha-1}\exp(\mathbf{x'_{ij}}\boldsymbol{\beta}) \quad . \tag{6}$$

In equation (6), $\omega_i \sim Ga(\kappa^{-1}, \kappa^{-1})$.

In reliability analyses, the lifetime data are usually incomplete, and only a portion of the individual lifetimes are known. Right-censored data are often called Type I censoring, and the corresponding likelihood construction problem has been extensively studied in the literature [21, 22]. Suppose the j^{th} individual in the i^{th} group has lifetime T_{ij} and censoring time L_{ij}. The observed lifetime $t_{ij} = \min(T_{ij}, L_{ij})$; therefore, the exact lifetime T_{ij} will be observed only if $T_{ij} \le L_{ij}$. In addition, the lifetime data involving right censoring can be represented by n pairs of random variables (t_{ij}, υ_{ij}), where $\upsilon_{ij} = 1$ if $T_{ij} \le L_{ij}$ and $\upsilon_{ij} = 0$ if $T_{ij} > L_{ij}$. This means that υ_{ij} indicates whether lifetime T_{ij} is censored or not. The likelihood function is deduced as

$$L(t) = \prod_{i=1}^{n}\prod_{j=1}^{m_i} [f(t_{ij})]^{\upsilon_{ij}} R(t_{ij})^{1-\upsilon_{ij}} \quad . \tag{7}$$

If we denote $\lambda_{ij} = \exp(\mathbf{x'_{ij}}\boldsymbol{\beta} + \log\gamma + \log\omega_i)$, equation (6) becomes $h(t_{ij}|\lambda_{ij},\alpha) = \lambda_{ij}\alpha t_{ij}^{\alpha-1}$, the Weibull regression model with a gamma frailty $W(\alpha, \lambda_{ij})$.

If we denote the model's dataset as $D = (n, \omega, \mathbf{t}, \mathbf{X}, \boldsymbol{\upsilon})$, following equation (7), the complete likelihood function $L(\boldsymbol{\beta}, \gamma, \alpha|D)$ for the individuals in the i^{th} group can be written as

$$L(\boldsymbol{\beta},\gamma,\alpha|D) \propto (\gamma\alpha t_{ij}^{\alpha-1}\omega_i \exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta}))^{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\upsilon_{ij}} \exp(-\sum_{i=1}^{n}\sum_{j=1}^{m_i}\gamma t_{ij}^{\alpha} \exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta})\omega_i) . \tag{8}$$

Let $\pi(\cdot)$ denote the prior or posterior distributions for the parameters. Then, the joint posterior distribution $\pi(\omega_i|\boldsymbol{\beta},\alpha,\gamma,D)$ for gamma frailties ω_i can be written as

$$\begin{aligned}\pi(\omega_i|\boldsymbol{\beta},\alpha,\gamma,D) &\propto L(\boldsymbol{\beta},\gamma,\alpha|D)\times\pi(\boldsymbol{\omega}|\kappa)\\ &\propto (\gamma\alpha t_{ij}^{\alpha-1}\omega_i \exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta}))^{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\upsilon_{ij}} \exp(-\sum_{i=1}^{n}\sum_{j=1}^{m_i}\gamma t_{ij}^{\alpha} \exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta})\omega_i)\times\prod_{i=1}^{n}\omega_i^{\kappa^{-1}-1}\exp(-\kappa^{-1}\omega_i)\\ &\propto \omega_i^{\kappa^{-1}+\sum_{j=1}^{m_i}\upsilon_{ij}-1} \exp\{-(\kappa^{-1}+\gamma\sum_{j=1}^{m_i}t_{ij}^{\alpha}\exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta}))\omega_i\}\\ &\sim Ga\{\kappa^{-1}+\sum_{j=1}^{m_i}\upsilon_{ij},\kappa^{-1}+\gamma\sum_{j=1}^{m_i}t_{ij}^{\alpha}\exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta})\}\end{aligned} \tag{9}$$

Equation (9) shows that the full conditional density of each ω_i is a gamma distribution.

Before continuing, we should discuss the selection of priors based on MCMC in our study. In Bayesian reliability inference, two kinds of priors are very useful: the conjugate prior and the non-informative prior. To apply MCMC methods, however, the "log-concave prior" is recommended.

The conjugate prior family is very popular, because it is convenient for mathematical calculation. If the posterior distributions are in the same family as the prior distributions, the prior and posterior distributions are called conjugate distributions, and the prior is called a conjugate prior. The Gaussian family is a conjugate of itself (or a self-conjugate) with respect to a Gaussian likelihood function: if the likelihood function is Gaussian, choosing a Gaussian prior distribution over the mean distribution will ensure that the posterior distribution is also Gaussian. This means that the Gaussian distribution is a conjugate prior for the likelihood function which is also Gaussian. Other examples include the following: the conjugate distribution of a Normal distribution is a Normal or inverse-Normal distribution; the Poisson and the Exponential distributions' conjugates both have a Gamma distribution, while the Gamma distribution is a self-conjugate; the Binomial and the negative Binomial distributions' conjugates both have a Beta distribution; the Polynomial distribution's conjugate is a Dirichlet distribution, etc. Non-informative prior refers to a prior for which we only know certain parameters' value ranges or their importance; for example, there may be a uniform distribution. A non-informative prior can also be called a vague prior, flat prior, diffuse prior, or ignorance prior, etc. There are many different ways to determine the distribution of a non-informative prior, including Bayes hypothesis, Jeffrey's rule, reference prior, inverse reference prior, probability-matching prior, Maximum entropy prior, relative likelihood approach, cumulative distribution function, Monte Carlo method, bootstrap method, random weighting simulation method, Harr invariant measurement, Laplace prior, Lindley rule, generalized maximum entropy principle, and the use of marginal distributions. From another perspective, the types of prior distribution also include informative prior, hierarchical prior, Power prior and non-parameter prior processes.

At this point, there are no set rules for selecting prior distributions. Regardless of the manner used to determine a prior's distribution, the selected prior should be both reasonable and convenient for calculation. Of the above, the conjugate prior is a common choice. To facilitate the calculation of MCMC, especially for adaptive rejection sampling and Gibbs sampling, a popular choice, as noted above, is log-concave prior distribution. Log-concave prior distribution refers to a prior distribution in which the natural logarithm is concave, i.e. the second derivative is non-positive. Common logarithmic concavity prior distributions include the normal distribution family, logistic distribution, student's t distribution, the exponential distribution family, the uniform distribution on a finite interval, greater than the gamma distribution with a shape parameter greater than 1, Beta distribution with a value interval (0 , 1), etc. As logarithmic concavity prior distributions are very flexible, this paper recommends their use in reliability studies.

Suppose γ has a gamma prior distribution, denoted by $\gamma \sim Ga(\rho_1, \rho_2)$. The full conditional density of γ is

$$
\begin{aligned}
\pi(\gamma|\boldsymbol{\beta},\alpha,\omega,D) &\propto L(\boldsymbol{\beta},\gamma,\alpha|D)\times\pi(\gamma|\rho_1,\rho_2) \\
&\propto (\gamma\alpha t_{ij}^{\alpha-1}\omega_i \exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta}))^{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\upsilon_{ij}} \exp(-\sum_{i=1}^{n}\sum_{j=1}^{m_i}\gamma t_{ij}^{\alpha}\exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta})\omega_i)\times\prod_{i=1}^{n}\gamma^{\rho_1-1}\exp(-\rho_2\gamma) \\
&\propto \gamma^{\rho_1+\sum_{i=1}^{n}\sum_{j=1}^{m_i}\upsilon_{ij}-1} \exp\{-(\rho_2+\sum_{i=1}^{n}\sum_{m=1}^{n_i}t_{ij}^{\alpha}\exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta})\omega_i)\gamma\} \\
&\sim Ga\{\rho_1+\sum_{i=1}^{n}\sum_{j=1}^{m_i}\upsilon_{ij},\rho_2+\sum_{i=1}^{n}\sum_{m=1}^{n_i}t_{ij}^{\alpha}\exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta})\omega_i\}
\end{aligned}
\tag{10}
$$

Equation (10) also shows that the full conditional density of γ is a gamma distribution. The full conditional density of $\boldsymbol{\beta}$ and α can be given by

$$
\pi(\boldsymbol{\beta}|\eta,\omega,\gamma,D)\propto\exp\{\boldsymbol{\beta}^{'}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\upsilon_{ij}x_{ij}-\gamma\sum_{i=1}^{n}\sum_{m=1}^{n_i}t_{ij}^{\alpha}\exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta})\omega_i\}\times\pi(\boldsymbol{\beta}) \tag{11}
$$

$$
\pi(\alpha|\boldsymbol{\beta},\gamma,\omega,D)\propto(\prod_{i=1}^{n}\prod_{j=1}^{m_i}t_{ij}^{\upsilon_{ij}})^{\alpha-1}\alpha^{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\upsilon_{ij}}\exp\{-\gamma\sum_{i=1}^{n}\sum_{m=1}^{n_i}t_{ij}^{\alpha}\exp(\mathbf{x}_{\mathbf{ij}}^{'}\boldsymbol{\beta})\omega_i\}\times\pi(\alpha)\,. \tag{12}
$$

3.2 Comparison study of degradation analyses

3.2.1 Degradation path and lifetime data

From the dataset, we obtain 5 to 6 measurements of the diameter of each wheel during its lifetime (in the period October 2010 to January 2012). By connecting these measurements, we can determine a degradation trend. In their analyses of train wheels, most studies (e.g., [2], [3], [5]) assume a linear degradation path. In this study, we apply a linear degradation path for each locomotive wheel. In Figure 1 we plot the degradation data for one locomotive wheel (I1H) in Fig.5 as an example. And in Table.1, the squares of their correlation coefficients (denoted as R^2) indicate that the linear assumption is a reasonable choice. Finally, the corresponding running distance (kilometres) is recognized as the lifetime. Note that the degradation path could be non-linear for other cases. We suggest comparing the R-square results in each case.

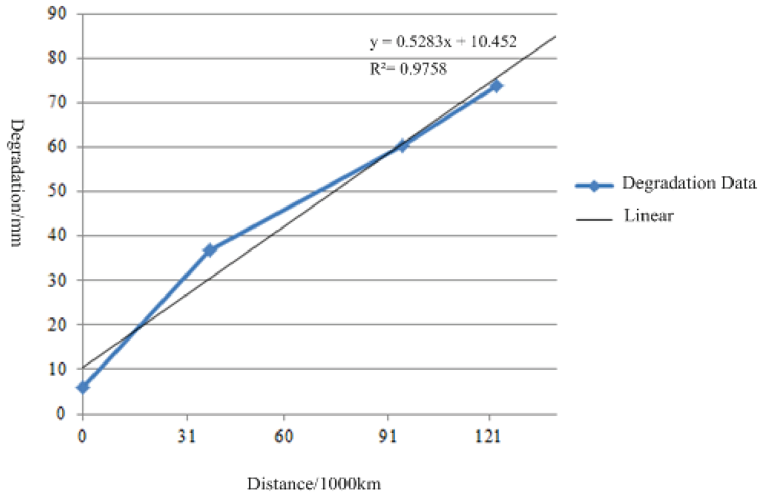


Fig.5 Plot of the wheel degradation data: an example (I1H)

Table.1 Statistics on degradation path and lifetime data

	Locomotive 1			Locomotive 2		
Positions	**Linear path**	**R²**	**Lifetime****	**Linear path**	**R²**	**Lifetime****
Bogie I	y = 0.5283x + 10.452	0.9758	* 169.50	y = 0.3195x + 16.67	0.9020	260.81
	y = 0.5174x + 10.867	0.9707	* 172.27	y = 0.319x + 16.717	0.9050	261.08
	y = 0.5293x + 10.306	0.9756	* 169.46	y = 0.3185x + 16.554	0.9039	262.00
	y = 0.528x + 10.327	0.9755	* 169.84	y = 0.318x + 16.673	0.9028	262.03
	y = 0.5581x + 10.528	0.9658	160.32	y = 0.3165x + 16.953	0.9073	262.39
	y = 0.5637x + 10.447	0.9701	158.87	y = 0.3168x + 16.901	0.9076	262.31
Bogie II	y = 0.4736x + 1.5766	0.9401	207.82	y = 0.3502x + 16.292	0.8972	239.03
	y = 0.4735x + 1.605	0.9424	207.80	y = 0.3511x + 16.198	0.8967	238.68
	y = 0.4994x + 1.5624	0.9537	197.11	y = 0.3861x + 11.828	0.8096	228.37
	y = 0.5002x + 1.5388	0.9553	196.84	y = 0.3831x + 12.323	0.8030	228.86
	y = 0.4979x + 4.0938	0.9702	192.62	y = 0.3425x + 17.353	0.8890	241.31
	y = 0.4926x + 3.8123	0.9738	195.27	y = 0.3427x + 17.34	0.8887	241.20

* Right-censored data; ** $\times 10^3$ kilometres

Note: some lifetime data are right-censored (denoted by the asterisk in Table.1). However, we know the real lifetimes will exceed the predicted lifetimes.

Following the above discussion, a wheel's failure condition is assumed to be reached if the diameter reaches y_0. We adopt the linear path for all wheels and set $y_0 = y$. The lifetimes for these wheels are now easily determined and are shown in the "Lifetime" columns of Table1. These lifetimes are the input of the Weibull frailty model, as discussed in Section 3.

3.2.2 Parameter Configuration

Following the discussion in 3.2.1, vague prior distributions are adopted in this paper as:

- Gamma frailty prior: $\omega_i \sim Ga$ (0.1, 0.1);
- Normal prior distribution: $\beta_0 \sim N$ (0.0, 0.001);
- Normal prior distribution: $\beta_1 \sim N$ (0.0, 0.001);
- Gamma prior distribution: $\alpha \sim Ga(0.1, 0.1)$;
- Gamma prior distribution: $\gamma \sim Ga(0.1, 0.1)$.

At this point, the MCMC calculations are implemented using the software WinBUGS [13]. We use a burn-in of 10,001 samples, along with an additional 10,000 Gibbs samples.

3.2.3 Results

Following the convergence diagnostics (*incl.*, checking dynamic traces in Markov chains, time series, and comparing the Monte Carlo (MC) error with Standard Deviation (SD); see [23]), we consider the following posterior distribution summaries (see Table.3): the parameters' posterior distribution mean, SD, MC error, and the 95% highest posterior distribution density (HPD) interval.

In Table.3, $\beta_1 < 0$ means that the wheels mounted in the first bogie (as $x = 1$) have a shorter lifetime than those in the second (as $x = 2$). However, the influence could possibly be reduced as more data are obtained in the future, because the 95% HPD interval includes a 0 point. In addition, the heterogeneity of the wheels on the two locomotives is significant. Nevertheless, $\omega_1 < 1$ suggests that the predictive lifetimes for the wheels mounted on the first locomotive are shorter when the frailties are considered; however, $\omega_2 > 1$ indicates the opposite conclusion.

Table.3 Posterior distribution summaries

Parameter	mean	SD	MC error	95% HPD Interval
β_0	-0.2826	31.36	0.3438	(-60.69,61.4)
β_1	-0.1613	31.62	0.3152	(-63.2,62.71)
α	1.035	3.329	0.03348	(2.449E-16,10.36)
γ	0.9726	3.101	0.02904	(1.683E-15,9.277)
ω_1	0.9709	2.999	0.02819	(1.738E-16,9.442)
ω_2	1.029	3.261	0.03392	(9.718E-16,10.21)

By considering the random effects resulting from the natural variability (explained by covariates) and the unobserved random effects within the same group (explained by frailties), we can determine other reliability characteristics of lifetime distribution. The statistics on reliability $R(t)$ for the two wheels mounted in different bogies are:

- $h(t_1) = 0.97 \times 1.035 \times 0.9709 \times t_1^{0.035} \exp(-0.2826 \times (-0.1613x))$
- $h(t_2) = 0.97 \times 1.035 \times 1.029 \times t_2^{0.035} \exp(-0.2826 \times (-0.1613x))$

3.3 Discussion

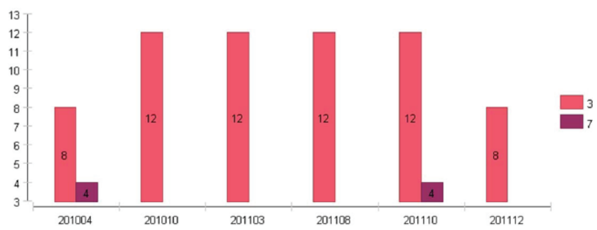
The above results can be applied to maintenance optimisation, including lifetime prediction and replacement, preventative maintenance, and re-profiling. More specifically, determining reliability characteristics distributed over the wheels' lifetime could be used to optimise replacement strategies and to support related predictions for spares inventory. With respect to preventative maintenance, the wheels installed in different bogies should be given more attention during maintenance. Especially when the wheels are re-profiled, they should be checked starting with the bogies to avoid duplication of efforts. Last but not least, as the operating environments are similar for the two locomotives considered here, the frailties between bogies could be caused by the locomotives themselves, the status of the bogies or spring systems, and human influences (including maintenance policies and the lathe operator).

4 Comparison study on re-profiling work orders

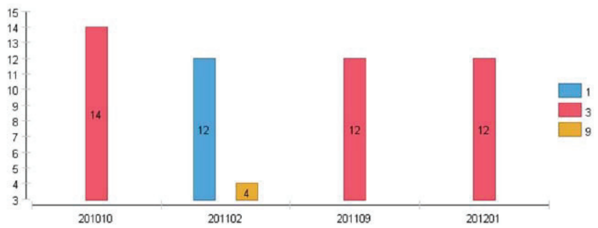
This section compares the work orders for wheel re-profiling by date (denoted as "by date" in Fig.6) and the corresponding bogies' total number of kilometres in operation (denoted as "by kilometres" in Fig.7), separately.

In Fig.6, the work order statistics for re-profiling are listed by date. The colour and the number of the bar represent the type of work order reported in the system. For instance, number 1 (blue) means the reason for re-profiling is a high flange; number 3 (red) represents the RCF problem; number 7 (purple) means the re-profiling is due to the dimension difference between wheels in a bogie; number 9 (yellow) denotes a thick flange. The work orders have 14 categories for re-profiling: high flange, thin flange, RCF, unbalanced wheel, QR measurements, out-of-round wheel, dimension difference in between wheels in same bogie, vibrations, thick flange, cracks, remarks from measurement of the wheel by Miniprof, other defects, to plant for re-profiling, and hollowware. These categories are determined by the operator and are listed in Appendix A. Take Fig.6 (a) for example. By April 2010, the wheels of Locomotive 1 have been re-profiled 12. Eight times it was related to category 3 (RCF problem), and four times it was in category 7 (the dimension difference between wheels in a bogie).

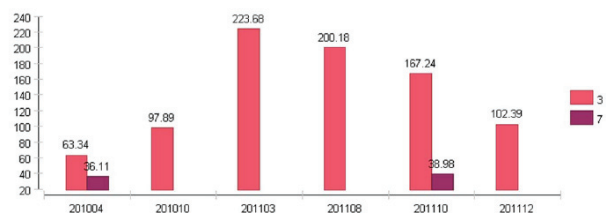
In Fig. 6 and Fig.7, the figures on the left side provide the statistics for locomotive 1, while those on the right are for locomotive 2. Note that in Fig.7, the work order statistics on re-profiling are listed by the corresponding bogies' total number of kilometres in operation on the reported date. In Fig.7 (b), the wheels have run 87721 kilometres and been re-profiled 16 times, 12 times due to category 1 (high flange) and 4 times due to category 9 (thick flange).



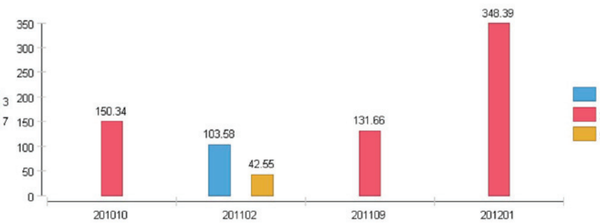
(a)Work order numbers for Locomotive 1



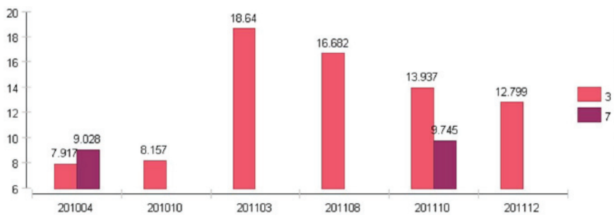
(b) Work order numbers for Locomotive 2



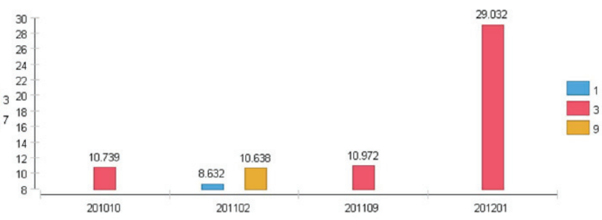
(c)Total re-profiling for Locomotive 1 /mm



(d) Total re-profiling for Locomotive 2 /mm



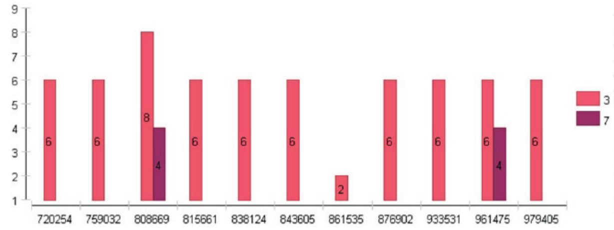
(e)Average re-profiling for Locomotive 1 /mm



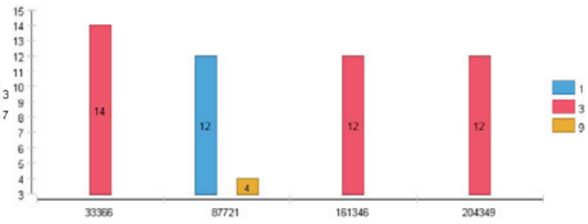
(f) Average re-profiling for Locomotive 2 /mm

Fig.6 Work order statistics on re-profiling by date

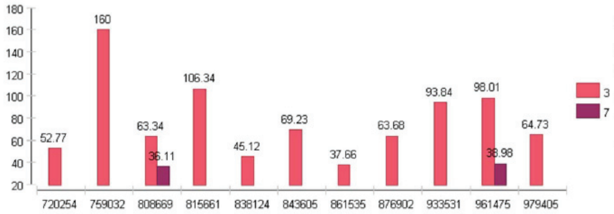
It should be pointed out that since October 2010, new wheels have been mounted on both locomotives. However, the selected work orders are from the beginning of 2010; therefore, more re-profiling has been done on locomotive 1.



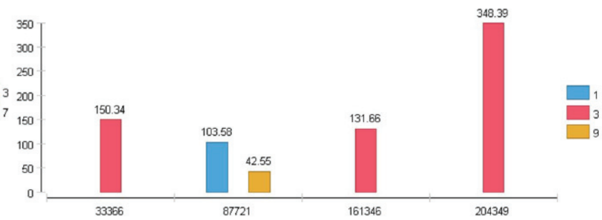
(a)Work order numbers for Locomotive 1



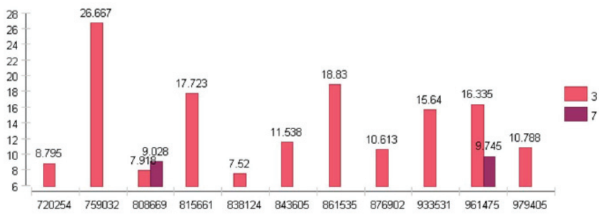
(b) Work order numbers for Locomotive 2



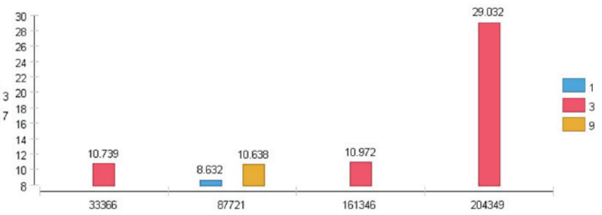
(c)Total re-profiling for Locomotive 1



(d) Total re-profiling for Locomotive 2



(e)Average re-profiling for Locomotive 1



(f) Average re-profiling for Locomotive 2

Fig.7 Work order statistics on re-profiling by kilometre

For locomotive 1, there are two failure modes: RCF and dimensional differences for wheels in the same bogie. The number of re-profiling work orders due to RCF is 64; the number due to dimensional differences for wheels in the same bogie is 8. Locomotive 2 shows three failure modes, high flange, RCF and thick flange. Again, the dominant failure mode is RCF with 38 re-profilings, followed by high flange with 12 re-profilings and thin flange with 4; see Fig.6 (b). Figs. 6 (c) and (d) show the amount of material removed at each re-profiling for all wheels. Even here, the RCF failure dominates with more material lost in re-profiling. Figs. 6 (e) and (f) show the mean cut deep for each re-profiling. The RCF failure mode has deeper cuts than other modes; the high flange failure mode has the smallest mean cut depth.

Fig. 7 shows the same information but uses the global traveling distance in kilometres (km). It should be pointed out that for Locomotive 1, Fig.7 has more bars on the left hand side because the axels have been changed and the recorded kilometres are different.

Generally speaking, RCF is the main type of work order for both locomotives. What should also be pointed out is that in the work order statistics, natural wear and the amount of re-profiling are considered simultaneously. Yet the trends in the amount of re-profiling are different. For instance, for locomotive 1, there is a decreasing trend for new wheels, while locomotive 2 shows an increasing trend.

During this investigation, we discovered a number of problems in the work orders. For example, some reported data cannot be recognised (e.g., some wheels are apparently re-profiled twice on one date; some reported wheel diameters after re-profiling are even larger than before re-profiling).

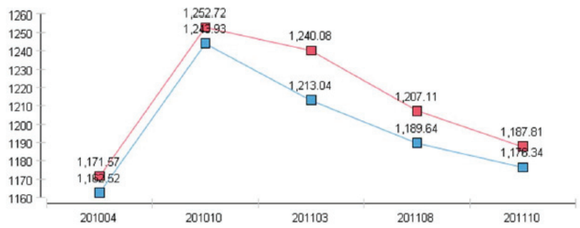
We suggest applying related KPIs to monitor the re-profiling work and the wheel performance in the future.

5 Comparison study on re-profiling parameters

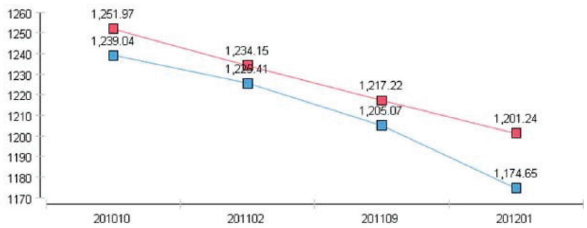
In this section, we compare the re-profiling parameters (the statistics before and after each re-profiling), including the diameter of the wheel (denoted as Rd), the flange thickness (denoted as Sd), the radial runout (denoted as Rr), and the axial runout (denoted as Rx).

5.1 Assessment of re-profiling parameters (Rd)

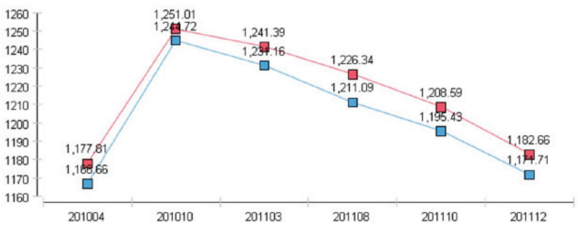
Starting in this section, we only include statistics by re-profiling date. In addition, due to the similarities of the wheels installed in the same bogie, we only list statistics for the chosen wheel within each bogie. The red line represents the statistics obtained before re-profiling; the blue line represents statistics after re-profiling. Fig.8 shows locomotive 1 on the left hand side and locomotive 2 on the right; for the graphs, the y-axle is the wheel diameter and the x-axle is the re-profiling date. For locomotive 1, the graphs start with the last re-profiling of an old wheel; step two is the first re-profiling with new wheels.



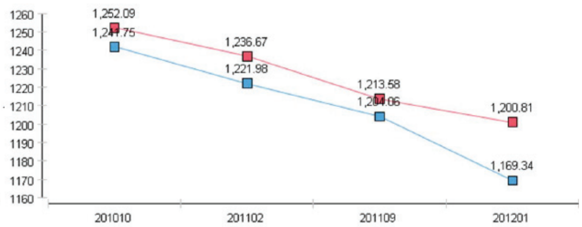
(a)Rd statistics for Locomotive 1 (I1H)



(b) Rd statistics for Locomotive 2 (I1H)



(c)Rd statistics for Locomotive 1 (II1H)



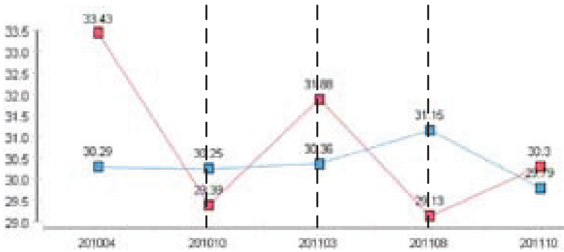
(d) Rd statistics for Locomotive 2 (II1H)

Fig.8 Rd statistics by date (before and after re-profiling): one example (I1H & II1H)

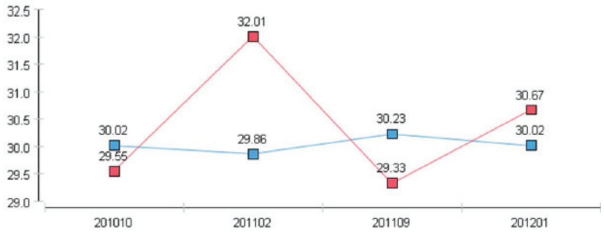
The wheels installed in the same bogie show similar trends in the before and after re-profiling statistics (denoted as Δ Rd). Δ Rd is decreasing for locomotive 1 and increasing for locomotive 2.

5.2 Assessment of re-profiling parameter (Sd)

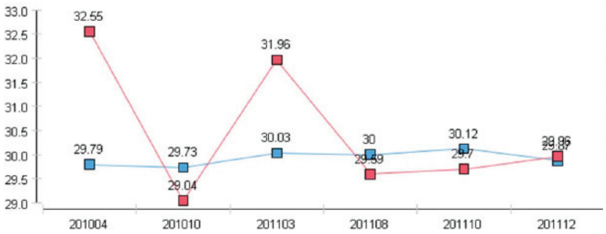
Fig.9 shows the statistics of the Sd for the selected wheels. Locomotive 1 is represented on the left hand side, with locomotive 2 on the rightt. For both, the flange thickness increases during winter and decreases in summer; this phenomenon is especially pronounced for locomotive 1 and the first bogie and first axle; see the dotted lines in Fig.9a.



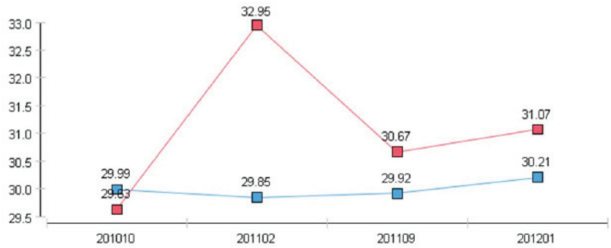
(a)Sd statistics for Locomotive 1 (I1H)



(b) Sd statistics for Locomotive 2 (I1H)



(c)Sd statistics for Locomotive 1 (II1H)



(d) Sd statistics for Locomotive 2 (II1H)

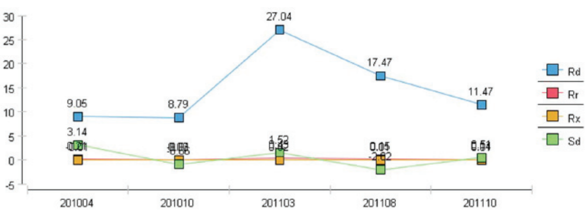
Fig.9 Sd statistics by date (before and after re-profiling): one example (I1H & II1H)

Like the Rd statistics, the Sd statistics for the wheels installed in the same bogie are quite similar. The “after” statistics (in blue) are stable. The “before” statistics (in red) are gradually becoming stable, which means the gap (denoted as Δ Sd) is decreasing.

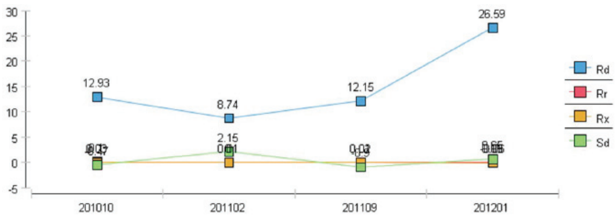
Note that if we check the before and after statistics in different seasons, we see that the flange thickness (red line) decreases in summer and increases in winter; see Fig.9 (a).

5.3 Assessment of re-profiling parameter (ΔRd, ΔSd, ΔRr, ΔRx)

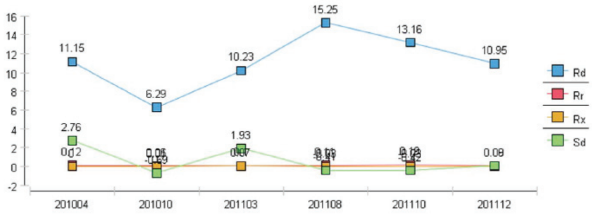
In this section, we simultaneously consider the gaps of the four parameters discussed above: ΔRd (blue), ΔSd (green), ΔRr (red), and ΔRx (yellow).



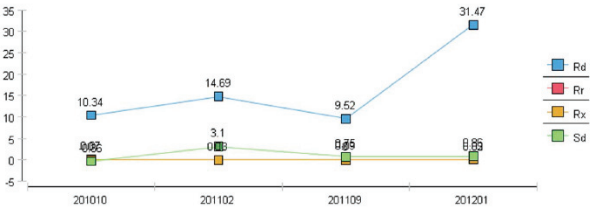
(a) Gap statistics for Locomotive 1 (I1H)



(b) Gap statistics for Locomotive 2 (I1H)



(c) Gap statistics for Locomotive 1 (II1H)



(d) Gap statistics for Locomotive 2 (II1H)

Fig.10 gap statistics by date (before and after re-profiling): one example (I1H & II1H)

As discussed above, the statistics for the wheels installed in the same bogie are quite similar. Among these four parameters, the changing of ΔRd is the most obvious one, with ΔSd coming second. The changing of ΔRr and ΔRx are random and the amount is quite small compared to the first two parameters. Therefore, we suggest applying the first two parameters to monitor the wheels' re-profiling performance in the future.

6 Comparison of wear rate

In this section, we compare the wheels' wear rates, shown in Tables.4 to 7. More details appear in Appendix B (Table B.1- B.20).

Table 4 shows locomotive 1, bogie 1 and the first axle on the right side; Table 5 shows locomotive 1, bogie 2 and the first axle on the right side; see Fig. 2 for the position of the bogies and axles. The number of re-profiling work orders is different between bogies: bogie 1 has 4 and bogie 2 has 5. The reason for the difference may be that bogie 1 was changed after the fourth re-profiling. The re-profiling at times 1 to 4 was done at the same time for both bogies, extending over 12 months.

As for locomotive 1, Table 4 shows that it has been running for 123.351 km; the mean distance between re-profiling is 41.117 km. The distance after the last re-profiling for bogie 2 was only 17.930 km, less than half of the average distance for re-profiling numbers 1 to 4; see Table 5. Tables 4 and 5 also show the diameter of the wheel before and after re-profiling and the amount of material removed at each re-profiling. The mean amount of material removed during re-profiling for bogie 1 is 16.193 mm and for bogie 2, 11.176 mm. Remarkably, the amount of re-profiling for bogie 2, step 2 is 27.04 mm, much more than the others; as noted above, the mean is 16.193 mm. If we compare natural wear with artificial wear, the former is between 15 mm and 20% of the total wear. In addition, the total wear rate for locomotive 2, bogie 1, is 0.619 mm/1000 km; for bogie 2, it is 0.393 mm/1000km.

Table.4 Statistics for wear rate: an example (locomotive 1, I1H)

Locomotive	1	Position	I1H		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201103	201108	201110	12 months
Reported kilometres /1000km	720.254	759.032	815.661	843.605	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	123.351
Diameters (before)/mm	1252.72	1240.08	1207.11	1187.81	/
Diameters (after)/mm	1243.93	1213.04	1189.64	1176.34	/
Re-profiling Amount/mm	8.79	27.04	17.47	11.47	64.77
Natural Wear/mm	0	3.85	5.93	1.83	11.61
Total Wear/mm	8.79	30.89	23.4	13.3	76.38
Re-profiling Amount %	1	0.875	0.747	0.862	0.848
Natural Wear %	0	0.125	0.253	0.138	0.152
WearRate_re-profiling	/	0.697	0.308	0.41	0.525
WearRate_Natural	/	0.099	0.105	0.065	0.094
WearRate_Total	/	0.797	0.413	0.476	0.619

Table.5 Statistics for wear rate: an example (locomotive 1, II1H)

Locomotive	1	Position	II1H			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres /1000km	838.124	876.902	933.531	961.475	979.405	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1251.01	1241.39	1226.34	1208.59	1182.66	/
Diameters (after)/mm	1244.72	1231.16	1211.09	1195.43	1171.71	/
Re-profiling Amount/mm	6.29	10.23	15.25	13.16	10.95	44.93
Natural Wear/mm	0	3.33	4.82	2.5	12.77	10.65
Total Wear/mm	6.29	13.56	20.07	15.66	23.72	55.58
Re-profiling Amount %	1	0.754	0.76	0.84	0.462	0.808
Natural Wear %	0	0.246	0.24	0.16	0.538	0.192
WearRate_re-profiling	/	0.264	0.269	0.471	0.611	0.318
WearRate_Natural	/	0.086	0.085	0.089	0.712	0.075
WearRate_Total	/	0.35	0.354	0.56	1.323	0.393

As mentioned, locomotive 1 and locomotive 2 have the same operating conditions (see Fig. 4 for the comparison), but the figures in Tables 6 and 7 show different results. Table 6 shows locomotive 2, the first bogie, the first axle, and the right hand side wheel; Table 7 shows the second bogie, the first axle, and the right hand side wheel. This locomotive has been re-profiled 4 times in 15 months; the mean distance between re-profiling is 56.990 km. The mean amount of material removed for re-profiling for bogie 1 is 15.10 mm; for bogie 2 it is 16.51 mm. The last re-profiling for the first bogie removed 26.59 mm and for the second bogie 31.47 mm. Finally, the total wear rate for locomotive 2, bogie 1, is 0.452 mm/1000 km and for bogie 2, 0.484 mm/1000km

Table.6 Statistics for wear rate: an example (locomotive 2, I1H)

Locomotive	2	Position	11H		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1251.97	1234.15	1217.22	1201.24	/
Diameters (after)/mm	1239.04	1225.41	1205.07	1174.65	/
Re-profiling Amount/mm	12.93	8.74	12.15	26.59	60.41
Natural Wear/mm	0	4.89	8.19	3.83	16.91
Total Wear/mm	12.93	13.63	20.34	30.42	77.32
Re-profiling Amount %	1	0.641	0.597	0.874	0.781
Natural Wear %	0	0.359	0.403	0.126	0.219
WearRate_re-profiling	/	0.161	0.165	0.618	0.353
WearRate_Natural	/	0.09	0.111	0.089	0.099
WearRate_Total	/	0.251	0.276	0.707	0.452

Table.7 Statistics for wear rate: an example (locomotive 2, II1H)

Locomotive	2	Position	II1H		Total/Average
Number	1	2	3	4	4 times
Date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1252.09	1236.67	1213.58	1200.81	/
Diameters (after)/mm	1241.75	1221.98	1204.06	1169.34	/
re-profiling Amount/mm	10.34	14.69	9.52	31.47	66.02
Natural Wear/mm	0	5.08	8.4	3.25	16.73
Total Wear/mm	10.34	19.77	17.92	34.72	82.75
re-profiling Amount %	1	0.743	0.531	0.906	0.798
Natural Wear %	0	0.257	0.469	0.094	0.202
WearSpeed_re-profiling	/	0.27	0.129	0.732	0.386
WearSpeed_Natural	/	0.093	0.114	0.076	0.098
WearSpeed_Total	/	0.364	0.243	0.807	0.484

Explanatory comments for Tables.4, 5, 6, 7 include the following:

- Absolute kilometres = the current reported kilometres – the previous reported kilometres;
- Re-profiling Amount = Diameters (before) - Diameters (after);
- Natural Wear = the previous Diameters (after) – the current Diameters (before);
- Total Wear = Re-profiling Amount + Natural Wear;
- Re-profiling Amount % = Re-profiling Amount / Total Wear;
- Natural Wear % = Natural Wear/ Total Wear;
- WearRate_Reprofiling = Re-profiling Amount / Absolute kilometres;
- WearRate_Natural = Natural Wear / Absolute kilometres;
- WearRate_Total = Total Wear / Absolute kilometres;
- Average of the total wear rate = the average of WearRate_Total.

In addition, by comparing the interval of the re-profiling date, we can simply divide each re-profiling episode into seasons (for instance, the summer and warmer times, the winter and cooler times).

In Table.8, we list the statistics for the WearRate_total of all the wheels for the two locomotives. The mean wear rates are 0.516 mm/1000km for locomotive 1 and 0.480 mm/1000km for locomotive 2; in other words, locomotive 1 has a 75% higher wear rate. Axles 1, 2 and 5 have 11.6 % higher wear rate than axles 3, 4 and 6.

Table.8 Statistics for total wear rates

	WearRate_total											
	11H	11V	12H	12V	13H	13V	21H	21V	22H	22V	23H	23V
Locomotive 1	0.619	0.607	0.614	0.605	0.542	0.533	0.393	0.404	0.467	0.467	0.467	0.472
Locomotive 2	0.452	0.439	0.448	0.448	0.449	0.448	0.484	0.482	0.568	0.575	0.487	0.476

By comparing the above parameters of the wheels installed in different positions on the locomotives, including the results shown in Tables 4 to 8, as well as the results shown in Appendix B (Table B.1 – B.20), we can reach the following additional conclusions:

- the average wear rate of the wheels on locomotive 1 is greater than for locomotive 2;
- the natural wear is about 10% ~ 25 % of the total wear; the re-profiling is about 75 %~ 90% of the total;
- the natural wear in winter time is slower than in summer;
- the re-profiling rate in winter is larger than in summer;
- the wheels installed on the second axel in the second bogie have an abnormal higher wear rate compared to the wheels installed in the same bogie but on the other axel; this requires more attention;
- The wheels installed in the same bogie perform similarly.

7 Conclusions

This paper compares the wheels of two selected locomotives on the Iron Ore Line in northern Sweden in an effort to explore their heterogeneity and their differences in reliability. To this end, it proposes integrating degradation data and re-profiling performance data to perform a reliability assessment.

The Weibull frailty model is used to analyse the wheels' degradation. The gamma shared frailties $\boldsymbol{\omega}_i$ are used to explore the influence of unobserved covariates within the same locomotive. By introducing covariate $\mathbf{x}_i$'s linear function $\mathbf{x}_i^{'}\boldsymbol{\beta}$, we can take into account the influence of the bogie in which a wheel is installed. The proposed framework can deal with small and incomplete datasets; it can also simultaneously consider the influence of various covariates. The MCMC technique is used to integrate high-dimensional probability distributions to make inferences and predictions about model parameters. Finally, we compare the statistics on re-profiling work orders, the performance of re-profiling parameters (denoted as ΔRd, ΔSd, ΔRr, ΔRx), and wear rates.

The results show the following for the two locomotives: 1) with the specified installation position and operating conditions, the Weibull frailty model is a useful tool to determine wheel reliability; 2) rolling contact fatigue (RCF) is the main type of re-profiling work order; 3) the re-profiling parameters can be applied to monitor both the wear rate and the re-profiling loss; 4) the total wear of the wheels can be determined by investigating natural wear and/or loss of wheel diameter through re-profiling loss, but these are different in different locomotives and under different operating conditions; 5) the bogie in which a wheel is installed is a key factor in assessing the wheel's reliability.

Finally, the approach discussed in this paper can be applied to cargo train wheels or to other technical problems (e.g. other industries, other components).

We suggest the following additional research:

- The covariates considered here are limited to the positions of the locomotive wheels; more covariates must be considered. For example, the braking forces and the curving forces should also be considered.
- We have chosen vague prior distributions for the case study. Other prior distributions, including both informative and non-informative prior distributions, should be studied.
- In subsequent research, we plan to use our results to optimise maintenance strategies and the related LCC (Life Cycle Cost) problem considering maintenance costs, particularly with respect to different maintenance inspection levels and inspection periods (long, medium and short term).

Acknowledgements

The authors would like to thank Luleå Railway Research Centre (Järnvägstekniskt Centrum, Sweden) and Swedish Transport Administration (Trafikverket) for initiating the research study and providing financial support. The authors would also like to thank Thomas Nordmark, Ove Salmonsson and Hans-Erik Fredriksson at LKAB for their support and their discussions of the locomotive wheels.

Reference

1. Yang C, Letourneau S. Learning to Predict Train Wheel Failures. In: The 11 *th* ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005). Chicago, Illinois, USA. August 21-22, 2005
2. Freitas MA, Colosimo EA, and Santos dos TR et al.. Reliability assessment using degradation models: Bayesian and classical approaches. *Pesquisa Operacional.* 2010; 30 (1):195-219
3. Freitas MA, Colosimo EA, and Santos dos TR et al.. Using Degradation Data to Assess Reliability: A Case Study on Train Wheel Degradation. *Journal of Quality and Reliability Engineering International*. 2009; 25: 607-629
4. Palo M, Schunnesson H, Kumar U, et al.. Rolling stock condition monitoring using wheel/rail forces. *Journal of Insight-Non-Destructive Testing and Condition Monitoring*. 2012; 54 (8): 451-455
5. Lin J, Asplund M, Parida A. Bayesian Parametric Analysis for Reliability Study of Locomotive Wheels. In:The 60 *th* Annual Reliability and Maintainability Symposium (RAMS® 2013). Orlando, FL, USA, January 28-31, 2003, paper no. 12E82.
6. Sahu SK, Dey DK, and Aslanidou H, et al.. A Weibull regression model with gamma frailties for multivariate survival data. *Lifetime Data Analysis*. 1997; 3:123-137
7. Aslanidou H, Dey DK, Sinha D. Bayesian analysis of multivariate survival data using Monte Carlo methods. *Canadian Journal of Statistics*. 1998; 26: 33-48
8. Clayton DG. A Model for Association in Bivariate Life Tales and its Application in Epi-demiological Studies of Familial Tendency in Chronic Disease Incidence. *Biometicka.* 1978; 65: 141-151
9. Oakes D. A model for association in bivariate survival data. Journal of Royal Statistics Society B. 1982; 44:414-422
10. Fröhling RD, Hettasch G. Wheel-rail interface management: a rolling stock perspective. Proceedings of the Institution of Mechanical Engineers. 2010; 224: 491-497
11. Clayton D G. A Monte Carlo method for Bayesian inference in frailty models. Biometrics. 1991, 47:467-485.
12. Sinha D. Semiparameteric Bayesian analysis of multiple event time data. Journal of the American Statistical Association. 1993, 88: 979-983.
13. Clayton D G, Cuzick J. Multivatiate Generalizations of the Proportional Hazards Model (with Discussion). Journal of the Royal Statistical Society A. 1985, 148: 82-117.
14. McGilchrist C A, Aisbett C W. Regression with Frailty in Survival Analysis. Biometrics. 1991, 47:461-466.
15. Hougaard P. Survival models for heterogeneous populations derived from stable distributions. Biometrika. 1986, 73: 387-396.
16. Crowder M. A multivariate distribution with Weibull connections. Journal of the Royal Statistical Society B. 1989, 51: 93-107.

17. Hougaard P. Frailty models for survival data. Lifetime Data Analysis.1995, 1:255-273.
18. Qiou Z, Ravishanker N, Dey D K. Multivariate survival analysis with positive frailties. Biometrika. 1999, 55: 637-644.
19. Pennell M L, Dunson D B. Bayesian Semi parametric Dynamic Frailty Models for Multiple Event Time Data. Biometrika. 2006, 62:1044-1052.
20. Hakon K G, Odd O A, Nilslid H. Frailty models based on levy processes. Advances in Applied Probability. 2003, 35: 532-550.
21. Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated Data. Springer-Verlag New York, Inc.1997, p.55.
22. Lawless J F. Statistical Models and Methods for Lifetime Data. John Wiley and Sons.1982, p.31
23. Spiegelhalter DJ, et al. Bayesian measures of model complexity and fit. *Journal of Royal Statist. Society Series B.* 2002; 64(3):583-639.

Appendix A

Table A.1 work order's categories

Code	Description	Code	Description	Code	Description
1	High flange	6	Out-of-round wheel	11	Measurements on the wheel, Miniprof
2	Thin flange	7	Dimension difference in between wheels in bogie	12	Other defect, pressure defect
3	RCF	8	Vibrations	13	Empty, no code
4	Unbalanced wheel	9	Thick flanges	14	Plant to be re-profiled
5	QR measurements	10	Cracks	15	Double flanges

Appendix B

Table B.1 Statistics for wear rate: locomotive 1. I1V

Locomotive	1	Position	I1V			Total/Average
Number of re-profiling	1	2	3	4	/	4 times
Re-profiling date	201010	201103	201108	201110	/	12 months
Reported kilometres /1000km	720.254	759.032	815.661	843.605	/	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	/	123.351
Diameters (before)/mm	1251.2	1241.18	1207.71	1187.39	/	/
Diameters (after)/mm	1243.88	1212.98	1188.83	1176.38	/	/
Re-profiling Amount/mm	7.32	28.2	18.88	11.01	/	65.41
Natural Wear/mm	0	2.7	5.27	1.44	/	9.41
Total Wear/mm	7.32	30.9	24.15	12.45	/	74.82
Re-profiling Amount %	1	0.913	0.782	0.884	/	0.874
Natural Wear %	0	0.087	0.218	0.116	/	0.126
WearRate_re-profiling	/	0.727	0.333	0.394	/	0.53
WearRate_Natural	/	0.07	0.093	0.052	/	0.076
WearRate_Total	/	0.797	0.426	0.446	/	0.607

Table B.2 Statistics for wear rate: locomotive 1. I2H

Locomotive	1	Position	I2H			Total/Average
Number of re-profiling	1	2	3	4	/	4 times
Re-profiling date	201010	201103	201108	201110	/	12 months
Reported kilometres /1000km	720.254	759.032	815.661	843.605	/	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	/	123.351
Diameters (before)/mm	1251.77	1241.42	1208.16	1188.77	/	/
Diameters (after)/mm	1244.06	1213.04	1190.09	1176.09	/	/
Re-profiling Amount/mm	7.71	28.38	18.07	12.68	/	66.84
Natural Wear/mm	0	2.64	4.88	1.32	/	8.84
Total Wear/mm	7.71	31.02	22.95	14	/	75.68
Re-profiling Amount %	1	0.915	0.787	0.906	/	0.883
Natural Wear %	0	0.085	0.213	0.094	/	0.117
WearRate_re-profiling	/	0.732	0.319	0.454	/	0.542
WearRate_Natural	/	0.068	0.086	0.047	/	0.072
WearRate_Total	/	0.8	0.405	0.501	/	0.614

Table B.3 Statistics for wear rate: locomotive 1. I2V

Locomotive	1	Position	I2V			Total/Average
Number of re-profiling	1	2	3	4	/	4 times
Re-profiling date	201010	201103	201108	201110	/	12 months
Reported kilometres /1000km	720.254	759.032	815.661	843.605	/	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	/	123.351
Diameters (before)/mm	1250.9	1241.76	1208.86	1188.44	/	/
Diameters (after)/mm	1244.05	1213.09	1190.08	1176.31	/	/
Re-profiling Amount/mm	6.85	28.67	18.78	12.13	/	66.43
Natural Wear/mm	0	2.29	4.23	1.64	/	8.16
Total Wear/mm	6.85	30.96	23.01	13.77	/	74.59
Re-profiling Amount %	1	0.926	0.816	0.881	/	0.891
Natural Wear %	0	0.074	0.184	0.119	/	0.109
WearRate_re-profiling	/	0.739	0.332	0.434	/	0.539
WearRate_Natural	/	0.059	0.075	0.059	/	0.066
WearRate_Total	/	0.798	0.406	0.493	/	0.605

Table B.4 Statistics for wear rate: locomotive 1. I3H

Locomotive	1	Position	I3H			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres /1000km	720.254	759.032	815.661	843.605	861.535	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1252.66	1235.89	1206.49	1185.39	1171.07	/
Diameters (after)/mm	1241.1	1213.05	1190.3	1176.09	1153.26	/
Re-profiling Amount/mm	11.56	22.84	16.19	9.3	17.81	59.89
Natural Wear/mm	0	5.21	6.56	4.91	5.02	16.68
Total Wear/mm	11.56	28.05	22.75	14.21	22.83	76.57
Re-profiling Amount %	1	0.814	0.712	0.654	0.78	0.782
Natural Wear %	0	0.186	0.288	0.346	0.22	0.218
WearRate_re-profiling	/	0.589	0.286	0.333	0.993	0.424
WearRate_Natural	/	0.134	0.116	0.176	0.28	0.118
WearRate_Total	/	0.723	0.402	0.509	1.273	0.542

Table B.5 Statistics for wear rate: locomotive 1. I3V

Locomotive	1	Position	I3V			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres /1000km	720.254	759.032	815.661	843.605	861.535	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1251.67	1237.74	1207.52	1187.95	1172.95	/
Diameters (after)/mm	1244.13	1212.87	1190.57	1176.31	1153.1	/
Re-profiling Amount/mm	7.54	24.87	16.95	11.64	19.85	61
Natural Wear/mm	0	6.39	5.35	2.62	3.36	14.36
Total Wear/mm	7.54	31.26	22.3	14.26	23.21	75.36
Re-profiling Amount %	1	0.796	0.76	0.816	0.855	0.809
Natural Wear %	0	0.204	0.24	0.184	0.145	0.191
WearRate_re-profiling	/	0.641	0.299	0.417	1.107	0.432
WearRate_Natural	/	0.165	0.094	0.094	0.187	0.102
WearRate_Total	/	0.806	0.394	0.51	1.294	0.533

Table B.6 Statistics for wear rate: locomotive 1. II1V

Locomotive	1	Position	II1V			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres /1000km	838.124	876.902	933.531	961.475	979.405	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1252.34	1240.81	1226.12	1209.16	1182.08	/
Diameters (after)/mm	1244.77	1231.11	1210.92	1195.29	1171.89	/
Re-profiling Amount/mm	7.57	9.7	15.2	13.87	10.19	46.34
Natural Wear/mm	0	3.96	4.99	1.76	13.21	10.71
Total Wear/mm	7.57	13.66	20.19	15.63	23.4	57.05
Re-profiling Amount %	1	0.71	0.753	0.887	0.435	0.812
Natural Wear %	0	0.29	0.247	0.113	0.565	0.188
WearRate_re-profiling	/	0.25	0.268	0.496	0.568	0.328
WearRate_Natural	/	0.102	0.088	0.063	0.737	0.076
WearRate_Total	/	0.352	0.357	0.559	1.305	0.404

Table B.7 Statistics for wear rate: locomotive 1. II2H

Locomotive	1	Position	II2H			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres /1000km	838.124	876.902	933.531	961.475	979.405	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1251.24	1241.86	1228.16	1194.45	1183.93	/
Diameters (after)/mm	1243.91	1231.35	1211.31	1185.28	1172.06	/
Re-profiling Amount/mm	7.33	10.51	16.85	9.17	11.87	43.86
Natural Wear/mm	0	2.05	3.19	16.86	1.35	22.1
Total Wear/mm	7.33	12.56	20.04	26.03	13.22	65.96
Re-profiling Amount %	1	0.837	0.841	0.352	0.898	0.665
Natural Wear %	0	0.163	0.159	0.648	0.102	0.335
WearRate_re-profiling	/	0.271	0.298	0.328	0.662	0.31
WearRate_Natural	/	0.053	0.056	0.603	0.075	0.156
WearRate_Total	/	0.324	0.354	0.932	0.737	0.467

Table B.8 Statistics for wear rate: locomotive 1. II2V

Locomotive	1	Position	II2V			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres /1000km	838.124	876.902	933.531	961.475	979.405	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1251.18	1241.53	1227.69	1194.44	1183.43	/
Diameters (after)/mm	1243.99	1231.34	1211.07	1185.26	1172.05	/
Re-profiling Amount/mm	7.19	10.19	16.62	9.18	11.38	43.18
Natural Wear/mm	0	2.46	3.65	16.63	1.83	22.74
Total Wear/mm	7.19	12.65	20.27	25.81	13.21	65.92
Re-profiling Amount %	1	0.806	0.82	0.356	0.861	0.655
Natural Wear %	0	0.194	0.18	0.644	0.139	0.345
WearRate_re-profiling	/	0.263	0.293	0.329	0.635	0.306
WearRate_Natural	/	0.063	0.064	0.595	0.102	0.161
WearRate_Total	/	0.326	0.358	0.924	0.737	0.467

Table B.9 Statistics for wear rate: locomotive 1. II3H

Locomotive	1	Position	II3H			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres /1000km	838.124	876.902	933.531	961.475	979.405	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1250.99	1239.23	1221.26	1204.72	1182.2	/
Diameters (after)/mm	1242.87	1227.73	1207.29	1185.02	1171.97	/
Re-profiling Amount/mm	8.12	11.5	13.97	19.7	10.23	53.29
Natural Wear/mm	0	3.64	6.47	2.57	2.82	12.68
Total Wear/mm	8.12	15.14	20.44	22.27	13.05	65.97
Re-profiling Amount %	1	0.76	0.683	0.885	0.784	0.808
Natural Wear %	0	0.24	0.317	0.115	0.216	0.192
WearRate_re-profiling	/	0.297	0.247	0.705	0.571	0.377
WearRate_Natural	/	0.094	0.114	0.092	0.157	0.09
WearRate_Total	/	0.39	0.361	0.797	0.728	0.467

Table B.10 Statistics for wear rate: locomotive 1. II3V

Locomotive	1	Position	II3V			Total/Average
Number of re-profiling	1	2	3	4	5	5 times
Re-profiling date	201010	201103	201108	201110	201112	14 months
Reported kilometres /1000km	838.124	876.902	933.531	961.475	979.405	/
Absolute kilometres /1000km	0	38.778	56.629	27.944	17.93	141.281
Diameters (before)/mm	1251.65	1240.1	1222.88	1205.08	1182.09	/
Diameters (after)/mm	1243.03	1228.55	1206.93	1184.9	1171.98	/
Re-profiling Amount/mm	8.62	11.55	15.95	20.18	10.11	56.3
Natural Wear/mm	0	2.93	5.67	1.85	2.81	10.45
Total Wear/mm	8.62	14.48	21.62	22.03	12.92	66.75
Re-profiling Amount %	1	0.798	0.738	0.916	0.783	0.843
Natural Wear %	0	0.202	0.262	0.084	0.217	0.157
WearRate_re-profiling	/	0.298	0.282	0.722	0.564	0.398
WearRate_Natural	/	0.076	0.1	0.066	0.157	0.074
WearRate_Total	/	0.373	0.382	0.788	0.721	0.472

Table B.11 Statistics for wear rate: locomotive 2. I1V

Locomotive	2	Position	I1V		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1249.91	1234.97	1216.01	1199.92	/
Diameters (after)/mm	1238.98	1225.44	1204.84	1174.88	/
Re-profiling Amount/mm	10.93	9.53	11.17	25.04	56.67
Natural Wear/mm	0	4.01	9.43	4.92	18.36
Total Wear/mm	10.93	13.54	20.6	29.96	75.03
Re-profiling Amount %	1	0.704	0.542	0.836	0.755
Natural Wear %	0	0.296	0.458	0.164	0.245
WearRate_re-profiling	/	0.175	0.152	0.582	0.331
WearRate_Natural	/	0.074	0.128	0.114	0.107
WearRate_Total	/	0.249	0.28	0.697	0.439

Table B.12 Statistics for wear rate: locomotive 2. I2H

Locomotive	2	Position	I2H		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1251.73	1235.75	1218.36	1202.83	/
Diameters (after)/mm	1239.55	1224.89	1205.41	1175.06	/
Re-profiling Amount/mm	12.18	10.86	12.95	27.77	63.76
Natural Wear/mm	0	3.8	6.53	2.58	12.91
Total Wear/mm	12.18	14.66	19.48	30.35	76.67
Re-profiling Amount %	1	0.741	0.665	0.915	0.832
Natural Wear %	0	0.259	0.335	0.085	0.168
WearRate_re-profiling	/	0.2	0.176	0.646	0.373
WearRate_Natural	/	0.07	0.089	0.06	0.076
WearRate_Total	/	0.27	0.265	0.706	0.448

Table B.13 Statistics for wear rate: locomotive 2. I2V

Locomotive	2	Position	I2V		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1251.55	1234.84	1218.37	1203.42	/
Diameters (after)/mm	1239.46	1224.7	1205.44	1174.98	/
Re-profiling Amount/mm	12.09	10.14	12.93	28.44	63.6
Natural Wear/mm	0	4.62	6.33	2.02	12.97
Total Wear/mm	12.09	14.76	19.26	30.46	76.57
Re-profiling Amount %	1	0.687	0.671	0.934	0.831
Natural Wear %	0	0.313	0.329	0.066	0.169
WearRate_re-profiling	/	0.187	0.176	0.661	0.372
WearRate_Natural	/	0.085	0.086	0.047	0.076
WearRate_Total	/	0.272	0.262	0.708	0.448

Table B.14 Statistics for wear rate: locomotive 2. I3H

Locomotive	2	Position	I3H		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1252.12	1233.64	1215.35	1201.39	/
Diameters (after)/mm	1239.89	1223.32	1205.37	1175.3	/
Re-profiling Amount/mm	12.23	10.32	9.98	26.09	58.62
Natural Wear/mm	0	6.25	7.97	3.98	18.2
Total Wear/mm	12.23	16.57	17.95	30.07	76.82
Re-profiling Amount %	1	0.623	0.556	0.868	0.763
Natural Wear %	0	0.377	0.444	0.132	0.237
WearRate_re-profiling	/	0.19	0.136	0.607	0.343
WearRate_Natural	/	0.115	0.108	0.093	0.106
WearRate_Total	/	0.305	0.244	0.699	0.449

Table B.15 Statistics for wear rate: locomotive 2. I3V

Locomotive	2	Position	I3V		Total/Average
Number	1	2	3	4	4 times
Date	201010	201102	201109	201201	15 months
Reported kilometers /1000km	33.366	87.721	161.346	204.349	/
Absolut kilometers /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1251.91	1234.58	1216.09	1202.29	/
Diameters (after)/mm	1239.96	1223.35	1208.38	1175.32	/
re-profiling Amount/mm	11.95	11.23	7.71	26.97	57.86
Natural Wear/mm	0	5.38	7.26	6.09	18.73
Total Wear/mm	11.95	16.61	14.97	33.06	76.59
re-profiling Amount %	1	0.676	0.515	0.816	0.755
Natural Wear %	0	0.324	0.485	0.184	0.245
WearSpeed_re-profiling	/	0.207	0.105	0.627	0.338
WearSpeed_Natural	/	0.099	0.099	0.142	0.11
WearSpeed_Total	/	0.306	0.203	0.769	0.448

Table B.16 Statistics for wear rate: locomotive 2. II1V

Locomotive	2	Position	II1V		Total/Average
Number of re-profiling	1	2	3	4	1
Re-profiling date	201010	201102	201109	201201	201010
Reported kilometres /1000km	33.366	87.721	161.346	204.349	33.366
Absolute kilometres /1000km	0	54.355	73.625	43.003	0
Diameters (before)/mm	1251.67	1237.44	1213.65	1201.64	1251.67
Diameters (after)/mm	1241.88	1222.01	1204.11	1169.24	1241.88
Re-profiling Amount/mm	9.79	15.43	9.54	32.4	9.79
Natural Wear/mm	0	4.44	8.36	2.47	0
Total Wear/mm	9.79	19.87	17.9	34.87	9.79
Re-profiling Amount %	1	0.777	0.533	0.929	1
Natural Wear %	0	0.223	0.467	0.071	0
WearRate_re-profiling	/	0.284	0.13	0.753	/
WearRate_Natural	/	0.082	0.114	0.057	/
WearRate_Total	/	0.366	0.243	0.811	/

Table B.17 Statistics for wear rate: locomotive 2. II2H

Locomotive	2	Position	II2H		Total/Average
Number of re-profiling	0	1	2	3	3 times
Re-profiling date	/	201102	201109	201201	11 months
Reported kilometres /1000km	/	87.721	161.346	204.349	/
Absolute kilometres /1000km	/	0	73.625	43.003	116.628
Diameters (before)/mm	/	1235.7	1215.56	1201.41	/
Diameters (after)/mm	/	1225.41	1204.04	1169.48	/
Re-profiling Amount/mm	/	10.29	11.52	31.93	53.74
Natural Wear/mm	/	0	9.85	2.63	12.48
Total Wear/mm	/	10.29	21.37	34.56	66.22
Re-profiling Amount %	/	1	0.539	0.924	0.812
Natural Wear %	/	0	0.461	0.076	0.188
WearRate_re-profiling	/	/	0.156	0.743	0.461
WearRate_Natural	/	/	0.134	0.061	0.107
WearRate_Total	/	/	0.29	0.804	0.568

Table B.18 Statistics for wear rate: locomotive 2. II2V

Locomotive	2	Position	II2V		Total/Average
Number of re-profiling	0	1	2	3	3 times
Re-profiling date	/	201102	201109	201201	11 months
Reported kilometres /1000km	/	87.721	161.346	204.349	/
Absolute kilometres /1000km	/	0	73.625	43.003	116.628
Diameters (before)/mm	/	1236.39	1215.64	1201.57	/
Diameters (after)/mm	/	1225.18	1204.09	1169.32	/
Re-profiling Amount/mm	/	11.21	11.55	32.25	55.01
Natural Wear/mm	/	0	9.54	2.52	12.06
Total Wear/mm	/	11.21	21.09	34.77	67.07
Re-profiling Amount %	/	1	0.548	0.928	0.82
Natural Wear %	/	0	0.452	0.072	0.18
WearRate_re-profiling	/	/	0.157	0.75	0.472
WearRate_Natural	/	/	0.13	0.059	0.103
WearRate_Total	/	/	0.286	0.809	0.575

Table B.19 Statistics for wear rate: locomotive 2. II3H

Locomotive	2	Position	II3H		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1252.38	1235.08	1214.07	1198.96	/
Diameters (after)/mm	1240.05	1224.93	1204.02	1169.13	/
Re-profiling Amount/mm	12.33	10.15	10.05	29.83	62.36
Natural Wear/mm	0	4.97	10.86	5.06	20.89
Total Wear/mm	12.33	15.12	20.91	34.89	83.25
Re-profiling Amount %	1	0.671	0.481	0.855	0.749
Natural Wear %	0	0.329	0.519	0.145	0.251
WearRate_re-profiling	/	0.187	0.137	0.694	0.365
WearRate_Natural	/	0.091	0.148	0.118	0.122
WearRate_Total	/	0.278	0.284	0.811	0.487

Table B.20 Statistics for wear rate: locomotive 2. II3V

Locomotive	2	Position	II3V		Total/Average
Number of re-profiling	1	2	3	4	4 times
Re-profiling date	201010	201102	201109	201201	15 months
Reported kilometres /1000km	33.366	87.721	161.346	204.349	/
Absolute kilometres /1000km	0	54.355	73.625	43.003	170.983
Diameters (before)/mm	1250.55	1236.41	1213.63	1198.7	/
Diameters (after)/mm	1248.02	1225.1	1204.04	1169.09	/
Re-profiling Amount/mm	2.53	11.31	9.59	29.61	53.04
Natural Wear/mm	0	11.61	11.47	5.34	28.42
Total Wear/mm	2.53	22.92	21.06	34.95	81.46
Re-profiling Amount %	1	0.493	0.455	0.847	0.651
Natural Wear %	0	0.507	0.545	0.153	0.349
WearRate_re-profiling	/	0.208	0.13	0.689	0.31
WearRate_Natural	/	0.214	0.156	0.124	0.166
WearRate_Total	/	0.422	0.286	0.813	0.476

Paper V

Bayesian Parametric Analysis for Reliability Study of Locomotive Wheels

- Lin Jing, Asplund Matthias, Parida Aditya.
- Published in Conference Proceedings.
- The 59th Annual Reliability and Maintainability Symposium (RAMS® 2013).
- January 28-31, Orlando, FL, USA.
- Oral Presentation.

Bayesian Parametric Analysis for Reliability Study of Locomotive Wheels

Jing Lin, PhD, Luleå University of Technology

Matthias Asplund, Luleå University of Technology

Aditya Parida, PhD, Luleå University of Technology

Key Words: Reliability analysis; Bayesian analysis; Locomotive; train wheels; Markov Chain Monte Carlo

SUMMARY & CONCLUSIONS

This paper proposes a new approach to study reliability of locomotive wheels with Bayesian framework, utilizing locomotive wheel degradation data sets that can be small or incomplete. In our study, a linear degradation path is assumed and locomotive wheels' installation positions are considered as covariates. A Markov Chain Monte Carlo (MCMC) computational method is also implemented. In the case study, data were collected from a Swedish railway company. This data includes, the diameter measurements of the locomotive wheels, total distances corresponding to their "time to maintenance", and the wheels' bill of material (BOM) data. During this study, likelihood functions were constructed for Expontional regression models, Weibull regression models, and lognormal regression models. The results show that the locomotive wheels' lifetimes are dependent on installation positions. For the studied locomotive wheels data, the Lognormal regression model is a better choice, because the model obtained the lowest Deviance Information Criterion (DIC) values. In addition, under current operation situation (e.g. topography) and current maintenance strategies (re-profiled, lubrication, etc.), the locomotive wheels installed in the second bogie have longer lifetimes than those installed in the first bogie; the wheels installed on the "back" axle have longer lifetimes than those on the "front" axle; and the right side wheels' lifetime is shorter than that for the left side under a given running situation.

1 INTRODUCTION

The service life of a railroad wheel can be significantly reduced due to failure or damage, leading to excessive cost and accelerated deterioration. Damage data show that a major proportion of wheel damage stems from degradation.

In order to monitor the performance of wheels and make replacement before adverse effects occur, the railway industry uses both preventive and predictive maintenance. [1-5] By predicting train wheel wear, fatigue, tribological aspects, and failures, the railway industry can formulate different preventive maintenance strategies under different time periods. [6] For predictive maintenance, wheel condition monitoring data have been studied to increase the lifetime by knowing the condition of the wheel profile. [7-9] A large number of related studies have been published in the last decade.

One common preventive maintenance strategy (that is used in the case study) is wheel re-profiling after running a certain distance, with its diameter being measured; if it reduces to a pre-specified diameter, the wheel will be replaced by a new one. In order to optimize such maintenance strategies, some researchers started looking into wheel degradation utilization data to determine reliability and failure distribution. [10] However, these studies cannot solve the combined problem of small data samples and incomplete data sets while simultaneously considering the influence of several covariates. For example, to avoid the potential influence of the different locations of wheels, Freitas [10] only consider those on the left side of axle number and on certain specified cars. Yang and Letourneau [5] suggest that certain attributes, including a wheel's installed position (right or left), might influence its wear rate, but they do not provide case studies.

To address the above issues, this paper undertakes a reliability study using a Bayesian survival analysis framework to explore the impact of the wheel's installed position on its service lifetime and to predict its reliability characteristics. [11] In section 2, the Expontional regression models, Weibull regression models, and lognormal regression models are used to establish the lifetime of locomotive wheels using degradation data and taking into account the position of the wheel. This position is described by three different discrete covariates: the bogie, the axel and the side of the locomotive where the wheel is mounted. In section 3, a linear degradation path is considered and the case study is performed using Markov Chain Monte Carlo methods. And finally, section 4 offers conclusions and comments.

2 BAYESIAN PARAMETRIC MODELS

In reliability analysis, the lifetime data set is usually incomplete, which means only a portion of the subsystem failures (i.e. wheels) are known. For other subsystems, the lifetimes are only known to exceed certain values. Take the locomotive wheels' degradation data for example. If the degradation data is less than the pre-specified diameter, the

978-1-4673-4711-2/13/$31.00 ©2013 IEEE

corresponding predicted lifetime is viewed as right-censored. The reason is that under a linear degradation path assumption, we just know those wheels' real lifetime will exceed the predicted lifetime.

Right-censored data are often called Type I censoring in the literature; the corresponding likelihood construction problem has been extensively studied. [12, 13] Suppose there are n individuals whose lifetimes and censoring times are independent. The ith individual has life time T_i and censoring time L_i. The T_i s are assumed to have probability density function $f(t)$ and reliability function $R(t)$. The exact lifetime T_i of an individual will be observed only if $T_i \le L_i$. The lifetime data involving right censoring can be conveniently represented by n pairs of random variables (t_i, υ_i), where $t_i = \min(T_i, L_i)$ and $\upsilon_i = 1$ if $T_i \le L_i$, and $\upsilon_i = 0$ if $T_i > L_i$. That is, υ_i indicates whether the lifetime T_i is censored or not. The likelihood function is deduced as [12, 13]

$$L(t) = \prod_{i=1}^{n} [f(t_i)]^{\upsilon_i} R(t_i)^{1-\upsilon_i} \tag{1}$$

2.1 Exponential regression model

Suppose the lifetimes $\mathbf{t} = (t_1, \cdots t_n)$ for n wheels are independent and identically distributed (iid), and the distribution being exponential distribution with a failure rate λ, where $\lambda > 0$. Therefore, the probability density function (pdf) is $f(t_i|\lambda) = \lambda \exp(-\lambda t_i)$, the cumulative distribution function (cdf) is $F(t_i|\lambda) = 1 - \exp(-\lambda t_i)$ and the reliability function is $R(t_i|\lambda) = 1 - R(t_i|\lambda)$. The incomplete indicators are denoted $\boldsymbol{\upsilon} = (\upsilon_1, \upsilon_2, \cdots, \upsilon_n)$ and the observed data set for the current study is $D_0 = (n, \mathbf{t}, \boldsymbol{\upsilon})$. From equation (1), the likelihood function of λ is

$$L(\lambda|D_0) = \prod_{i=1}^{n} \left(\lambda \exp(-\lambda t_i)\right)^{\upsilon_i} \left(\exp(-\lambda t_i)\right)^{1-\upsilon_i} \tag{2}$$

The $p \times 1$ vector of covariates for the ith wheel is denoted $\mathbf{x_i} = (x_{1i}, \cdots x_{pi})'$. Similarly, $\boldsymbol{\beta}$ is a $p \times 1$ vector of regression coefficients, which represents the degree of influences of covariates. Let $\lambda_i = \exp(\mathbf{x}_i'\boldsymbol{\beta})$ and the observed data set for current study is denoted by $D = (n, \mathbf{t}, \mathbf{X}, \boldsymbol{\upsilon})$. The likelihood function for the regression coefficients is given by

$$L(\boldsymbol{\beta}|D) = \exp\left(\sum_{i=1}^{n} \upsilon_i \mathbf{x_i}'\boldsymbol{\beta}\right) \exp\left(-\sum_{i=1}^{n} \exp(\mathbf{x_i}'\boldsymbol{\beta}) t_i\right) \tag{3}$$

The prior distributions should be realistic and computationally feasible. There are two common choices for $\boldsymbol{\beta}$'s prior distributions. [11] One is uniform improper prior distribution, for example, $\pi(\boldsymbol{\beta}) \propto 1$ .The other is the normal distribution. As proved by Ibrahim [11], it's a log-concave prior and such kind of choice will be convenient for posterior's computation. To implement the MCMC simulation more easily, a multinomial prior $\boldsymbol{\beta} \sim N_p(\boldsymbol{\mu_0}, \boldsymbol{\Sigma_0})$ with mean $\boldsymbol{\mu_0}$ and covariance matrix $\boldsymbol{\Sigma_0}$ is assumed. Let $\pi(\cdot)$ denote the prior or posterior distributions for the parameters. The posterior distribution, $\pi(\boldsymbol{\beta}|D)$, can be written as

$$\begin{aligned}\pi(\boldsymbol{\beta}|D) &\propto L(\boldsymbol{\beta}|D) \times \pi(\boldsymbol{\beta}|\boldsymbol{\mu_0}, \boldsymbol{\Sigma_0}) \\ &\propto \exp\left(\sum_{i=1}^{n} \upsilon_i \mathbf{x_i}'\boldsymbol{\beta} - \sum_{i=1}^{n} \exp(\mathbf{x_i}'\boldsymbol{\beta}) t_i - \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu_0})' \boldsymbol{\Sigma_0^{-1}} (\boldsymbol{\beta} - \boldsymbol{\mu_0})\right)\end{aligned} \tag{4}$$

It is not easy to get the exact integration results for $\pi(\boldsymbol{\beta}|D)$ due to its complexity. Therefore, we select the MCMC method, which has been widely applied to Bayesian statistics since 1990s, to carry out the posterior inference.

2.2 Weibull regression model

Suppose, the lifetimes $\mathbf{t} = (t_1, \cdots t_n)$ for n individuals are iid, and the distribution is Weibull, $W(\alpha, \gamma)$, where $\alpha > 0$ and $\gamma > 0$. The pdf is $f(t_i|\alpha, \gamma) = \alpha \gamma t_i^{\alpha-1} \exp(-\gamma t_i^{\alpha})$ while the cdf is $F(t_i|\alpha, \gamma) = 1 - \exp(-\gamma t_i^{\alpha})$ and the reliability function $R(t_i|\alpha, \gamma) = 1 - F(t_i|\alpha, \gamma)$. To facilitate the analysis, let $\xi = \ln(\gamma)$, permitting the following representation:

$$f(t_i|\alpha, \xi) = \alpha t_i^{\alpha-1} \exp(\xi - \exp(\xi) t_i^{\alpha}) \tag{5}$$

Similarly, we can get $F(t_i|\alpha, \xi)$ and $R(t_i|\alpha, \xi)$.
From equation (1), the joint likelihood function for α and ξ is

$$\begin{aligned}L(\alpha, \xi|D_0) &= \alpha^{\sum_{i=1}^{n} \upsilon_i} \\ &\times \exp\left(\sum_{i=1}^{n} \upsilon_i \xi + \sum_{i=1}^{n} \upsilon_i (\alpha - 1)\ln(t_i) - \exp(\xi) t_i^{\alpha}\right)\end{aligned} \tag{6}$$

To construct the Weibull Regression Model, covariates are introduced through ξ. With $\xi_i = \mathbf{x_i}'\boldsymbol{\beta}$, the likelihood function is given by

$$\begin{aligned}L(\alpha, \boldsymbol{\beta}|D) &= \alpha^{\sum_{i=1}^{n} \upsilon_i} \\ &\times \exp\left(\sum_{i=1}^{n} \upsilon_i \left(\mathbf{x_i}'\boldsymbol{\beta} + \upsilon_i (\alpha - 1)\ln(t_i) - \exp(\mathbf{x_i}'\boldsymbol{\beta}) t_i^{\alpha}\right)\right)\end{aligned} \tag{7}$$

In this paper, it is assumed that α and ξ are independent. Furthermore, it is assumed that the prior distribution of α is a gamma distribution, denoted by $G(a_0, b_0)$ The prior distribution can be written as $\pi(\alpha|a_0, b_0) \propto \alpha^{a_0 - 1} \exp(-b_0 \alpha)$. Then, the posterior distribution of α and $\boldsymbol{\beta}$ is:

$$\begin{aligned}\pi(\alpha, \boldsymbol{\beta}|D) &\propto \alpha^{a_0 - 1 + \sum_{i=1}^{n} \upsilon_i} \\ &\times \exp\left(\sum_{i=1}^{n} \left(\upsilon_i \mathbf{x_i}'\boldsymbol{\beta} + \upsilon_i (\alpha - 1)\ln(t_i) - \exp(\mathbf{x_i}'\boldsymbol{\beta}) t_i^{\alpha}\right) \right. \\ &\left. - b_0 \alpha - \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu_0})' \boldsymbol{\Sigma_0^{-1}} (\boldsymbol{\beta} - \boldsymbol{\mu_0})\right)\end{aligned} \tag{8}$$

2.3 Lognormal regression model

Suppose the lifetimes $\mathbf{t} = (t_1, \cdots t_n)$ of n wheels are iid, with $\ln(t)$ being normally distributed according to $N(\mu, \sigma^2)$. This implies that t_i is lognormally distributed with parameters μ and σ^2, denoted by $LN(\mu, \sigma^2)$. The pdf and reliability functions for t_i are

$$f(t_i|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma t_i} \exp\left(-\frac{1}{2\sigma^2}(\ln(t_i) - \mu)^2\right) \tag{9}$$

$$R(t_i|\mu, \sigma^2) = 1 - \Phi\left(\frac{\ln(t_i) - \mu}{\sigma}\right) \tag{10}$$

From equation (1), the joint likelihood function for μ and σ given an incomplete data set is

$$L(\mu,\sigma|D_0) = (2\pi\sigma^2)^{-\frac{1}{2}\sum_{i=1}^{n}\upsilon_i} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\upsilon_i(\ln(t_i)-\mu)^2\right) \times\prod_{i=1}^{n} t_i^{-\upsilon_i}\left(1-\Phi\left(\frac{\log(t_i)-\mu}{\sigma}\right)\right)^{1-\upsilon_i} \quad (11)$$

To construct a lognormal regression model covariates that are realized through μ are introduced by defining $\mu_i = \mathbf{x}_i^{'}\boldsymbol{\beta}$. By defining $\tau = 1/\sigma^2$, the likelihood function is given can be written as

$$L(\boldsymbol{\beta},\tau|D) = (2\pi\tau^{-1})^{-\frac{1}{2}\sum_{i=1}^{n}\upsilon_i} \exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}\upsilon_i(\ln(t_i)-\mathbf{x}_i^{'}\boldsymbol{\beta})^2\right) \times\prod_{i=1}^{n} t_i^{-\upsilon_i}\left(1-\Phi\left(\frac{\ln(t_i)-\mathbf{x}_i^{'}\boldsymbol{\beta}}{\tau^{-1/2}}\right)\right)^{1-\upsilon_i} \quad (12)$$

A typical prior distribution for τ is a gamma prior distribution. [11] In this paper, it is supposed that $\tau \sim G(a_0/2, b_0/2)$, $\boldsymbol{\beta}$ has a multinormal prior distribution with p vector, denoted by $N_p(\boldsymbol{\mu}_0, \tau^{-1}\boldsymbol{\Sigma}_0)$.[11] Therefore, the posterior distribution for τ and $\boldsymbol{\beta}$ can be is

$$\pi(\boldsymbol{\beta},\tau|D) \propto \tau^{\frac{a_0+\sum_{i=1}^{n}\upsilon_i}{2}-1} \times\exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}\upsilon_i(\ln(t_i)-\mathbf{x}_i^{'}\boldsymbol{\beta})^2+(\boldsymbol{\beta}-\boldsymbol{\mu}_0)^{'}\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\beta}-\boldsymbol{\mu}_0)+b_0\right) + \prod_{i=1}^{n} t_i^{-\upsilon_i}\left(1-\Phi\left(\tau^{1/2}(\ln(t_i)-\mathbf{x}_i^{'}\boldsymbol{\beta})\right)\right)^{1-\upsilon_i} \quad (13)$$

3 EXAMPLE

This paper focuses on the wheels of the locomotive of a cargo train. While two types of locomotives with the same type of wheels are used in cargo trains, we consider only one.

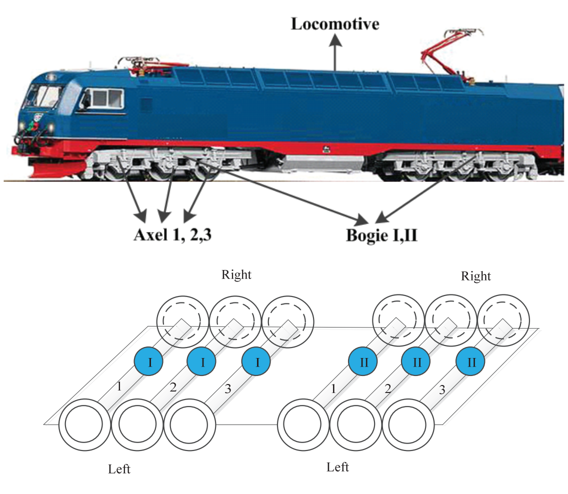


Fig.1 Locomotive wheels' installation positions

There are two bogies for each locomotive and three axels for each bogie (Fig.1). The installed position of the wheels on a particular locomotive is specified by a bogie number (I, II-number of bogies on the locomotive), an axel number (1, 2, 3-number of axels for each bogie) and the side of the wheel on the axle (right or left) where each wheel is mounted.

The diameter of a new locomotive wheel in the studied railway company is 1250 mm. In the company's current maintenance strategy, a wheel's diameter is measured after running a certain distance. If it is reduced to 1150 mm, the wheel is replaced by a new one. Otherwise, it is re-profiled or other maintenance strategies are implemented. A threshold level for failure, is defined as 100 mm (= 1250 mm -1150 mm). The wheel's failure condition is assumed to be reached if the diameter reaches 100mm.

The company's complete database also includes the diameters of all locomotive wheels at a given observation time, the total running distances corresponding to their "time to be maintained", and the wheels' bill of material (BOM) data, from which we can determine their positions.

Two assumptions are made: 1) for each censored datum it is supposed that the wheel is replaced; 2) degradation is linear. Only one locomotive is considered in this example to ensure that 1) all wheel's maintenance strategies are the same; 2) the axle load and running speed are obviously constant; and 3) the operational environments including routes, climates and exposure are common for all wheels.

The data set contains 46 datum points (n =46) of a single locomotive throughout period November 2010 to January 2012. We take the following steps to obtain locomotive wheels' lifetime data (Fig.2):

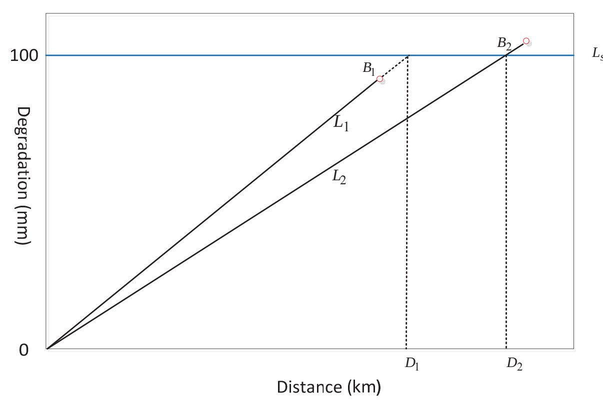


Fig.2 Plot of the wheel degradation data: one example

- Establish a threshold level L_s, where L_s =100 mm (1250 mm – 1150 mm).
- Transfer observed 90 records of wheel diameters at reported time t to degradation data; this equals to 1250mm minus the corresponding observed diameter.
- Assume a liner degradation path and construct a degradation line L_i (e.g. L_1, L_2) using the origin point and the degradation data (e.g. B_1, B_2).
- Set $L_s = L_i$, get the point of intersection and the corresponding lifetimes data (e.g. D_1, D_2).

For each reported datum, a wheel's installation position is documented, and as mentioned above, positioning is used in this study as a covariate. As discussed in section 3, the wheel's position (bogie, axel, and side) or covariate $\mathbf{X}$ is denoted by x_1 (bogie I: x_1=1, bogie II: x_1=2), x_2 (axel 1: x_2=1, axel 2: x_2=2, axel 3, x_2=3) and x_3 (right: x_3=1, left: x_3=2). Correspondingly, the covariates' coefficients are represented by β_1, β_2, and β_3. In addition, β_0 is defined as random effect.

The calculations are implemented with the software WinBUGS[14] . A burn-in of 10,001 samples is used, with an additional 10,000 Gibbs samples for each Markov chain. Vague prior distributions are adopted here as the following: For exponential regression: $\boldsymbol{\beta} \sim N(0,0.0001)$; for Weibull regression: $\alpha \sim G(0.2,0.2)$, $\boldsymbol{\beta} \sim N(0,0.0001)$; for lognormal Regression: $\tau \sim G(1,0.01)$, $\boldsymbol{\beta} \sim N(0,0.0001)$.

Following the convergence diagnostics (including to check Markov chains' dynamic trace, time series, Gelman-Rubin-Statistics, as well as to compare the MC error with Standard Deviation (SD)), [14] we consider following posterior summaries of parameters as shown in tables 1, 2 and 3 for our models with censored data, including the parameters' posterior mean, standard deviation, Monte Carlo (MC) error, and 95% HPD (highest posterior distribution density) interval. In below tables, the mean values for $\boldsymbol{\beta}$ seem quite small, that is because the measurement unit to locomotive wheels lifetime is by thousands kilometers ($\times 10^3$ km).

Table.1 Posteriors Summaries - Exponential Regression Model

Parameter	Mean	SD	MC error	95 % HPD
β_0	-5.862	0.7355	0.02299	(-7.366,-4.452)
β_1	-0.07207	0.3005	0.007269	(-0.6672,0.5104)
β_2	-0.03219	0.1858	0.003797	(-0.3889,0.3325)
β_3	-0.0124	0.2973	0.00726	(-0.5954,0.5787)

Table.2 Posteriors Summaries - Weibull Regression Model

Parameter	Mean	SD	MC error	95 % HPD
α	10.08	0.9674	0.05559	(8.234,11.76)
β_0	-60.47	5.977	0.3434	(-71.01,-49.16)
β_1	-0.07775	0.306	0.008339	(-0.6845,0.5156)
β_2	-0.146	0.2231	0.005801	(-0.5878,0.2856)
β_3	-0.05026	0.2982	0.007143	(-0.6356,0.5324)

Table.3 Posteriors Summaries - Log-normal Regression Model

Parameter	Mean	SD	MC error	95 % HPD
β_0	5.864	0.05341	0.001622	(5.76,5.97)
β_1	0.06733	0.02174	5.042E-4	(0.02492,0.1103)
β_2	0.02077	0.01373	2.765E-4	(-0.00629,0.0478)
β_3	0.001102	0.02175	5.007E-4	(-0.0412,0.04444)
τ	187.5	39.84	0.3067	(118.3,273.5)

Accordingly, the locomotive wheels' reliability functions are:

- Exponential Regression Model:

$$R(t_i \mid \mathbf{X}) = \exp\left[-\exp(-5.862 - 0.072x_1 - 0.032x_2 - 0.012x_3) \times t_i\right]$$

- Weibull Regression Model:

$$R(t_i \mid \mathbf{X}) = \exp\left[-\exp(-60.47 - 0.078x_1 - 0.146x_2 - 0.050x_3) \times t_i^{10.08}\right]$$

- Log-normal Regression Model:

$$R(t_i \mid \mathbf{X}) = 1 - \Phi\left[\frac{\ln(t_i) - (5.864 + 0.067x_1 + 0.02x_2 + 0.001x_3)}{(187.5)^{-1/2}}\right]$$

Obviously, other quantities regarding lifetime distribution, including MTTF can be determined.

For model comparison, usually two main aspects are considered: the model's measure of fit and its complexity. In this paper, we adopt the Deviance Information Criterion (DIC), which utilizes the model's deviance to evaluate its measure of fit, and the effective number of parameters to evaluate its complexity. [14]

Define a Bayesian model's Bayesian deviance, denoted as $D(\theta)$, as $D(\theta) = -2\log(p(D|\theta))$; Define the effective number of parameters, denoted as p_d, as:

$$p_d = \overline{D}(\theta) - D(\overline{\theta}) = -\int 2\ln(p(D|\theta))d\theta - (-2\ln(p(D|\overline{\theta})))$$

Then, $DIC = D(\overline{\theta}) + 2p_d = \overline{D}(\theta) + p_d$. We calculate the DIC values for the above three Bayesian parametric models separately, as shown in Table 4.

Table.4 DIC Summaries

Model	$\overline{D}(\theta)$	$D(\overline{\theta})$	p_d	DIC
Exponential	648.98	645.03	3.95	652.93
Weibull	472.22	467.39	4.83	477.05
Log-normal	442.03	436.87	5.16	447.19

Our results show that the DIC for Log-normal Regression Model is the lowest (447.19), and it is a better choice. The prediction of the locomotive wheels MTTF, following Bayesian Lognormal regression model, appears in Table.5.

It should be pointed out that the 95% HPD interval in Bayesian Lognormal regression model for β_2 and β_3 includes 0 (Table.3). This means that, although the positioning does have an influence, in some instances, the impact on the wheel's service lifetime is not significantly strong. In our case, the bogies have more impact on service lifetime than axels or sides. Given this conclusion, we can deal with such covariates better in our future research. Besides above, other conclusions include: 1) the lifetime of the wheel installed in the second bogie is longer than that of the wheel installed in the first one; 2) the wheel installed in the third axel has a longer lifetime than that installed in the second axel, and the wheel in the second axel has a longer lifetime than the one in the first axel; 3) the right side wheel's lifetime is shorter than the left side.

(Researchers from Norwegian National Rail Administration cited previously concur with this. Using condition monitoring methods on train wheels operating on the same route, they found that the wheel forces on the right and the left sides can be different, even for wheels in the same axel.). Possible causes include the influence of the earth's rotation, topographical complexity, and the position of the locomotive's centre of gravity.

Table.5 MTTF statistics based on Bayesian Lognormal Regression Model

Bogie	Axel	Side	μ_i	MTTF ($\times 10^3$ km)
I (x_1=1)	1 (x_2=1)	Right(x_3=1)	5.9532	387.03
		Left (x_3=2)	5.9543	387.46
	2 (x_2=2)	Right(x_3=1)	5.9740	395.16
		Left (x_3=2)	5.9751	395.60
	3 (x_2=3)	Right(x_3=1)	5.9947	403.43
		Left (x_3=2)	5.9958	403.87
II (x_1=2)	1 (x_2=1)	Right(x_3=1)	6.0205	413.97
		Left (x_3=2)	6.0216	414.43
	2 (x_2=2)	Right(x_3=1)	6.0413	422.67
		Left (x_3=2)	6.0424	423.14
	3 (x_2=3)	Right(x_3=1)	6.0621	431.56
		Left (x_3=2)	6.0632	432.03

4 CONCLUSIONS AND FUTURE RESEARCH

This paper proposes three parametric Bayesian models for locomotive wheels' reliability analysis using degradation data: Bayesian Exponential Regression Model, Bayesian Weibull Regression Model, and Log-normal Regression Model. By introducing the covariate $\mathbf{x}_i$'s linear function $\mathbf{x}_i^{'}\boldsymbol{\beta}$, these three parameter models are constructed depending on the failure rate λ_i in the exponential model, the log of the rate parameter $\ln(\gamma_i)$ in the Weibull model and the logarithmic mean μ_i in the log-normal models. The proposed Bayesian survival models can deal with small and incomplete data sets and simultaneously consider the influence of several covariates. The MCMC technique via the Gibbs sampler is used here to achieve models' posteriors estimations.

The case study's results suggest that the locomotive wheels' lifetimes are different with different installed positions. In addition, the approach discussed in this paper can also be applied for analyzing cargo train wheels.The work presented also leads to the implementation of additional research:

- The assumed liner degradation path was a simple one. For more complex path models, more degradation paths need to be studied, including considering different wear rates.
- The covariates considered here are only limited to locomotive wheels' installed positions, more covariates needs to be considered later, like: temperature, applied loading, train speed, *etc.*
- We have chosen general prior distributions for the case study. As more information can be utilized, how to integrate different prior also need to be studied. Besides above limitations, in our later research, we also plan to consider utilization of utilization of our results to optimize maintenance strategies and related LCC (Life Cycle Cost) problems with consideration of maintenance cost.

REFERENCES

1. Braghin F, et al. A Mathematical Model to Predict Railway Wheel Profile Evolution Due to Wear. Journal of Wear. 2006. 261: 1253-1264
2. Tassini N, et al. A Numerical Model of Twin Disc Test Arrangement for the Evaluation of Railway Wheel Wear Prediction Methods. Journal of Wear. 2010. 268: 660-667
3. Bernasconi A, et al. An Integrated Approach to Rolling Contact Sub-surface Fatigue assessment of Railway Wheels. Journal of Wear. 2005. 258: 973-980
4. Clayton P. Tribological Aspects of Wheel-Rail Contact: A Review of Recent Experimental Research. Journal of Wear. 1996. 191: 170-183
5. Yang C, Letourneau S. Learning to Predict Train Wheel Failures. Conference Proceedings. The 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005). Chicago, Illinois, USA.
6. Pombo J, Ambrosio J, Pereira M. A Railway Wheel Wear Prediction Tool based on A Multibody Software. Journal of Theoretical and Applied Mechanics. 2010. 48, 3:751-770
7. Donato P, et al. Design and Signal Processing of A Magnetic Sensor Array for Train Wheel Detection. Journal of Senors and Actuators A. 2006. 132: 516-525
8. Stratman B, Liu Y, Mahadevan S. Structural Health Monitoring of Railroad Wheels Using Wheel Impact Load Detectors. Journal of Failure Analysis and Prevention. 2007. 7(3):218-225
9. Palo M. Condition Monitoring of Railway Vehicles: A Study on Wheel Condition for Heavy Haul Rolling Stock. Licentiate Thesis. Luleå University of Technology, Sweden. 2012
10. Freitas M A, et al. Using Degradation Data to Assess Reliability: A Case Study on Train Wheel Degradation. Journal of Quality and Reliability Engineering International. 2009, 25: 607-629
11. Ibrahim J G, Chen M H, Sinha D. Bayesian Survival Analysis. New York: Berlin Heidelberg, 2001
12. Klein J P, Moeschberger M L. Survival Analysis: Techniques for Censored and Truncated Data. Springer-Verlag New York, Inc.1997
13. Lawless. Statistical Models and Methods for Lifetime Data. John Wiley and Sons.1982
14. Spiegelhalter D, et al. WinBUGS User Manual (Version 1.4). January, 2003. http:// www.mrc-bsu.cam.ac.uk/bugs

ACKNOWLEDGEMENTS

The authors would like to thank Luleå Railway Research Centre (Järnvägstekniskt Centrum, Sweden) for initiating the research study and Swedish Transport Administration (Trafikverket) for providing financial support. Also we would like to thank the editor and anonymous referees for their constructive comments.

BIOGRAPHIES

Jing Lin, PhD
Division of Operation, Maintenance and Acoustics
Luleå University of Technology
Luleå, Norrbotten 97187 Sweden

e-mail: janet.lin@ltu.se

Dr. Jing Lin is currently a researcher in the Division of Operation, Maintenance and Acoustic, at Luleå University of Technology (LTU), Sweden. She obtained her PhD degree in Management from Nanjing University of Science and Technology (NJUST), China, in April 2008; and she received the bachelor degree in Management from (NJUST) in 2003. After the college, she worked 3 years for SKF Co., Ltd as an Asset Management Consultant. Dr. Lin's research interests primarily lie in asset management and reliability. She has published 35 peer reviewed Journal and Conference papers and 1 monograph in related topics.

Matthias Asplund, PhD candidate
Division of Operation, Maintenance and Acoustics
Luleå University of Technology
Luleå, Norrbotten 97187 Sweden

e-mail: Matthias.asplund@ltu.se

Matthias Asplund is a PhD student in the Division of Operation, Maintenance and Acoustic, at Luleå University of Technology (LTU), Sweden. since 2011. His research area is RAMS with railway topics in focus. He has twelve years working experience from product development, lean production, maintenance and railway engineering. He got his Master degree in Mechanical Engineering with focus on Applied Mechanics from Luleå Technical University. His last work before studies to PhD was Track Engineering for the Swedish Infrastructure Manager.

Aditya Parida, PhD
Division of Operation, Maintenance and Acoustics
Luleå University of Technology
Luleå, Norrbotten 97187 Sweden

e-mail: Aditya.parida@ltu.se

Dr. Aditya Parida is an Associate Professor in the Division of Operation, Maintenance and Acoustic, at Luleå University of Technology, Sweden. He obtained his PhD in Operation and Maintenance Engineering. His area of research is Asset Management, Maintenance Performance Measurement and model, RCM and eMaintenance. Besides teaching, he is actively involved in research and projects. He is the author of 65 peer reviewed Journal and Conference papers, besides 3 book chapters and guest editors of four special issues of International Journals.