

**HYBRID BAYESIAN-WIENER PROCESS IN
TRACK GEOMETRY DEGRADATION ANALYSIS**

by

Silvia Adriana Galván Núñez

A dissertation submitted to the Faculty of the University of Delaware in
partial fulfillment of the requirements for the degree of Doctor of Philosophy in Civil
Engineering

Summer 2017

© 2017 Silvia Adriana Galván Núñez
All Rights Reserved

**HYBRID BAYESIAN-WIENER PROCESS IN
TRACK GEOMETRY DEGRADATION ANALYSIS**

by

Silvia Adriana Galván Núñez

Approved: _____
Sue McNeil, Ph.D.
Chair of the Department of Civil and Environmental Engineering

Approved: _____
Babatunde A. Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: _____
Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Nii O. Attoh-Okine, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Sue McNeil, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Allan Zarembski, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Javier Garcia-Frias, Ph.D.
Member of dissertation committee

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my advisor, Prof. Attoh-Okine, for his constant guidance, encouragement, and mentorship during this process.

I would also like to thank each of the members of my dissertation committee for providing extensive professional and academic advice.

Special thanks to my colleagues and research group members for their constructive criticisms and contributions.

I would like to thank family for their unconditional support and love in everything I pursue.

Finally, I would like to dedicate this accomplishment to my parents. They are my ultimate role models.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
ABSTRACT	xvi
 Chapter	
1 INTRODUCTION	1
1.1 Background & Motivation	1
1.2 Statement of the Problem	2
1.3 Objective of the Study	2
1.4 Research Approach	3
1.5 Dissertation Structure	4
1.5.1 Chapter 1: Introduction	4
1.5.2 Chapter 2: Background	4
1.5.3 Chapter 3: Bayesian Inference	4
1.5.4 Chapter 4: Wiener Process for Degradation Analysis	5
1.5.5 Chapter 5: Exploratory Data Analysis	5
1.5.6 Chapter 6: Hybrid Markov Chain Monte Carlo and Wiener Process in Railway Track Geometry Degradation Analysis	5
1.5.7 Chapter 7: First Hitting Time in Railway Track Geometry Degradation Analysis	5
1.5.8 Chapter 8: Conclusions and Recommendations	5
1.6 Contributions of the Dissertation and Summary of Publications, Conferences, and Awards	6
1.6.1 Peer-Review Papers	6
1.6.2 Conferences and Presentations	6
1.6.3 Awards	7
 REFERENCES	 8

2	BACKGROUND	9
2.1	Introduction	9
2.2	Track Characterization	9
2.3	Track Failure Mechanisms	10
2.3.1	Track Structural Defects	13
2.4	Track Geometry	13
2.4.1	Track Geometry Parameters	13
2.5	Track Maintenance Methods	14
2.6	Track Quality Index (TQI)	15
2.7	Track Geometry Degradation Models	17
2.7.1	Regression Models	17
2.7.2	Exponential Smoothing	20
2.7.3	Hierarchical Bayesian Model	21
2.7.4	Neural Networks	21
2.7.5	Stochastic Process	22
2.8	Remarks	25
	REFERENCES	26
3	BAYESIAN INFERENCE	30
3.1	Introduction	30
3.2	Frequentist and Bayesian Approaches	31
3.3	Representation of Bayesian Estimation	32
3.3.1	Conjugate Priors	34
3.3.2	Jeffreys Priors	35
3.3.3	Non-informative Priors	35
3.3.4	Informative Priors	36
3.4	Markov Chain Monte Carlo Methods	36
3.4.1	Markov Chains	36
3.4.2	Metropolis-Hastings Algorithm	38

3.4.3	Gibbs Sampling Algorithm	40
3.4.4	Population-Based Markov Chain Monte Carlo Methods	41
3.4.4.1	Mutation	41
3.4.4.2	Crossover	42
3.5	Illustration of MCMC Output Analysis in Track Geometry Degradation	43
3.6	Remarks	46
	REFERENCES	47
4	WIENER PROCESS FOR DEGRADATION ANALYSIS	49
4.1	Introduction	49
4.2	Degradation in Reliability Analysis	49
4.3	Review on Stochastic Processes	51
4.3.1	Random Variable	51
4.3.1.1	Discrete Random Variables	52
4.3.1.2	Continuous Random Variables	52
4.3.2	Stochastic Process	53
4.4	Stochastic Differential Equations	54
4.4.1	Motivation	54
4.4.2	Itô Integral	56
4.4.3	Stratonovich Integral	58
4.4.4	Strong Law of Large Numbers and Central Limit Theorem	59
4.4.4.1	Strong Law of Large Numbers	59
4.4.4.2	Central Limit Theorem	60
4.5	Wiener Process	60
4.5.1	First Hitting Time	63
4.6	Wiener Process in Track Geometry Degradation	67
4.7	Remarks	69
	REFERENCES	70

5	EXPLORATORY DATA ANALYSIS	74
5.1	Introduction	74
5.2	Data Set Description	74
5.2.1	Types of Data Set of EDA	77
5.3	Graphical Methods for EDA	78
5.3.1	Histogram and Quantile-Quantile Plot	78
5.3.1.1	Histograms and QQ Plots for Longitudinal Data	79
5.3.1.2	Histograms and QQ Plots for Cross-Sectional Data	80
5.3.2	Box Plot	82
5.3.2.1	Box Plot for Longitudinal Data	83
5.3.2.2	Box Plot for Cross-Sectional Data	86
5.3.3	Correlation Plot	87
5.4	Remarks	88
	REFERENCES	90
6	HYBRID MARKOV CHAIN MONTE CARLO AND WIENER PROCESS IN RAILWAY TRACK GEOMETRY DEGRADATION ANALYSIS	91
6.1	Introduction	91
6.2	Data Preprocessing	91
6.3	Metropolis-Hastings Algorithm Implementation and Output Analysis	94
6.4	Wiener Process Sample Paths	109
6.5	Model Validation	114
6.6	Remarks	115
	REFERENCES	117

7	FIRST HITTING TIME IN RAILWAY TRACK GEOMETRY DEGRADATION ANALYSIS	118
7.1	Introduction	118
7.2	Data Preprocessing	118
7.3	Estimation of the FHT for TQI	119
7.3.1	Confidence Limits for FHT	122
7.4	FHT for Raw Data	124
7.5	Remarks	125
8	CONCLUSIONS AND RECOMMENDATIONS	126
8.1	Introduction	126
8.2	Conclusions	126
8.3	Recommendations	128
Appendix		
A	EXPLORATORY DATA ANALYSIS PLOTS	130
A.1	Foot-By-Foot Measurements	130
A.2	Box Plots for Longitudinal Data	139
B	MCMC OUTPUT PLOTS	143
B.1	Output Plots for 500-foot sections	143
C	PERMISSIONS	157

LIST OF TABLES

2.1	Some track quality indices	16
2.2	Selected references on track geometry degradation models	24
3.1	Differences between Bayesian and frequentist approaches	32
3.2	Conjugate prior distributions (Attoh-Okine, 2017)	34
5.1	Sample of longitudinal data	77
5.2	Sample of cross-sectional data	78
6.1	Output summary for crosslevel track sections in the geocell zone . .	99
6.2	Output summary for surface left (124 ft) track sections in the geocell zone	102
6.3	Output summary for alignment left (124 ft) track sections in the geocell zone	105
6.4	Output summary for warp (62 ft) track sections in the geocell zone	108
6.5	Mean squared error for testing data set	115
7.1	Skewness and kurtosis for FHT	122
7.2	Illustration of the confidence intervals for the FHT for surface data	123

LIST OF FIGURES

1.1	Train accidents over ten years by major cause from FRA data	2
1.2	Research approach	4
2.1	Typical track structure (Attoh-Okine, 2017)	10
2.2	Track degradation models classification	11
2.3	Classification of rail track degradation based on structural and geometrical components	12
2.4	Some track geometry parameters	14
2.5	Classification of track geometry degradation models	17
2.6	Linear representation of track geometry degradation and restoration	18
2.7	Nonlinear representation of track geometry degradation and restoration	19
3.1	Interpretation of confidence and credible intervals	32
3.2	Illustration of the Bayesian estimation process	33
3.3	Representation of accept-reject regions	39
3.4	Representation of mutation operator	42
3.5	Representation of crossover operator	43
3.6	Representation burn-in period	44
3.7	Kernel plot	45
3.8	Trace plot	45

3.9	Autocorrelation plot	46
4.1	Representation of a degradation process with failure threshold . . .	51
4.2	Illustration of a random variable defined as a function in the sample space S and takes values on \mathbb{R}	53
4.3	Kernel density for surface right 62 ft	54
4.4	Trajectory of a ODE initial condition $x(0)$	55
4.5	Schematic representation of a trajectory of a ODE and a SDE given initial condition $x(0)$	56
4.6	Realization of the Brownian motion	62
4.7	Wiener process sample paths at different number of steps (N) . . .	64
4.8	Schematic representation of the threshold-regression model	65
4.9	Representation of track geometry with shock events	68
5.1	Illustration of foot-by-foot measurements for some track geometry parameters	76
5.2	Illustration of surface right (62 ft) data	77
5.3	Illustration of histogram and QQ plot for crosslevel, surface right (62 ft), and alignment right (62 ft) at a specific track location	79
5.4	Illustration of histogram and QQ plot for surface left (62 ft), alignment left (62 ft), and warp (62 ft) at a specific track location	80
5.5	Illustration of histogram and QQ plot for crosslevel, surface right (62 ft), and alignment right (62 ft) at a specific inspection date	81
5.6	Illustration of histogram and qq plot for surface left (62 ft), alignment left (62 ft), and warp (62 ft) at a specific inspection date	82
5.7	Illustration of box plot for crosslevel	84
5.8	Illustration of box plot for surface right (62 ft)	85

5.9	Illustration of box plot for alignment right (62 ft)	86
5.10	Illustration of box plot for selected geometry parameters at a specific inspection time	87
5.11	Correlation matrix and scatter plots between selected track geometry parameters at a specific inspection date	88
6.1	General overview for the implementation of the hybrid Bayesian-Wiener process for a single track geometry parameter	91
6.2	Degradation plot for multiple track geometry parameters at a specific 150-foot track section	92
6.3	Degradation plot for multiple track geometry parameters at a specific 500-foot track section	93
6.4	Crosslevel: MCMC posterior plots for drift parameter	97
6.5	Crosslevel: MCMC posterior plots for diffusion parameter	98
6.6	Surface left (124 ft): MCMC posterior plots for drift parameter	100
6.7	Surface left (124 ft): MCMC posterior plots for diffusion parameter	101
6.8	Alignment Left (124 ft): MCMC posterior plots for drift parameter	103
6.9	Alignment Left (124 ft): MCMC posterior plots for diffusion parameter	104
6.10	Warp (62 ft): MCMC posterior plots for drift parameter	106
6.11	Warp (62 ft): MCMC posterior plots for diffusion parameter	107
6.12	Sample paths and observed data for selected track section for crosslevel	110
6.13	Sample paths and observed data for selected track section for standard deviation of surface left (124 ft)	111
6.14	Sample paths and observed data for selected track section for standard deviation of alignment left (124 ft)	112

6.15	Sample paths and observed data for selected track section for standard deviation of warp (62 ft)	113
6.16	Illustration of training and testing data sets for a track section for parameter surface left (124 ft)	114
7.1	General overview for the estimation of the first hitting time for a single track geometry parameter	118
7.2	Schematic representation of the threshold-regression model	119
7.3	PDF and CDF of FHT for crosslevel	120
7.4	PDF and CDF of FHT for surface left (124 ft)	120
7.5	PDF and CDF of FHT for alignment left (124 ft)	121
7.6	PDF and CDF of FHT for warp (62 ft)	121
7.7	Theoretical and simulated FHT at four adjacent locations for surface left (62 ft)	124
A.1	Illustration of crosslevel data at multiple inspection dates	130
A.2	Illustration of surface right (124 ft) data at multiple inspection dates	131
A.3	Illustration of surface left (124 ft) data at multiple inspection dates	132
A.4	Illustration of gage data at multiple inspection dates	133
A.5	Illustration of warp (62 ft) data at multiple inspection dates	134
A.6	Illustration of alignment left (62 ft) data at multiple inspection dates	135
A.7	Illustration of alignment right (62 ft) data at multiple inspection dates	136
A.8	Illustration of alignment right (124 ft) data at multiple inspection dates	137
A.9	Illustration of alignment left (124 ft) data at multiple inspection dates	138
A.10	Illustration of box plot for gage	139

A.11	Illustration of box plot for surface left (124 ft)	140
A.12	Illustration of box plot for alignment right (124 ft)	141
A.13	Illustration of box plot for alignment left (124 ft)	142
B.1	Alignment left (62 ft): MCMC posterior plots for drift parameter .	143
B.2	Alignment left (62 ft): MCMC posterior plots for diffusion parameter	144
B.3	Alignment left (124 ft): MCMC posterior plots for drift parameter .	145
B.4	Alignment left (124 ft): MCMC posterior plots for diffusion parameter	146
B.5	Alignment right (62 ft): MCMC posterior plots for drift parameter	147
B.6	Alignment right (62 ft): MCMC posterior plots for diffusion parameter	148
B.7	Alignment right (124 ft): MCMC posterior plots for drift parameter	149
B.8	Alignment right (124 ft): MCMC posterior plots for diffusion parameter	150
B.9	Crosslevel: MCMC posterior plots for drift parameter	151
B.10	Crosslevel: MCMC posterior plots for diffusion parameter	152
B.11	Surface left (124 ft): MCMC posterior plots for drift parameter . .	153
B.12	Surface left (124 ft): MCMC posterior plots for diffusion parameter	154
B.13	Surface right (124 ft): MCMC posterior plots for drift parameter .	155
B.14	Surface right (124 ft): MCMC posterior plots for diffusion parameter	156
C.1	Permission to use Figure 1.1 and Table 3.2	157

ABSTRACT

Globally, track-caused accidents are a major factor of train derailments. Rail fatigue, rail wear, and track geometry defects are examples of track failure mechanisms. These mechanisms are usually modeled separately due to their individual characteristics, so maintenance activities are normally targeted to repair specific track structure components. Modeling track degradation and estimation of the failure time of the track is critical for safety and derailment purposes.

In particular, the use of railway track geometry degradation models has played an important role in railway engineering. It helps in establishing track infrastructure maintenance policies and the output can be used to address derailment potential. Most track geometry degradation models are not stochastic and fail to account for small variations of the degradation values. On the other hand, failure time has been traditionally modeled using defect data. However, unless it is an accident due to extreme events, track geometry reaches a threshold as a result of an underlying degradation process. This dissertation focuses on the formulation of track geometry degradation and its first hitting time, in which two case studies were conducted using U.S. Class I railroad inspection data.

The first case study formulates track geometry degradation as a Wiener process. The Wiener process is a stochastic process that models degradation for non-strictly monotonic increasing functions. Based on the characteristics of the track geometry data, the Wiener process appears to be suitable for modeling the degradation process. The model parameters were estimated using an adaptive Markov chain Monte Carlo algorithm. The second case study estimates the first hitting time (FHT) for each track geometry parameter and track section. The FHT is referred to as the probability distribution of the time at which the degradation path first reaches a safety threshold.

The underlying degradation path is modeled as a Wiener process with drift and the FHT follows an inverse Gaussian distribution.

Results from this dissertation provide a better understanding of track geometry degradation and failure by accounting for the inherent uncertainty in this process and by providing an alternative approach to identify track sections that require more attention for maintenance activities, considering each track geometry parameter.

Chapter 1

INTRODUCTION

1.1 Background & Motivation

The railway network is one of the most important transportation systems in the United States. It is composed of approximately 140,000 miles of private-owned track and moves nearly 40% of the U.S. freight (FRA, 2017). Maintaining railway infrastructure is critical. To keep infrastructure in good condition, railroads in 2015 invested about \$27.1 billion in maintaining, upgrading, and expanding the network, representing 42% of the total annual expenditure (AAR, 2016). These efforts for maintaining and improving railway infrastructure influence, among others, the number of accidents due to the condition of the infrastructure. Figure 1.1 presents the breakdown of the major cause of train accidents for a ten-year time frame from the Federal Railroad Administration (FRA) data. It can be seen from the figure that overall the total number of accidents and those resulting in train derailments was reduced. Despite that reduction, the ranking of the major causes of accidents was constant; track is the second most reported cause, after human factor.

A degraded track affects ride comfort and may lead to accidents if not maintained on time. Track geometry is a key component for keeping track infrastructure in good condition, and the data collection process has a variety of uses (Lindamood et al., 2003): (i) it helps to identify defects that are present in the track, (ii) it contributes to the assessment of track condition, that is, track quality indices (TQIs), and (iii) it can be used as an input for forecasting models.

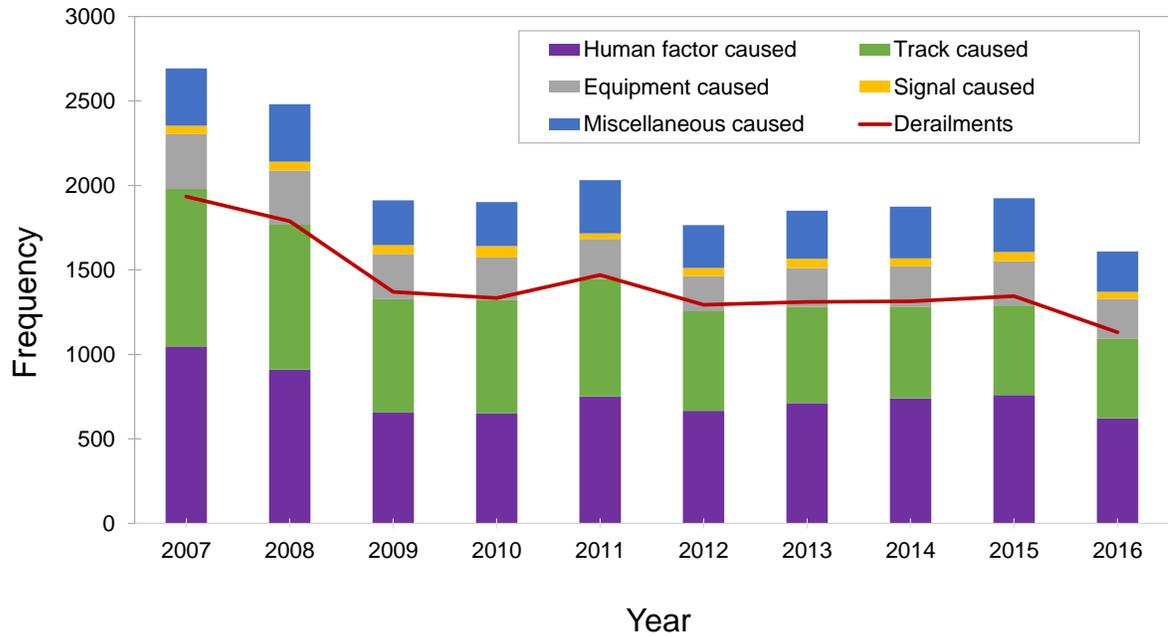


Figure 1.1: Train accidents over ten years by major cause from FRA data

1.2 Statement of the Problem

Predicting track geometry degradation and failure time require accounting for the inherent uncertainty of the unobserved condition of the track. Traditionally, track geometry degradation is quantified using foot-by-foot measurements and failure time is estimated using track geometry defect data only. When degradation data is available as a function of time or cumulative tonnage, extrapolation of the degradation paths can be mathematically modeled as a stochastic process and the estimation of the time-to-failure can be obtained for predefined confidence limits. There is, therefore, a need to address questions regarding the prediction of track geometry degradation and failure as a stochastic process that can provide insights of how new safety standards can be created.

1.3 Objective of the Study

Based on the statement of the problem presented in Section 1.2, the main objective of this dissertation is to formulate and implement a threshold-regression model

for analyzing track geometry degradation and its failure time. To achieve this main objective, the following sub-objectives are defined:

1. To conceptualize track geometry degradation as a stochastic process
2. To formulate and implement Markov chain Monte Carlo for parameter estimation of track geometry degradation model
3. To estimate the time at which the degradation paths first reach a threshold
4. To conduct analysis of track geometry degradation and failure using United States track geometry inspection data

1.4 Research Approach

This dissertation was conducted in three stages as presented in Figure 1.2. The first stage deals with the conceptualization of the mathematical approaches for estimating track geometry degradation model parameters and stochastic processes for predicting track degradation. Based on literature review in various disciplines, not only railway track geometry, Bayesian approaches and stochastic processes were selected for analyzing the available data set. The second stage examines in more detail the different approaches for both Bayesian approaches and Wiener process. In this stage, a variation of the Metropolis-Hastings algorithm and the Wiener process with drift for the prediction of track geometry degradation and the estimation of the first hitting time were considered. Finally, the third stage consists of the implementation of the designed methodology in railway infrastructure, with a main focus on track geometry degradation.

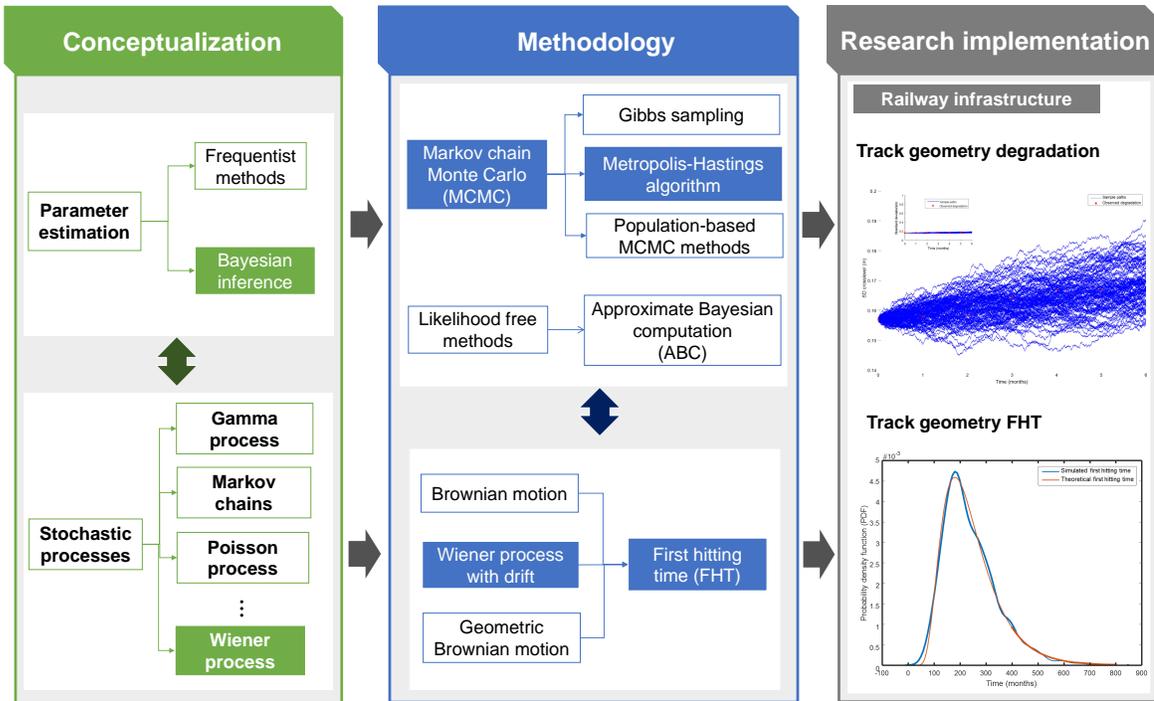


Figure 1.2: Research approach

1.5 Dissertation Structure

1.5.1 Chapter 1: Introduction

This chapter presents the motivation of this research, as well as the structure and the contributions of this dissertation.

1.5.2 Chapter 2: Background

This chapter presents the definitions of track geometry parameters and a literature review on track geometry degradation models. Also, the gaps in the literature are discussed.

1.5.3 Chapter 3: Bayesian Inference

This chapter deals with the definition of Bayesian inference for parameter estimation. The fundamental differences with frequentist methods are discussed. Also, this chapter presents three variations of the Markov chain Monte Carlo algorithms: Metropolis-Hastings algorithm, Gibbs sampling algorithm, and population-based Markov

chain Monte Carlo methods. This chapter also presents the contributions of parameter estimation in track geometry degradation models using Markov chain Monte Carlo methods. Finally, an illustration of an output analysis in the context of track geometry degradation is presented.

1.5.4 Chapter 4: Wiener Process for Degradation Analysis

The aim of this chapter is to introduce the concepts of stochastic process and highlights the importance of using the Wiener process in track geometry degradation and the estimation of the first hitting time. Also, this chapter formulates the Wiener process and first hitting time in track geometry degradation.

1.5.5 Chapter 5: Exploratory Data Analysis

The goal of this chapter is to describe the data set utilized in this research, and to present the findings of the exploratory data analysis for the data used.

1.5.6 Chapter 6: Hybrid Markov Chain Monte Carlo and Wiener Process in Railway Track Geometry Degradation Analysis

This chapter presents the implementation of the hybrid Markov chain Monte Carlo and Wiener process using the data set available in this research. A detailed discussion of the findings are also presented.

1.5.7 Chapter 7: First Hitting Time in Railway Track Geometry Degradation Analysis

This chapter presents the estimation of the first hitting time using the predicted sample paths from Chapter 6. Also, the discussion of the results for track geometry parameters and their confidence levels are highlighted.

1.5.8 Chapter 8: Conclusions and Recommendations

This chapter summarizes the research conducted. Recommendations and future work are discussed.

1.6 Contributions of the Dissertation and Summary of Publications, Conferences, and Awards

Contributions of this dissertation appear in the following peer-review papers and conferences.

1.6.1 Peer-Review Papers

1. **S. Galván-Núñez** and N. Attoh-Okine. “A Threshold-Regression Model for Track Geometry Degradation,” Submitted to Journal of Rail and Rapid Transit, 2017.
2. **S. Galván-Núñez** and N. Attoh-Okine. “Hybrid Markov Chain Monte Carlo and Wiener Process in Railway Track Geometry Degradation Analysis,” Submitted to IEEE Transactions on Reliability, 2017.
3. **S. Galvan-Nunez** and N. Attoh-Okine, Hybrid Particle Swarm Optimization and K-Means Analysis for Bridge Clustering Based on National Bridge Inventory Data,” ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng., p. F4016001, 2016.
4. **S. Galván-Núñez** and N. Attoh-Okine. “Approximate Bayesian Computation for Parameter Estimation in Track Degradation Models,” 2017. (In Progress)

1.6.2 Conferences and Presentations

1. **S. Galvan-Nunez** and N. Attoh-Okine. “Estimating the First Hitting Time for Track Geometry Degradation,” 13th Annual Inter-University Symposium on Infrastructure Management, West Lafayette, IN, United States, 2017.
2. **S. Galvan-Nunez** and N. Attoh-Okine. “Predicting Track Geometry Degradation Using Wiener Process,” Delaware Data Science Symposium, Newark, DE, United States, 2017.
3. **S. Galvan-Nunez** and N. Attoh-Okine. “A Wiener Process for Modelling Railway Track Geometry Degradation,” 7th Annual Grad Forum, University of Delaware, Newark, DE, United States, 2017.
4. **S. Galvan-Nunez** and N. Attoh-Okine. “Assessing Uncertainty of Track Geometry Degradation Based on Evolutionary Markov Chain Monte Carlo,” Transportation Research Board 95th Annual Meeting, Washington, D.C., United States, 2017.

5. **S. Galvan-Nunez**. “Women of Color and the PhD: A Latin American Perspective,” “ADVANCE Conference—Women of Color in the Academy: What’s Next?”, Newark, DE, United States, 2016.
6. **S. Galvan-Nunez** and N. Attoh-Okine. “Assessing Uncertainty of Track Geometry Degradation Based on Evolutionary Markov Chain Monte Carlo,” 12th Annual Inter-University Symposium on Infrastructure Management, Still Water, OK, United States, 2016.
7. **S. Galvan-Nunez** and N. Attoh-Okine. “Markov Chain Monte Carlo in Rail Track Defect,” Transportation Research Board 95th Annual Meeting, Washington, D.C., United States, 2016.
8. N. Attoh-Okine and **S. Galvan-Nunez**. “Introducing Big Data Concepts in Resilience Engineering Analysis,” 2nd National Symposium on Resilient Critical Infrastructure (NSRCI), Philadelphia, PA, United States, 2015.
9. **S. Galvan-Nunez** and N. Attoh-Okine. “Markov Chain Monte Carlo in Rail Degradation Analysis,” 11th Annual Inter-University Symposium on Infrastructure Management, Alexandria, VA, United States, 2015.
10. **S. Galvan-Nunez** and N. Attoh-Okine. “Metaheuristics in Big Data: An Approach to Railway Engineering,” IEEE International Conference on Big Data, Washington D.C, United States, 2014.

1.6.3 Awards

1. Best Presentation. 13th Annual Inter-University Symposium on Infrastructure Management, West Lafayette, IN, United States, June 2017.
2. Best Presentation. 12th Annual Inter-University Symposium on Infrastructure Management, Still Water, OK, United States, June 2016.
3. Best Presentation. 11th Annual Inter-University Symposium on Infrastructure Management, Alexandria, VA, United States, May 2015.
4. Professional Development Award, University of Delaware, United States, October 2014.

REFERENCES

- AAR. Total Annual Spending 2015 Data, 2016. URL <https://www.aar.org/FactSheets/Safety/AARAnnualSpending{ }2016Update{ }7.15.16.pdf>.
- FRA. FRA (Federal Railroad Administration), 2017. URL <https://www.fra.dot.gov/Page/P0001>.
- Brian Lindamood, James Strong, and James McLeod. Railway Track Design. In *Practical Guide to Railway Engineering*, volume 6, page 48. 2003. ISBN 2160-2514. URL <https://www.arema.org/publications/pgre/>.

Chapter 2

BACKGROUND

2.1 Introduction

Railway is one of the most important freight and passenger transportation systems around the world, and in order to keep track infrastructure in good conditions, different studies have focused on maintenance and track degradation prediction models. Railway track geometry degradation is a phenomenon that if not detected and corrected on time, can lead to extensive maintenance costs and safety issues. Despite the extensive contributions in the literature regarding track geometry degradation models, there is still no consensus about the model parameters that best represent degradation process. It has been shown that model parameters selection has a major influence on the accuracy of the track degradation prediction used for long-term maintenance decisions. The aim of this chapter is to present a state-of-the-art survey on track geometry degradation models, highlighting their main characteristics including their shortcomings. The outcome of the survey will provide researchers and practitioners a better understanding of track geometry behavior that can be used for maintenance planning and other decision making. This chapter also includes a general overview of the track components and failure mechanisms.

2.2 Track Characterization

A ballasted track can be classified in two groups: (i) track superstructure, that consists of the rail, fastening system and ties, and (ii) track substructure that includes the ballast, subballast and subgrade. Figure 2.1 illustrates the track components.

- **Rail:** rails are defined as longitudinal steel that are designed to provide support and guide the train. For ballasted track, the most common rail shape is the standard tee rail section.

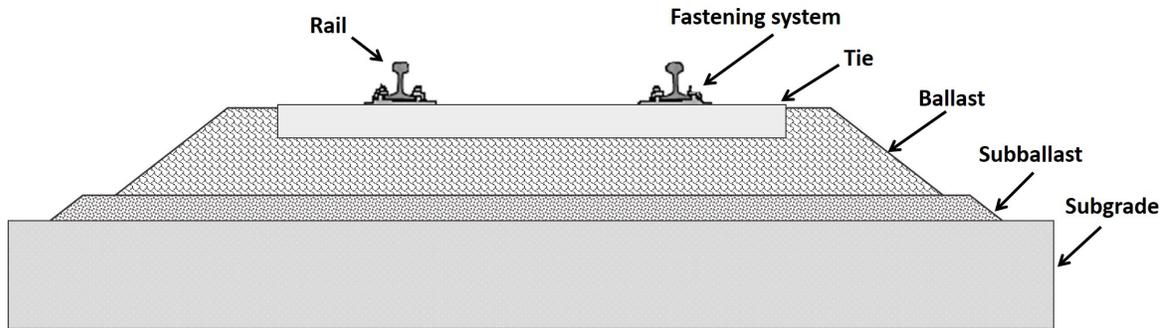


Figure 2.1: Typical track structure ([Attoh-Okine, 2017](#))

- **Ties:** ties provide support to the rails and contribute to keep the designed rail gauge. Ties are typically made of wood (the most common in the United States), concrete or steel.
- **Fastening system:** fasteners are used to firmly hold rails on top of the ties to ensure they not move vertically, longitudinally or laterally ([Indraratna et al., 2011](#)). For wood tie, the most common fastener is cut spike, whose are distributed two spikes per plate and up to four for sharp curves. For concrete ties, elastic fastenings are used at the rail hold down location in order to provide sufficient tie load to the rail base and to provide longitudinal restraint for the rail.
- **Ballast:** ballast is defined as the coarse aggregate that is place under the ties and above the subballast and it is designed to resist applied vertical, lateral, and longitudinal loads. It is also used to maintain track position, so track deformation can be prevented. In addition, it provides rapid drainage, that is, without a good drainage system, there is a high risk of the void areas to fill by the action of the combination of water and small particles.
- **Subballast:** subballast is the granular layer placed below the ballast and above the subgrade. Similar as the ballast, the subballast contributes to load and stress reduction, provides drainage, protects from frost to the subgrade, etc.
- **Subgrade:** subgrade, the foundation upon which everything above depends for support, is often the most variable and potentially the weakest of track components ([Li et al., 2015](#)).

2.3 Track Failure Mechanisms

When referring to track degradation, there is a wide range of possible failure mechanisms that can influence track performance. Based on that, there is more than one definition of track degradation, that is why it is important to specifically define

which part of the track is subject to degradation, what their key parameters are, and what the relationship with other track components is. In the literature there are different classifications of track degradation models. [Oberg \(2006\)](#) presents an extensive review of different track degradation models and classifies them in five groups (Figure 2.2):

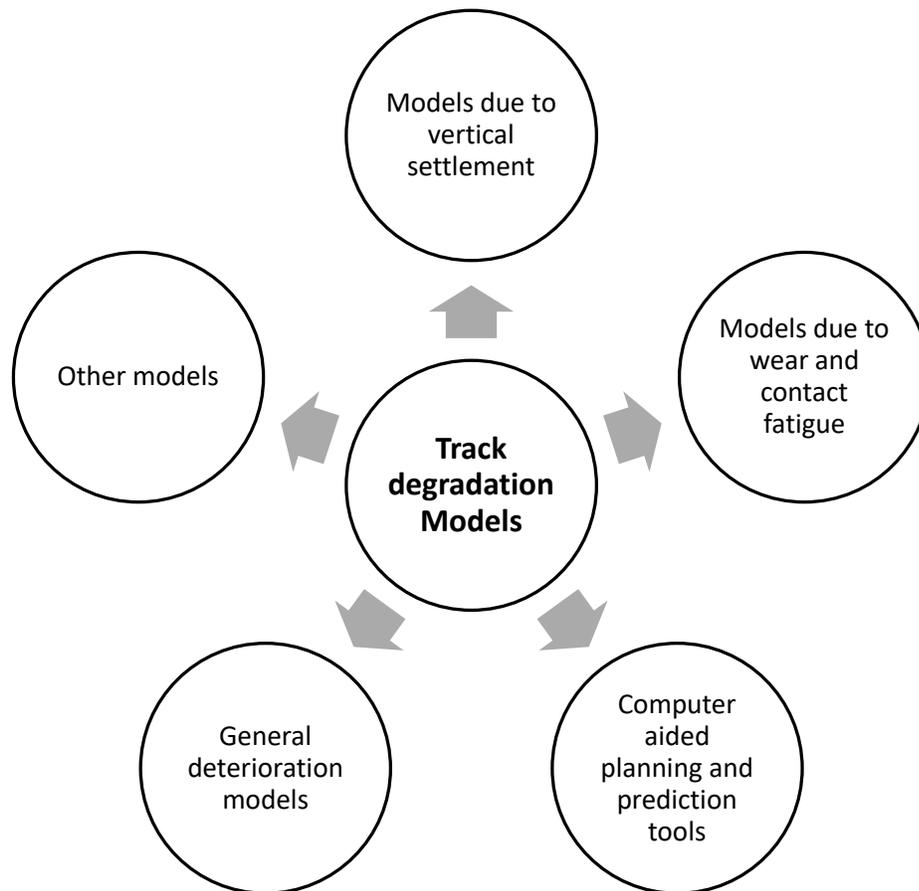


Figure 2.2: Track degradation models classification

- Models due to vertical settlement for ballasted tracks
- Models due to wear and contact fatigue: this type of models only consider surface defects of rails
- Computer aided planning and prediction tools: this group presents the computational tools or software used for predicting track degradation, as well as scheduling and cost analysis for track maintenance

- General degradation models: they take under consideration more than one track failure mechanism, including vertical settlement, wear and fatigue
- Other models: they attempt to update track maintenance activities due to changes in the traffic condition

Another relevant classification on track failure mechanisms is presented by [Zarembski and Palese \(2006\)](#). The authors discussed three categories of track failure or degradation that compromise safety as shown below:

- Broken rail risk model: this model intends to estimate the probability of a broken rail that allows railways to schedule ultrasonic test activities to reduce that probability
- Track buckling risk model: this model allows to identify and classify track locations of high risk of buckling (track lateral misalignment mainly caused by high compressive forces, weakened track conditions, and vehicle loads)
- Vehicle/track geometry risk model: this model allows to identify and prioritize track locations with high potential of vehicle/track geometry related derailments

On the other hand, [He et al. \(2013\)](#) classify track degradation in two main groups called structural and geometry defects (Figure 2.3).

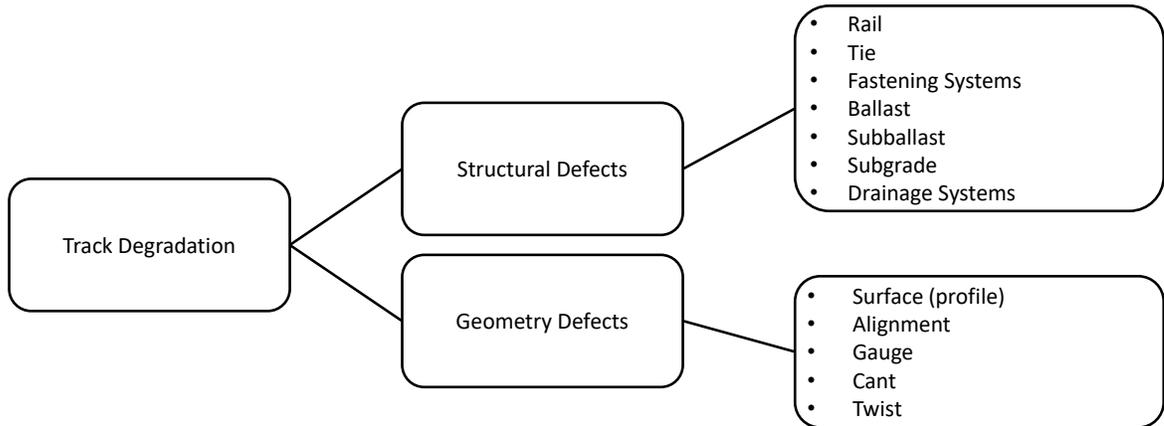


Figure 2.3: Classification of rail track degradation based on structural and geometrical components

2.3.1 Track Structural Defects

Structural defects are those that are generated by the structural components of the track, such as rail, ties, fastening systems, ballast, subballast, subgrade and drainage systems. There are different failure mechanisms for each one these components, which if not detected and corrected on time, can lead to the reduction of the track system performance or in a derailment.

2.4 Track Geometry

In the United States, deviation from the allowable standards for track geometry measurements are established by the Federal Railroad Administration (FRA), although railroads usually consider narrower standards. Track geometry defects are obtained by visual inspection, and the frequency follows FRA requirements, usually two times per week for mainline in Class 4 and Class 5. However, automatic inspection is widely used for measuring track geometry, which is performed by using a track geometry car, and defects are defined based on FRA standards. The FRA operates the Automated Track Inspection Program (ATIP), in which the ATIP cars use the Track Geometry Measurement System (TGMS) whose main function is to generate signals collected from the measurement and subsequently perform an online signal processing, producing as an output a graphic report regarding track geometry measurements. Track geometry defects are identified and displayed in tabular form. Once the data are collected and processed, the ATIP cars compare the current track condition and compare it to the FRA standards.

2.4.1 Track Geometry Parameters

Track geometry can be defined based on the following parameters ([U.S Army, 2008](#)) (Figure 2.4).

- Alignment: alignment is the relative position of the rails in its horizontal plane, measured at the midpoint of a 32-foot, 62-foot, or 124-foot chord. For tangent track, the alignment is equal to zero. For curved track, the alignment is equal to the degree of curvature.

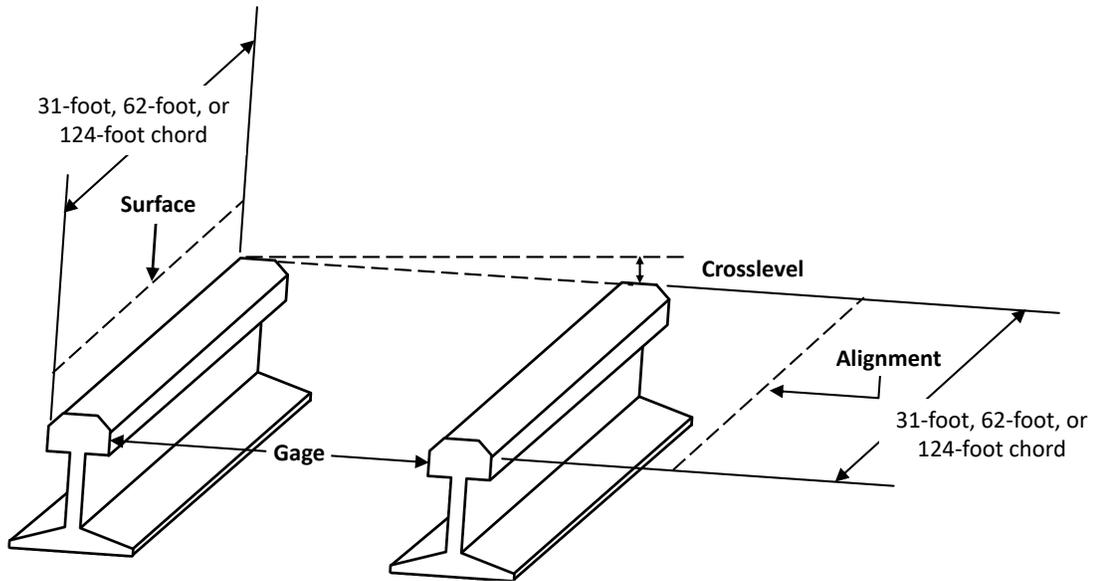


Figure 2.4: Some track geometry parameters

- Crosslevel: crosslevel is the difference in elevation between the top surface of the two rails at any point of railroad track. Crosslevel is measured at right angles to the track.
- Gage: gage is the distance between two rail heads at right angles to the rails in a plane five-eighths of an inch below the top of the rail head. In the United States the distance used is standard gage which is equal to 56.5 inches for tracks containing up to 12 degrees of curvature.
- Surface or Profile: surface (longitudinal leveling) is the relative elevation of the two rails along the track. The profile measurement is usually carried out at the midpoint of the 32-foot, 62-foot, or 124-foot mid chord.
- Twist: twist is the difference in crosslevel between two points of a fixed distance.
- Warp: warp is the difference in crosslevel between any two points less than or equal to 62 feet apart.

2.5 Track Maintenance Methods

This section deals with the description of some track maintenance techniques.

- Tamping: this technology has traditionally been used in the United States to correct surface geometry defects in North America. Tamping consists on squeeze the ballast under the ties using tamping tools, correcting track geometry surface. It is a quick technology to implement; however, the ballast is disturbed as a result of the insertion of tamping tools ([Zarembski and Newman, 2008](#)).

- Design over-lift tamping: this alternative consists on lifting the track over the design profile in order to compensate the rapid ballast settlement that occurs once the traffic resumes, so the ballast settlement reaches the design profile. The advantage of using this alternative in comparison to conventional tamping is that when conventional tamping actions are performed, the ballast settles to the prior rough shape once the traffic resumes. The opposite occurs with design over-lift tamping that allows a better shape after traffic resumes.
- Stone blowing: this technology is defined as the process of track geometry correction that adds rocks to the surface of the ballast right under the lifted tie. The idea of this alternative is that ballast particles are rearranged in order to fill the voids under the lifted tie. The advantages of performing stone blowing over conventional tamping are the following: (i) because stone blowing pre-measure track geometry, the stone blower performs corrections only in the place where it is needed. This contributes to minimize the ballast requirements, and (ii) the ballast damage is lower compared to conventional tamping, making the geometry correction last longer (Li et al., 2015).

2.6 Track Quality Index (TQI)

Track quality is usually assessed by putting a single or a set of track geometry parameters in a metric or set of metrics. These metrics are known as Track Quality Index (TQI). The TQI provides information regarding track geometry and it is used to determine track interventions, as well as track performance, and also to compare track condition before and after interventions (Fortunato et al., 2007). Each TQI is unique and vary as shown in Table 2.1.

Table 2.1: Some track quality indices

TQI	Equation	Parameters	Comments
J synthetic coefficient (Madejski and Grabczyk, 2002)	$J = \frac{S_x + S_y + S_w + 0.5S_e}{3.5}$	<ul style="list-style-type: none"> • S_z: surface (vertical irregularities) • S_w: alignment (horizontal irregularities) • S_y: twist • S_e: gage 	Considers the standard deviation of surface, alignment, twist, and gage.
Track roughness index (Liu et al., 2015)	$R^2 = \sum_{i=1}^n \frac{d_i^2}{n}$	<ul style="list-style-type: none"> • d_i: deviation i of track geometry parameter • n: number of observations 	Quantifies the quality of the track as the sum of squares of the deviation of a track section divided by the number of observations.
Federal Railroad Administration TQI (Zhang et al., 2004)	$TQI = \left[\frac{L_s}{L_0} - 1 \right] \times 10^6$	<ul style="list-style-type: none"> • L_s: curve traced length • L_0: theoretical length of a track section 	Considers the curve traced length and the theoretical length over a section of track for individual track geometry parameters.
Track geometry index (Mundry, 2000)	$TGI = \frac{2UI+TI+6AI+GI}{10}$	<ul style="list-style-type: none"> • UI: surface (unevenness) • AI: alignment • TI: twist • GI: gage 	Considers the standard deviation of surface, alignment, twist, and gage. Individual track geometry parameters are exponentially distributed.
Overall TQI (Sadeghi and Askarinejad, 2010)	$OTGI = \frac{\frac{5}{2} \times GI^+ + \frac{1}{2} GI^- + b \times AI + c \times PI + d \times TI}{\frac{a+d}{2} + b + c + d}$	<ul style="list-style-type: none"> • GI^+ and GI^-: positive and negative gage indices respectively • AI: alignment index • PI: surface (profile) index • TI: twist index • a, a', b, c and d: model coefficients. 	Quantifies the quality of the track as a function of surface, alignment, gage, and twist. Track geometry parameters are normally distributed.
Sweden TQI (Andersson, 2002)	$150 - 100 \left[\frac{\sigma_H}{\sigma_{Hlim}} + 2 \times \frac{\sigma_S}{\sigma_{Slim}} \right] / 3$	<ul style="list-style-type: none"> • σ_H: standard deviation surface (profile) • σ_S: standard deviation of crosslevel, gage, and alignment • σ_{Hlim}: standard deviation of the allowable σ_H • σ_{Slim}: allowable value of σ_S 	Considers the standard deviation of crosslevel, gage, and alignment.

2.7 Track Geometry Degradation Models

Based on the existing literature, track geometry degradation models can be classified in five groups (Figure 2.5): regression models, exponential smoothing, hierarchical Bayesian, artificial neural networks, and stochastic processes. There is an important number of contributions in this area, where different data analysis techniques have been performed to predict track degradation that can be used as an input for determining optimal maintenance schedule activities. The works of [Dahlberg \(2001\)](#), [Oberg \(2006\)](#), [Guler \(2013b\)](#), and [Soleimanmeigouni et al. \(2016a\)](#) are examples of extensive literature reviews in this area. In this section, an overview of the contributions in literature regarding these track degradation models is presented, highlighting the main characteristics of each model.

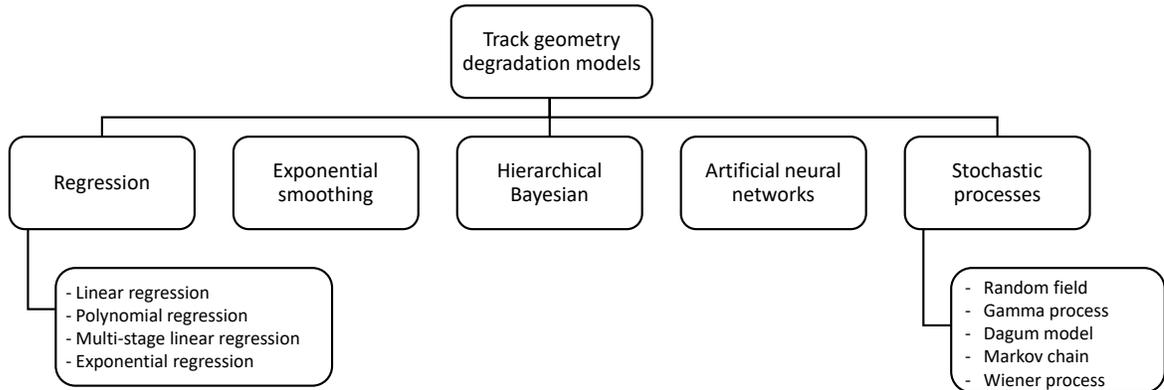


Figure 2.5: Classification of track geometry degradation models

2.7.1 Regression Models

Different types of regression models have been used in track geometry degradation. Examples of regression models include: linear regression, polynomial regression, multi-stage regression, and exponential regression. Figures 2.6 and 2.7 show an example of linear and nonlinear regression.

In the linear regression approach, it is assumed that the track degrades at a constant rate (the gradient of the equation). The constant of the equation is based on the TQI value after restoration. Generally, the intersection of the linear equation

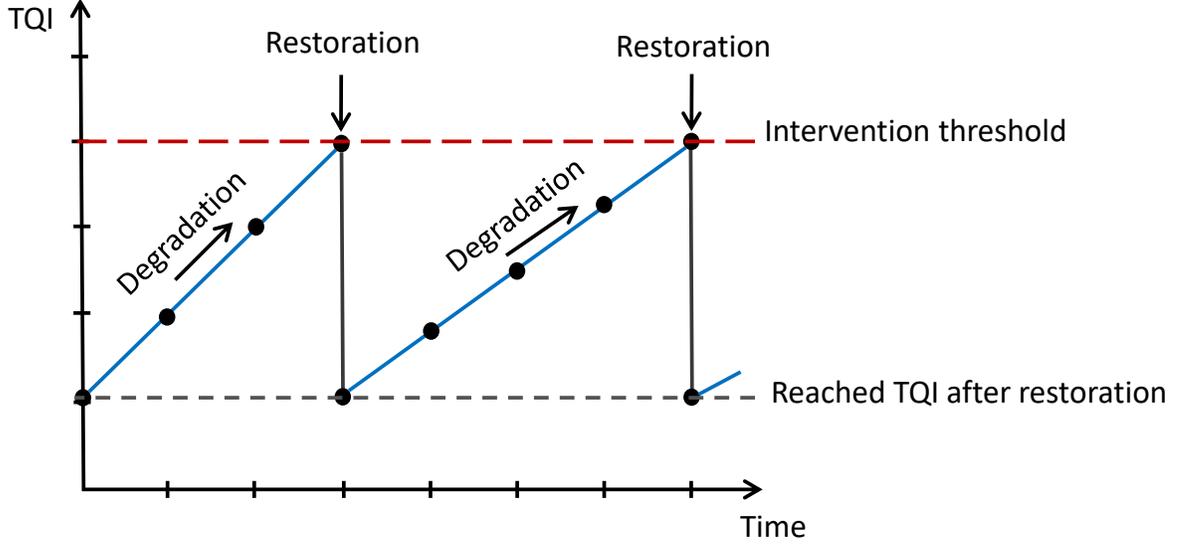


Figure 2.6: Linear representation of track geometry degradation and restoration

and the intervention threshold serves as the restoration time. [Chang et al. \(2010\)](#) highlighted the characteristics of the track geometry degradation model using multi-linear components of the track models. The authors defined the track geometry in terms of three elements: (i) periodicity, defined as the change of surface values over the track, which are the same between two adjacent track maintenance activities, (ii) multi-stage, this is the case when degradation rates vary from the initial to final maintenance cycle, and (iii) exponential characteristics, which is the case when the degradation model is in the form of an exponential function. Equation (2.1) shows the multi-stage model

$$\sigma_{TLD} = a_n + b_n T, \quad (2.1)$$

where:

- σ_{TLD} : standard deviation of parameter surface
- b_n : slope of line n
- a_n : intercept of line n
- T : cumulative passing tonnage from last maintenance to the present day

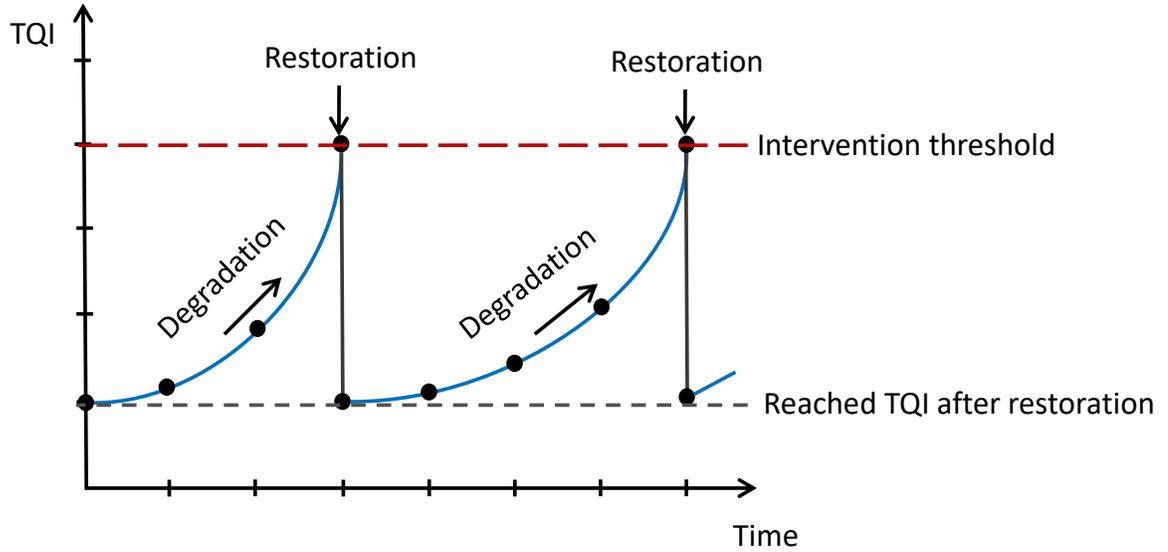


Figure 2.7: Nonlinear representation of track geometry degradation and restoration

Andrade and Teixeira (2011) presented track geometry degradation as a function of the standard deviation of surface and the cumulative tonnage as presented in equation 2.2.

$$\sigma_{LD} = c_1 + c_0T, \quad (2.2)$$

where:

- σ_{LD} : standard deviation of surface
- c_1 : initial standard deviation of surface right after upgrade actions
- c_0 : degradation rate
- T : cumulative tonnage after track upgrade (100 MGT)

The authors used Monte Carlo simulation to address the uncertainty on the model regarding tamping cycle for each track section group and specific quality levels. The research concluded that bridges and switches require tamping activities more often compared to stations and tangent track.

In a follow up paper, Andrade and Teixeira (2012) used Markov Chain Monte Carlo (MCMC) for the estimation of the degradation model parameters. The aim of

this analysis was to assess the uncertainty behavior of the track infrastructure through its life cycle.

[Audley and Andrews \(2013\)](#) similarly presented a linear model (equation 2.3) and used the maximum likelihood method to estimate the model parameters.

$$\sigma = A + Bt, \tag{2.3}$$

where:

- σ : surface (vertical alignment)
- A : intercept
- B : degradation rate

[Xu et al. \(2011\)](#) approximated the nonlinear track geometry degradation for each tamping cycle for multiple linear models selected in a short-range section. Each linear model is updated based on the change of track geometry degradation historical data. The linear model parameters were estimated using the frequentist approach least squares.

[Berawi et al. \(2010\)](#) and [Quiroga and Schnieder \(2012\)](#) developed nonlinear models in track geometry degradation. [Berawi et al. \(2010\)](#) assumed nonlinearity of track geometry degradation in terms of the TQIs and they compared their results to the European Standard EN 13848-5, the TQIs J synthetic coefficient, and TGI. In their formulation, [Quiroga and Schnieder \(2012\)](#) assumed an exponential form, considering parameter surface as the dependent variable of the model.

2.7.2 Exponential Smoothing

Simple exponential smoothing is more appropriate for forecasting a time series, where there is absence of trend or seasonal pattern, and the mean slowly changes over time. In general, exponential smoothing methods give larger weights to more recent observations, and the weights decrease exponentially as the observations become more distant, that is, the older observations. [Oyama and Miwa \(2006\)](#) proposed a model for

measuring the parameter surface and predicting track maintenance operation effects based on a logistic distribution. The exponential smoothing was utilized to predict the increasing trend parameters associated with the standard deviation of logistic distribution and to express the typical characteristic of track surface. The authors' model is accurate enough to estimate the future changes of track surface by forecasting a selected parameter value which is the function of a logistic parameter and an integer greater or equal to zero.

2.7.3 Hierarchical Bayesian Model

Hierarchical Bayesian model is a statistic technique based on graph-probabilistic models. The technique is presented in a form of a directed acyclic graph in multiple layers. In many applications, the model parameters have significant dependencies on each other and there is strong need to find and address these dependencies. In the hierarchical Bayesian methods, the directed graph helps to factorize the model using the independent properties of the graph. Once the full conditionals are established, the posterior distributions can be estimated.

[Andrade and Teixeira \(2013\)](#) used the hierarchical Bayesian approach in track geometry degradation. The authors assumed that the standard deviation of parameter surface is normally distributed, and the mean is composed by the following elements: (i) constant linear evolution with MGT, (ii) initial standard deviation of surface, (iii) disturbance effect of the initial standard deviation of surface after each tamping operation, and (iv) renewed track and non-renewed track sections. Non-informative priors using inverse gamma distributions were assumed by the authors. The model proposed by [Andrade and Teixeira \(2013\)](#) alerted railway researchers and practitioners to address spatial dependencies in rail track geometry degradation modeling and analysis.

2.7.4 Neural Networks

Neural networks (NN) are biologically inspired machine learning techniques. NN are data-driven self-adaptive, allowing few assumptions regarding the model ([Zhang](#)

et al., 1998). Guler (2013a) used neural networks to model track geometry degradation. In particular, the author applied the methodology to study the effects of the degree of curvature, crosslevel, speed, age, rail type, rail length, etc. in the neural network approach. However, the author did not fully discuss the learning approach, as well as the criteria for the selection of the appropriate parameters for training, testing, and validation. It is very important to note that there are different NN architectures, but it is not clear which of them will be more appropriate for track degradation models. This warrants further studies.

2.7.5 Stochastic Process

Stochastic processes account for uncertainty in track geometry degradation Chapter 4 presents a detailed definition on this topic. Selected contributions in this area are presented below.

Iyengar and Jaiswal (1995) used a random field model in track geometry degradation. The authors used a stationary Gaussian random field defined with the power spectra density function applied to Indian railways tracks. In terms of Markov models, Shafahi and Hakhamaneshi (2009) predicted track degradation in Iranian railways by defining six states. The transition matrix was defined using historical data composed of about 2627 miles (4228 kilometers) in the time interval 2002-2006. The results were utilized as an input for a dynamic programming model which was used to schedule track maintenance activities. Yousefikia et al. (2014) also presented a Markov model to assess tram track condition and predict maintenance actions for Melbourne tram track data. The Markov chain state space is composed by the following:

- Normal
- Maintenance limit: degraded failure undetected
- Maintenance limit: degraded failure inspected
- Action limit: degraded failure undetected
- Action limit: degraded failure inspected

- Safety limit
- Repaired

[Meier-Hirmer et al. \(2006\)](#) presented a Gamma process for track degradation and a classification method based on regression trees using environmental variables such as type and height of ballast, maximum speed, weather conditions, type of rail, accumulated tonnage since ballast renewal, among others. In the paper, the authors considered surface (longitudinal leveling) defects as the track failure mechanism for track geometry defects. [Mercier et al. \(2012\)](#) implemented an extended version of [Meier-Hirmer et al. \(2006\)](#) by modeling track geometry degradation for parameter surface, using a bivariate Gamma process where the maximum likelihood estimation was implemented to estimate the model parameters. In addition, [Vale and Lurdes \(2013\)](#) performed a stochastic model based on the Dagum distribution, which is used for describing track geometry degradation process over time. Dagum distribution is a function of the input data and the model parameters. For parameter estimation, the authors used the maximum likelihood method.

[Soleimanmeigouni et al. \(2016b\)](#) modeled track degradation as a Wiener process, considering as the failure mechanism the standard deviation of the parameter surface. The main limitation of this paper is that the model parameters were estimated using maximum likelihood method which is a deterministic method that provides a single point estimate, instead of the probability distribution of all the possible values for the parameters.

Table 2.2: Selected references on track geometry degradation models

Reference	Degradation model	Model variables
Iyengar and Jaiswal (1995)	Random field	- Surface (absolute vertical profile) - Unevenness data
Meier-Hirmer et al. (2006)	Gamma process	Surface (longitudinal leveling)
Oyama and Miwa (2006)	Exponential smoothing	- Standard deviation of logistic distribution/track surface irregularities
Chang et al. (2010)	Multi-stage linear regression	- Standard deviation of surface (longitudinal level irregularity) - Slope of line - Intercept - Cumulative passing tonnage from last maintenance to the present day
Berawi et al. (2010)	Nonlinear regression	Surface, alignment, twist, and gage
Xu et al. (2011)	Linear regression	Standard deviation of gage, crosslevel, surface, alignment, and twist
Mercier et al. (2012)	Bivariate gamma process	Surface (longitudinal leveling) and alignment (transverse leveling)
Quiroga and Schnieder (2012)	Exponential function	- Surface (longitudinal leveling) - Time - Log-normally distributed variable - Measurement noise - Number of tamping interventions
Andrade and Teixeira (2011)	Linear regression	- Initial standard deviation after renewal or tamping operations - Degradation rate (mm/100 MGT) - Cumulative tonnage between tamping operations (100 MGT)
Andrade and Teixeira (2012)	Linear regression	- Initial standard deviation after renewal or tamping operations - Degradation rate (mm/100 MGT) - Cumulative tonnage between tamping operations (100 MGT)
Andrade and Teixeira (2013)	Hierarchical Bayesian	- Standard deviation of surface (longitudinal leveling) - Accumulated tonnage since last tamping or renewal operations - Degradation rate - Initial standard deviation of surface (longitudinal level defects) - Initial standard deviation of surface (longitudinal level) after each tamping operations - Number of tamping operations performed since last renewal
Vale and Lurdes (2013)	Dagum model	- Standard deviation of surface (longitudinal level) Dagum distribution shape and scale parameters
Audley and Andrews (2013)	Linear regression	surface (vertical alignment)
Yousefikia et al. (2014)	Markov chain	Markov chain states: - Normal - Maintenance limit. Degraded failure undetected - Maintenance limit. Degraded failure inspected - Action limit. Degraded failure undetected - Action limit. Degraded failure inspected - Safety limit. Repaired
Guler (2013a)	Artificial neural networks	- Gradient (%) - Curvature (1/R)(1/m) - Crosslevel (mm) - Speed (km/h) - Age (years) - Rail type (kg/m) - Rail length (m) - Tie type

2.8 Remarks

Based on the literature review, the following conclusions are made:

1. Track geometry degradation models were classified in five groups (Figure 2.5) where regression models is the most common approach. Table 2.2 presents a summary of selected references on track geometry degradation where the type of degradation model and model variables are highlighted. The literature review findings show that there is still no consensus on the best representation of track geometry degradation phenomenon, where different assumptions have been made for each model. This makes track geometry degradation an area with potential for improvement.
2. Model parameters' were mainly estimated using frequentits methods such as maximum likelihood estimation and least squares method. As will be discussed in Chapter 3, these methods do not capture the uncertainty associated with model parameters, and probabilistic approaches (such as Markov chain Monte Carlo, implemented in few studies) require further study.
3. In terms of track geometry parameters utilized in degradation models, track quality indices (TQIs), in which a single or multiple parameters are included, have been used to represent a single track section. However, the literature review shows that parameter surface (profile) is the dominant parameter and the most used for representing track geometry degradation.
4. Stochastic processes have been used to model track geometry degradation. These approaches exhibit high potential because it accounts for the uncertainty in the degradation process itself, making it more robust compared to other models. However, despite the important contributions in the literature, the advantages of stochastic processes have not been fully exploited. Failure time is a critical aspect of track geometry degradation for maintenance and safety purposes.

REFERENCES

- M Andersson. Strategic planning of track maintenance. State of the art. Technical report, 2002.
- A Andrade and P Teixeira. Uncertainty in Rail-Track Geometry Degradation : Lisbon-Oporto Line Case Study. *Journal of Transportation Engineering. ASCE.*, 687 (March):193–200, 2011. doi: 10.1061/(ASCE)TE.1943-5436.0000206.
- A. Andrade and P. F. Teixeira. Hierarchical Bayesian modelling of rail track geometry degradation. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(4):364–375, may 2013. ISSN 0954-4097. doi: 10.1177/0954409713486619.
- Antonio Andrade and Paulo Teixeira. A Bayesian model to assess rail track geometry degradation through its life-cycle. *Research in Transportation Economics*, 36(1):1–8, sep 2012. ISSN 07398859. doi: 10.1016/j.retrec.2012.03.011.
- N. Attoh-Okine. *Big data and differential privacy: analysis strategies for railway track engineering*. John Wiley & Sons, 2017. ISBN 1119229057, 9781119229056.
- M. Audley and J. Andrews. The effects of tamping on railway track geometry degradation. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(4):376–391, mar 2013. ISSN 0954-4097. doi: 10.1177/0954409713480439.
- Abdul Rohim Boy Berawi, Raimundo Delgado, Rui Calcada, and Cecilia Vale. Evaluating Track Geometrical Quality Through Different Methodologies. *International Journal of Technology*, 1:38–47, 2010.
- H. Chang, R. Liu, and Q. Li. A multi-stage linear prediction model for the irregularity of the longitudinal level over unit railway sections, 2010. ISSN 17433509.
- T Dahlberg. Some Railroad Settlement ModelsA Critical Review. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 215 (4):289–300, jan 2001. ISSN 0954-4097. doi: 10.1243/0954409011531585.
- E Fortunato, A Pinelo, J Lobo Costa, D Goncalves, and A Pratas. Some results of performance indicators of a railway track obtained during its renewal process. In *4th International SIIV Congress*, Palermo, 2007.

- Hakan Guler. Prediction of railway track geometry deterioration using artificial neural networks: a case study for Turkish state railways. *Structure and Infrastructure Engineering*, 10(5):1–13, 2013a. ISSN 1573-2479. doi: 10.1080/15732479.2012.757791.
- Hakan Guler. Decision Support System for Railway Track Maintenance and Renewal Management. *Journal of Computing in Civil Engineering*, 27(3):292–306, may 2013b. ISSN 0887-3801. doi: 10.1061/(ASCE)CP.1943-5487.0000221.
- Qing He, Hongfei Li, Debarun Bhattacharjya, Dhairat P. Parikh, and Arun Hampapur. Railway Track Geometry Defect Modeling: Deterioration, Derailment Risk, and Optimal Repair. *Transportation Research Board 92st Annual Meeting*, pages 1–20, 2013.
- B Indraratna, W Salim, and C Rujikiatkamjorn. *Advanced rail geotechnology: ballasted track*. CRC press London, 2011. ISBN 9780415669573.
- R. N. Iyengar and O. R. Jaiswal. Random Field Modeling of Railway Track Irregularities. *Journal of Transportation Engineering*, 121(4):303–308, may 1995. ISSN 0733-947X. doi: 10.1061/(ASCE)0733-947X(1997)123:3(245.2).
- Dingqing Li, James P. Hyslip, Theodore R. Sussmann, and Steven Chrismer. *Railway geotechnics*. CRC Press, 2015. ISBN 9780415695015.
- Reng-Kui Liu, Peng Xu, Zhuang-Zhi Sun, Ce Zou, Quan-Xin Sun, Reng-Kui Liu, Peng Xu, Zhuang-Zhi Sun, Ce Zou, and Quan-Xin Sun. Establishment of Track Quality Index Standard Recommendations for Beijing Metro. *Discrete Dynamics in Nature and Society*, 2015:1–9, 2015. ISSN 1026-0226. doi: 10.1155/2015/473830.
- Janusz Madejski and Juliusz Grabczyk. Continuous geometry measurement for diagnostics of tracks and switches. In Delft University of Technology, editor, *Proceedings of the International Conference on Switches*. Delft University of Technology, 2002.
- C. Meier-Hirmer, A. Senée, G. Riboulet, F. Sourget, and M. Roussignol. A decision support system for track maintenance. In *Computers in Railways X*, volume 1 of *WIT Transactions on The Built Environment, Vol 88*, pages 217–226, Southampton, UK, jun 2006. WIT Press. ISBN 1845641779. doi: 10.2495/CR060221.
- Sophie Mercier, Carolina Meier-Hirmer, and Michel Roussignol. Bivariate Gamma wear processes for track geometry modelling, with application to intervention scheduling. *Structure and Infrastructure Engineering*, 8(4):357–366, apr 2012. ISSN 1573-2479. doi: 10.1080/15732479.2011.563090.
- J. S. Mundrey. *Railway track engineering*. Tata McGraw-Hill Pub, 2000. ISBN 007463724X.

- Johan Oberg. *Track Deterioration of Ballasted Tracks: Marginal Cost Models for Different Railway Vehicles*. Rail Vehicles, Aeronautical and Vehicle engineering, Royal Institute of Technology, 2006. ISBN 917178537X.
- Tatsuo Oyama and Masashi Miwa. Mathematical modeling analyses for obtaining an optimal railway track maintenance schedule. *Japan Journal of Industrial and Applied Mathematics*, 23(2):207–224, jun 2006. ISSN 0916-7005. doi: 10.1007/BF03167551.
- L. M. Quiroga and E. Schnieder. Monte Carlo simulation of railway track geometry deterioration and restoration. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 226(3):274–282, 2012. ISSN 1748-006X. doi: 10.1177/1748006X11418422.
- Javad Sadeghi and Hossein Askarinejad. Development of Improved Railway Track Degradation Models. *Structure and Infrastructure Engineering*, 6(6):675–688, 2010. ISSN 1573-2479. doi: 10.1080/15732470801902436.
- Y Shafahi and R Hakhamaneshi. Application of a Maintenance Management Model for Iranian Railways Based on the Markov Chain and Probabilistic Dynamic Programming. *Transaction A: Civil Engineering*, 16(1):87–97, 2009.
- I. Soleimanmeigouni, A. Ahmadi, and U. Kumar. Track geometry degradation and maintenance modelling: A review. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 0(0):1–30, jul 2016a. ISSN 0954-4097. doi: 10.1177/0954409716657849.
- Iman Soleimanmeigouni, Alireza Ahmadi, Christophe Letot, Arne Nilsson, and Uday Kumar. Cost-based optimization of track geometry inspection. In *11th World Congress on Railway Research*, pages 1–7, 2016b.
- U.S Army. Unified facilities criteria: Railroad track maintenance & safety standards, 2008.
- Cecília Vale and Simões Lurdes. Stochastic model for the geometrical rail track degradation process in the Portuguese railway Northern Line. *Reliability Engineering and System Safety*, 116:91–98, 2013. ISSN 09518320. doi: 10.1016/j.res.2013.02.010.
- P Xu, Q Sun, R Liu, and F Wang. A Short-Range Prediction Model for Track Quality Index. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 225(3):277–285, 2011. ISSN 0954-4097. doi: 10.1177/2041301710392477.
- Mohammad Yousefikia, Sara Moridpour, Sujeeva Setunge, and Ehsan Mazloumi. Modeling Degradation of Tracks for Maintenance Planning on a Tram Line. *Journal of Traffic and Logistics Engineering*, 2(2):86–91, 2014. ISSN 23013680. doi: 10.12720/jtle.2.2.86-91.

- Allan M. Zarembski and G Robert Newman. Comparative Technical and Economic Analysis of Stoneblowing vs. Tamping. In *AREMA 2008 Annual Conference*, pages 1–26. AREMA, 2008.
- Allan M Zarembski and Joseph W Palese. Managing risk on the railway infrastructure. In *7th World Congress on Railway Research (WCRR)*, pages 1–18, 2006.
- Guoqiang Zhang, B. Eddy Patuwo, and Michael Y. Hu. Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting*, 14(1): 35–62, 1998. ISSN 01692070. doi: 10.1016/S0169-2070(97)00044-7.
- Yu-Jiang Zhang, Magdy El-Sibaie, and Sung Lee. FRA track quality indices and distribution characteristics. In *AREMA 2004 Annual Conference*, 2004.

Chapter 3

BAYESIAN INFERENCE

3.1 Introduction

Bayesian inference has gained high relevance in statistical inference. It allows quantifying of the uncertainty associated with measurements caused by non-calibrated devices or processes that are not completely controlled or understood ([Gamerman and Lopes, 2006](#)). Also, Bayesian inference is a powerful mechanism for incorporating information from previous studies or experts' opinions regarding the model parameters ([Dunson, 2001](#)). In particular, Markov chain Monte Carlo methods have been utilized in different fields to facilitate the Bayesian analysis, by using numerical methods that provide an approximation of the posterior density of the model parameters. Applications include epidemiology ([Dunson, 2001](#)), water resources ([Lu et al., 2012](#)), computer science ([Zhang, 2001](#)), and civil infrastructure ([Mills et al., 2012](#); [Andrade and Teixeira, 2012](#); [Mokhtarian et al., 2013](#)), among others.

In railway track geometry, large amounts of data collected from different sensors measuring track geometry defects, measurement errors, and modeling uncertainties heavily influence maintenance decisions. Therefore, accurate methods of analysis and interpretation of the data are critical. Few studies have focused on uncertainties associated with track geometry degradation using Markov chain Monte Carlo approaches ([Andrade and Teixeira, 2012, 2013](#)). However, few contributions in terms of output analysis have been performed. This chapter presents the general concepts of Bayesian inference with a main focus on Markov chain Monte Carlo methods, in which their advantages with respect to classical inference methods are discussed. Also, an implementation of track geometry degradation model parameters estimation is presented.

3.2 Frequentist and Bayesian Approaches

Frequentist and Bayesian approaches have played an important role when it is necessary to estimate model parameters. There are fundamental differences in both approaches regarding each approaches' assumptions, as well as the results interpretation. This section broadly discusses these differences and presents the motivation of using Bayesian approaches to estimate model parameters for track geometry degradation.

Frequentist approaches are developed based on the assumption that data come from a repeatable random sample and the model parameters are constant during this process. Since the unknown parameters are assumed to be constant, their uncertainty are expressed not in terms of a probability distribution, but as confidence limits and significant levels (p-values) instead (Cox, 2005). Examples of frequentist approaches include least squares, maximum likelihood estimation, etc.

Bayesian approaches consider observed data that come from a sample (data are fixed) and the model parameters are unknown and follow a probability distribution. In general, Bayesian inference estimates the probability distribution of the model parameters by combining the prior knowledge of the parameters (prior distribution) and the probability of the data given the parameters (likelihood distribution). Both the prior and likelihood are put together, so the probability distributions of the parameters given the observed data are estimated (posterior distribution). Equation 3.2 presents an approximation for representing the estimation of the posterior distribution as a function of the prior and likelihood distributions.

$$Posterior \approx Prior \times Likelihood. \quad (3.1)$$

In the frequentist approach, for example, a *confidence interval* of 95% means that in a collection of intervals, 95% of them contain the true parameter θ (left side of Figure 3.1). In the Bayesian approach, a *credible interval* means that there is a 95% chance that the interval contains the true parameter (right side of Figure 3.1).

Table 3.1 summarizes the main differences between Bayesian and frequentist

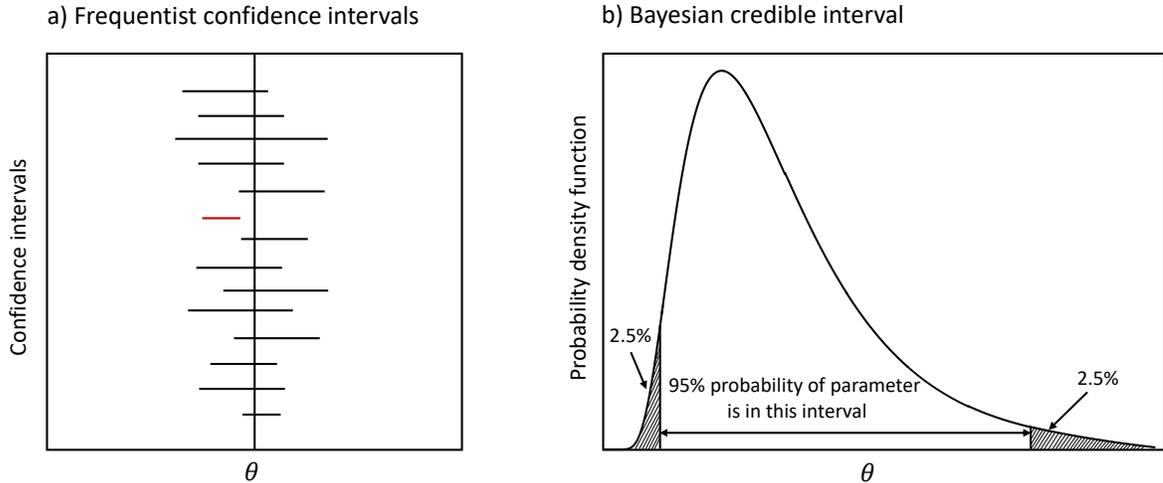


Figure 3.1: Interpretation of confidence and credible intervals

approaches.

Table 3.1: Differences between Bayesian and frequentist approaches

Bayesian	Frequentist
Data are observed from the realized sample the studies are fixed	Data are repeatable random sample there is a frequency the studies are repeatable
Parameters are unknown and assumed to follow a probability distribution	Underlying parameters remain constant during this repeatable process
Data are fixed	Parameters are fixed
Uses probabilities for the parameters and data	Does not use or estimates the probability of a parameter
Depends on the prior and likelihood of observed data	Depends on the likelihood for observed and unobserved data
Uses prior information of the parameters	Does not use prior information of the parameters

3.3 Representation of Bayesian Estimation

As presented in Section 3.2, Bayesian inference assesses the uncertainty of model parameters given observed data and it is a function of the prior and likelihood functions.

Formally it can be defined using the Bayes' rule:

$$f(\theta | X) = \frac{f(\theta)f(X | \theta)}{\int f(\theta)f(X | \theta)d\theta}, \quad (3.2)$$

where

- $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$: parameter vector
- X : observed data
- $f(\theta)f(X | \theta)$: joint distribution
- $\int f(\theta)f(X | \theta)d\theta$: normalizing constant

The joint distribution is composed by two functions: $f(X | \theta)$ is defined as the *likelihood* function. $f(\theta)$ is known as the *prior* function and is defined as the probability of the model parameters before observing the data. Figure 3.2 presents a schematic representation of the Bayesian estimation. This probability usually comes from experts' opinions, previous studies, and conjugacy, etc., and it can be classified in four groups: (i) conjugate priors, (ii) Jeffreys priors, (iii) non-informative priors, and (iv) informative priors. Description of each prior class is presented below.

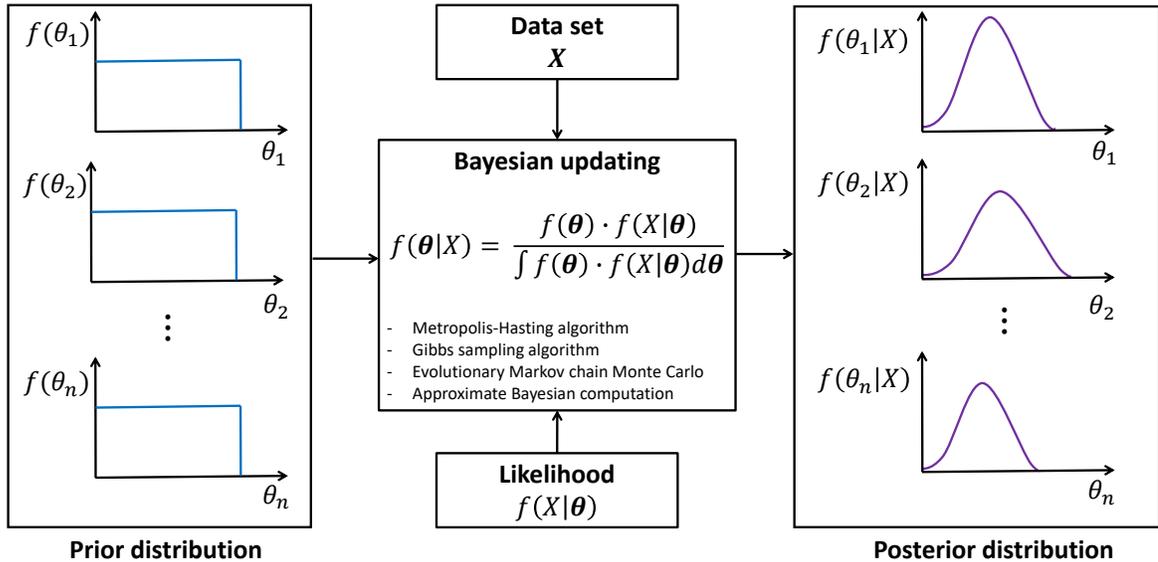


Figure 3.2: Illustration of the Bayesian estimation process

3.3.1 Conjugate Priors

In a large number of practical applications, it is not feasible to find a closed-form solution to the Bayes' theorem. The integral expressed in the denominator of equation 3.2 can be challenging to solve, especially in high-dimensional parameter spaces (Stauffer, 2007). However, in some applications, a closed-form solution can be obtained by using *conjugate priors*. Under this approach, the prior distribution is chosen from specific probability distribution priors that can provide the posterior distribution. Table 3.2 presents a summary of some common cases of conjugate priors.

Table 3.2: Conjugate prior distributions (Attoh-Okine, 2017)

Prior	Likelihood	Posterior
Normal	Normal	Normal
Normal	Gamma	Gamma
Bernoulli	Beta	Beta
Poisson	Gamma	Gamma
Binomial	Beta	Beta

An example of a conjugate prior is presented as follows:

The data set X is normally distributed with parameters $N(\theta, \sigma^2)$, and the prior distribution of the parameters θ is $N(\mu, b^2)$. Then the likelihood function is:

$$f(X | \theta) = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^2 \exp \left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \theta}{\sigma} \right)^2 \right). \quad (3.3)$$

In addition, the prior distribution of θ is:

$$f(\theta) = \frac{1}{b\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{\theta - \mu}{b} \right)^2 \right), \quad (3.4)$$

so the posterior distribution is:

$$f(\theta | X) = \frac{1}{b\sqrt{2\pi}} \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^2 \exp \left(-\frac{1}{2} \left[\sum_{i=1}^n \left(\frac{x_i - \theta}{\sigma} \right)^2 + \left(\frac{\theta - \mu}{b} \right)^2 \right] \right). \quad (3.5)$$

$$f(\theta | X) \propto \exp \left[-\frac{1}{2} \left(\frac{\theta^2(\sigma^2 + nb^2) - 2\theta(\sigma^2\mu + nb^2\bar{x}) + (\sigma^2\mu^2 + b^2 \sum x_i^2)}{\sigma^2 b^2} \right) \right]. \quad (3.6)$$

$$f(\theta | X) \propto \exp \left[-\frac{1}{2} \left(\frac{\theta - \frac{\mu\sigma^2 + nb^2\bar{x}}{\sigma^2 + nb^2}}{\sqrt{\frac{b^2\sigma^2}{\sigma^2 + nb^2}}} \right)^2 \right], \quad (3.7)$$

therefore,

$$f(\theta | X) \sim N \left(\frac{\mu\sigma^2 + nb^2\bar{x}}{\sigma^2 + nb^2}, \frac{b^2\sigma^2}{\sigma^2 + nb^2} \right). \quad (3.8)$$

In this case, it can be noticed that both, the prior and posterior are normally distributed, and therefore, are conjugates.

3.3.2 Jeffreys Priors

An approach that allows estimation of the posterior distribution when conjugate priors are not suitable is the Jeffreys priors. This class of priors are defined as those that might vary when a transformation is done. For example, if parameter θ_i follows an uniform distribution, it is not necessary that θ_i^2 is also uniformly distributed. The idea of the Jeffreys priors is to define posterior distributions that are invariant to any transformation of the model parameters (Stauffer, 2007). This is done by choosing a prior based on the shape of the likelihood function.

3.3.3 Non-informative Priors

Non-informative priors are mainly used when there is not enough knowledge about the model parameters, so a flat distribution is defined in order to reduce the impact over the posterior distribution. A common choice of this class of priors includes the uniform distribution, because it assigns the same probability to each value of the parameter. Although this approach allows the reduction of subjective assessment of the model parameters, problems might arise when the parameter is not bounded, so the sum of all the probability values for parameter θ_i does not sum up to 1. This case

is known as *improper priors*. However, having an improper prior would not produce an improper posterior, if the likelihood function has bounds that support where the function is nonzero (Stauffer, 2007).

3.3.4 Informative Priors

In contrast to non-informative priors, informative priors are not dominated by the likelihood function. This type of priors are usually defined when there is a high knowledge of the model parameters based on previous studies or experts' opinions.

3.4 Markov Chain Monte Carlo Methods

There are alternative methods to Bayesian inference techniques including Monte Carlo integration and rejection sampling, among others. These methods are characterized to produce sets of independent simulated values from the desired probability distribution (Stauffer, 2007). On the other hand, Markov chain Monte Carlo (MCMC) methods provide simulated chains that are slightly dependent on each other. This is because the current value of the chain depend on the previous one. The MCMC idea is to simulate a Markov chain whose stationary distribution is the posterior distribution of the model parameters θ . To reach this stationary distribution, the Markov chain has to be ergodic and irreducible. When the posterior distribution cannot be estimated analytically, numerical methods-based MCMC are used to provide an approximation of the posterior distribution. In this section, the Metropolis-Hastings algorithm, Gibbs sampling algorithm, and population-based MCMC methods are discussed.

3.4.1 Markov Chains

Markov chains is a class of stochastic processes and it has been applied in different areas. Formally, a Markov chain is defined as follows:

Definition 3.4.1. *A stochastic process $X = (X_j)_{j \in \mathbb{N}_0}$ with values in a set S is a Markov chain if (Voss, 2013)*

$$P(X_j \in A \mid X_{j-1} \in A_{j-1}, X_{j-1} \in A_{j-1}, \dots, X_0 \in A_0) = \quad (3.9)$$

$$P(X_j \in A_j \mid X_{j-1} \in A_{j-1}), \forall A_0, A_1, \dots, A_j \subseteq S, \text{ and } j \in \mathbb{N}. \quad (3.10)$$

Equation 3.9 states that the condition of the state at time j only depends on the condition at time $j - 1$.

The probability of moving from state i to state j is represented by the transition probability matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{pmatrix}. \quad (3.11)$$

Markov chain states (e_j) have the following properties (Neapolitan, 2009):

- e_j is *reachable* from e_i in $n \geq$ steps such as $p_{ij}^n > 0$. A Markov chain is called *irreducible* if every state is reachable from every other state.
- e_j is *periodic* if it has a period $t > 1$ with $p_{ij}^{(n)} = 0$, unless $n = mt$ for any integer m , and t is the largest integer with this property. In its counterpart, e_j is *aperiodic* if $t > 1$ does not exist.
- e_j is *persistent* if $f_{ii} = 1$ and is *transient* if $f_{ii} < 1$. Where f_{ii} is the probability of starting from state e_i and entry to state e_i . The persistent state e_j is also called *null* if its mean recurrence time ($\mu_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)} = \infty$) and otherwise it is called *non-null*.
- e_j is *ergodic* if it is aperiodic, persistent, and non-null. A Markov chain is said to be ergodic if all states are ergodic.

The stationary distribution of the Markov chain is presented as follows (Neapolitan, 2009):

Theorem 3.4.1. *For an ergodic and irreducible Markov chain the limits*

$$r_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)} \quad (3.12)$$

exist and are independent of the initial state e_i . Also, for $r_j > 0$

$$\sum_j r_j = 1, \quad (3.13)$$

$$r_j = \sum_i r_i p_{ij}, \quad (3.14)$$

and

$$r_j = \frac{1}{\mu_j}, \quad (3.15)$$

where μ_j is the mean recurrence time of e_j .

Therefore, the probability distribution $P(E = e_j) \equiv r_j$ is called the stationary distribution of the Markov chain.

3.4.2 Metropolis-Hastings Algorithm

The Metropolis-Hastings (MH) algorithm is based in the accept-reject (AR) method. The AR method generates non-Markov samples. The objective is to generate samples from a continuous *target function* $\pi(x) = f(x)/K$, where $x \in \mathbb{R}^d$, $f(x)$ is the unnormalized density, and K is the normalizing constant (Chib and Greenberg, 1995). Also, let the *envelope function* $h(x)$ be a density where samples can be generated from and there is a constant c such that $f(x) \leq ch(x)$, $\forall x$ (Figure 3.3). To obtain the target function, the accept-reject algorithm is presented in Algorithm 1.

Algorithm 1 Accept-reject algorithm

- 1: Propose a candidate H from $h(\cdot)$
 - 2: Generate a value u from an uniform distribution $U(0, 1)$
 - 3: **if** $u \leq f(H)/ch(H)$ **then**
 - 4: Set $H = y$
 - 5: **else**
 - 6: Go to step 1
 - 7: **end if**
-

The accepted value y is a random variable from the target function $\pi(\cdot)$. As presented by [Chib and Greenberg \(1995\)](#), constant c must be carefully selected because the accept-reject routine may result in a large number of rejected samples.

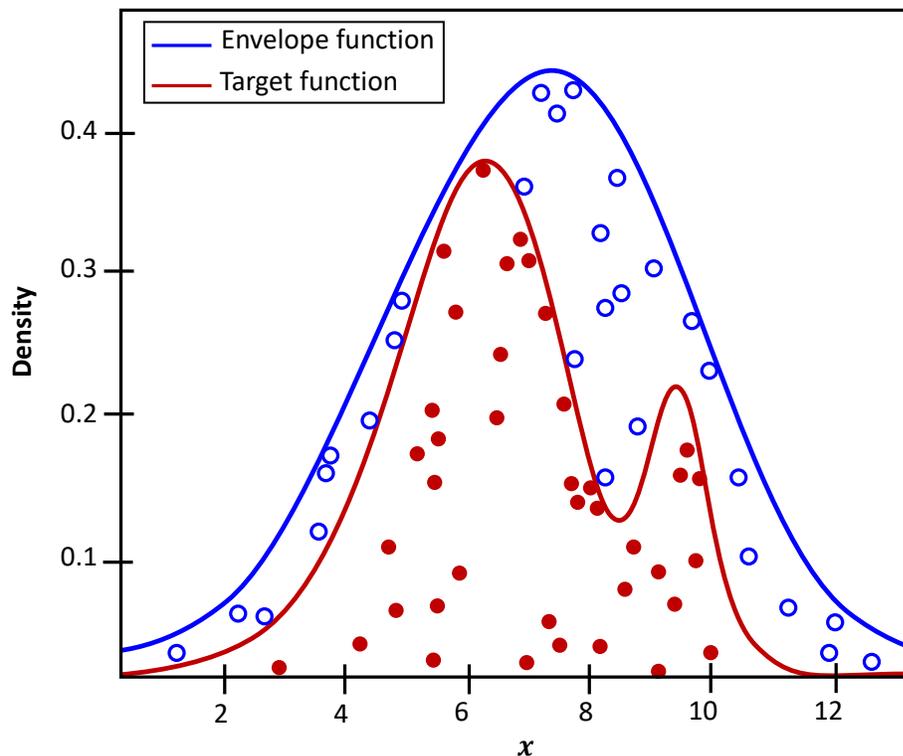


Figure 3.3: Representation of accept-reject regions

The M-H algorithm simulates samples from a probability distribution by using the joint density function and proposal densities for each parameter to be estimated as presented in Algorithm 2 ([Yildirim, 2012](#)).

Algorithm 2 Metropolis-Hastings routine

- 1: Initialize $x^{(0)} \sim q(x)$
- 2: **for** iteration $i = 1, 2, \dots$ **do**
- 3: Draw $x^{cand} \sim q(x^{(i)} | x^{(i-1)})$
- 4: Compute M-H ratio

$$\alpha(x^{(cand)} | x^{(i-1)}) = \min \left\{ 1, \frac{q(x^{(i-1)} | x^{(cand)}) \pi(x^{(cand)})}{q(x^{(cand)} | x^{(i-1)}) \pi(x^{(i-1)})} \right\}$$

- 5: Draw $u \sim U(0, 1)$
 - 6: **if** $U < \alpha$ **then**
 - 7: Accept proposal: $x^{(i)} = x^{cand}$
 - 8: **else**
 - 9: Reject proposal: $x^{(i)} = x^{(i-1)}$
 - 10: **end if**
 - 11: **end for**
-

3.4.3 Gibbs Sampling Algorithm

The Gibbs sampling algorithm is a special case of the M-H algorithm and it has been widely used when sampling from a multivariate posterior is not feasible, but sampling from the conditional distributions for each parameter of interest is feasible (Lynch, 2007). In contrast to the M-H algorithm, the Gibbs sampling has an acceptance rate of 100%, meaning that the M-H ratio is not computed. Instead, each parameter θ_i sample is generated by calculating the conditional probability of a specific parameter given the current value of the remaining parameters and repeating the process until reaching a predetermined number of iterations. Algorithm 3 presents the Gibbs sampling routine (Lynch, 2007).

Algorithm 3 Gibbs sampling algorithm

- 1: Assign a vector of starting values S to the parameter vector θ
 - 2: Set $i = i + 1$
 - 3: Sample $\left(\theta_1^j | \theta_2^{(j-1)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)} \right)$
 - 4: Sample $\left(\theta_2^j | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)} \right)$
 - \vdots
 - 5: Sample $\left(\theta_k^j | \theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_{k-1}^{(j)} \right)$
-

3.4.4 Population-Based Markov Chain Monte Carlo Methods

Another variation of MCMC methods include hybrid approaches, in which meta-heuristic approaches are combined with the MH and Gibbs sampling algorithm. These methods are called *population MCMC*. This subsection presents a variation of these methods, where the evolutionary Markov chain Monte Carlo (EMCMC) is presented.

The EMCMC samples a population of Markov chains in parallel, where each individual within the population is coded as a finite real-coded sequence. The population evolves by applying the genetic operators crossover and mutation (Liang and Wong, 2001).

Formally, the EMCMC in its real-coded version is defined as follows: a distribution on a space of finite real sequence is represented by Liang and Wong (2000) as shown below.

$$f(x) = \exp\{-H(x)/\tau\}, \quad (3.16)$$

where:

- x : d -dimensional vector $x = (\beta_1, \beta_2, \dots, \beta_d)$, $\beta_i \in \mathbb{R}$
- $H(x)$: fitness function. It is defined as the negative of the log-density of x

3.4.4.1 Mutation

A chromosome x_i is uniformly chosen from the current population $\mathbf{x} = \{x_1, \dots, x_N\}$, then mutated to a new chromosome y_k by reversing the values of some bits that are also chosen randomly (Figure 3.4). A new population is proposed as $\mathbf{y} = \{x_1, \dots, y_k, \dots, x_N\}$, and it is accepted or rejected according to the Metropolis-Hastings rule. A new chromosome is generated by adding a random vector e_k (Liang and Wong, 2000):

$$y_k = x_k + e_k, \quad (3.17)$$

where e_k is usually chosen for the mutation operation to have moderate acceptance probability (Liang and Wong, 2001). The new population \mathbf{y} is accepted with probability $\min(1, r_m)$ according to the Metropolis-Hastings rule:

$$r_m = \frac{f(\mathbf{y}) T(\mathbf{x} | \mathbf{y})}{f(\mathbf{x}) T(\mathbf{y} | \mathbf{x})} = \exp \left\{ (H(\mathbf{y}_k)) - \frac{H(\mathbf{x}_k)}{t_k} \right\} \frac{T(\mathbf{x} | \mathbf{y})}{T(\mathbf{y} | \mathbf{x})}, \quad (3.18)$$

where $T(\cdot | \cdot)$ denotes the transition probability between populations.

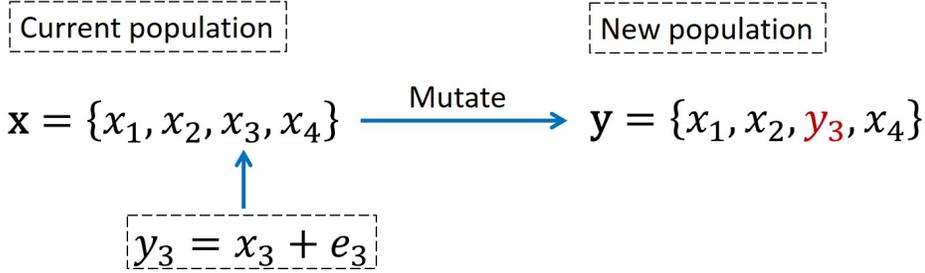


Figure 3.4: Representation of mutation operator

3.4.4.2 Crossover

Two chromosomes x_i and x_j are chosen from the current population \mathbf{x} according to some selection procedure, such as random selection, tournament, or roulette wheel selection (Liang and Wong, 2000). The new population \mathbf{y} is accepted with probability $\min(1, r_c)$ according to the Metropolis-Hastings rule:

$$r_c = \frac{f(\mathbf{y}) T(\mathbf{x} | \mathbf{y})}{f(\mathbf{x}) T(\mathbf{y} | \mathbf{x})} = \exp \left\{ -\frac{H(y_i) - H(x_i)}{t_i} - \frac{H(y_j) - H(x_j)}{t_j} \right\} \frac{T(\mathbf{x} | \mathbf{y})}{T(\mathbf{y} | \mathbf{x})}, \quad (3.19)$$

where $T(\cdot | \cdot)$ denotes the selection probability of (x_i, x_j) from the population \mathbf{x} , and $P(\cdot | \cdot)$ denotes the probability of (y_i, y_j) from the parents chromosomes (x_i, x_j) . Figure 3.5 illustrates the crossover operator.

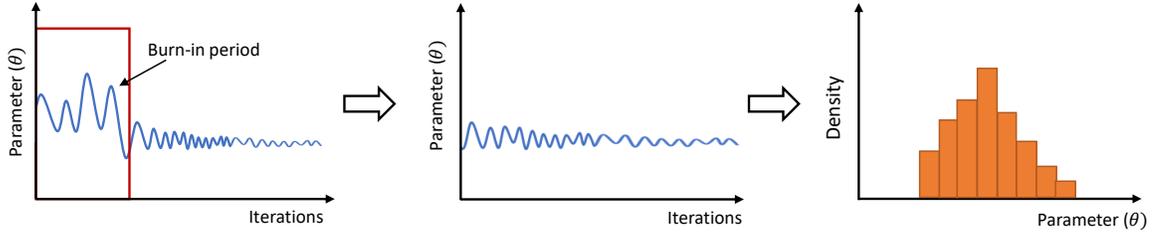


Figure 3.6: Representation burn-in period

In track geometry degradation, as discussed in chapter 2, most of the studies utilized frequentist approaches to estimate model parameters. [Andrade and Teixeira \(2012, 2013\)](#) implemented MCMC for the estimation of the linear track geometry degradation model parameters. The objective was to evaluate the uncertainty behavior of the track infrastructure through its life-cycle. Despite the contribution in this area, these studies did not discuss the MCMC output analysis, which is important to define the convergence of the simulated Markov chain in the steady state. Elements such as good or poor mixing in the parameters simulated values over the iterations, Kernel densities, and autocorrelation plots, are necessary to evaluate the MCMC routines.

To illustrate the MCMC output analysis, a linear representation of track geometry degradation using U.S. class 1 track geometry inspection data is presented as shown in equation [3.20](#).

$$\sigma = \theta_1 + \theta_2 t, \quad (3.20)$$

where:

- σ : standard deviation of surface (in)
- θ_1 : intercept (in)
- θ_2 : degradation rate (in/month)
- t : inspection time (month)

The parameters vector include the intercept and the degradation rate $\theta = [\theta_1, \theta_2]$.

In this example, the burn-in period was defined as 5000 iterations after a trial and error procedure. Figures 3.7 and 3.8 show the trace and density for parameters θ_1 and θ_2 respectively. With regard to the prior distribution, non-informative priors were considered. The motivation of using non-informative priors was to reduce subjective assessments that can impact the posterior distribution. As presented in this chapter, non-informative priors may lead to improper distributions. The prior distribution for both parameters follows a lognormal distribution with parameters $\theta_1 \sim LN(0.1, 0.001)$ and $\theta_2 \sim LN(0.1, 0.001)$ were defined.

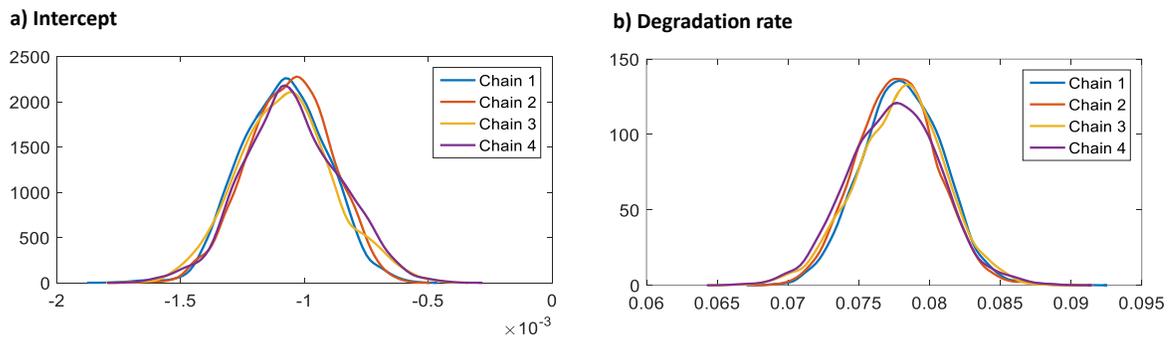


Figure 3.7: Kernel plot

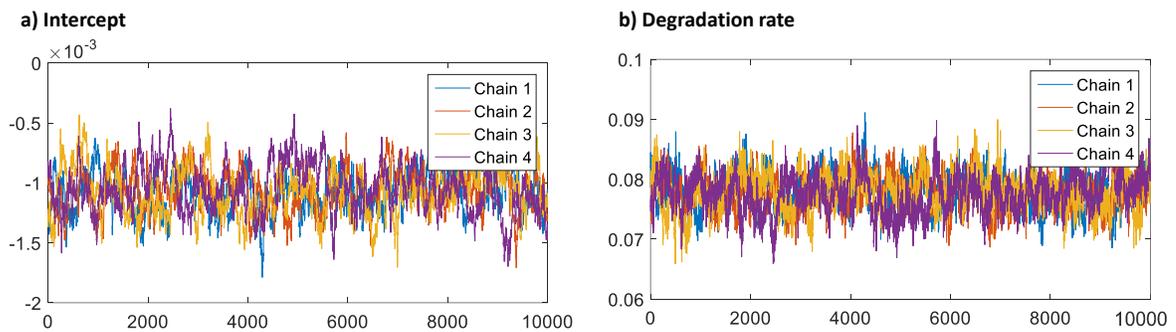


Figure 3.8: Trace plot

The MCMC convergence is another point to be analyzed in the MCMC output. Time series plays an important role in this case, because it represents the algorithm convergence through iterations. One of the drawbacks is that the time series are autocorrelated, so the informative properties of the output are not the best. One way to

increase the information of the time series is to eliminate the autocorrelation, by making each iteration independent from the posterior distribution. The strategy consists of lengthening each MCMC run by a factor m and taking every m -th iteration. This process is known as *thinning*. In this example, the total number of iterations and the thinning values were set as 50,000 and 10 iterations, respectively. As seen in Figure 3.9, the autocorrelation values decrease as the lag increases for all the chains, ensuring independence of the initial values.

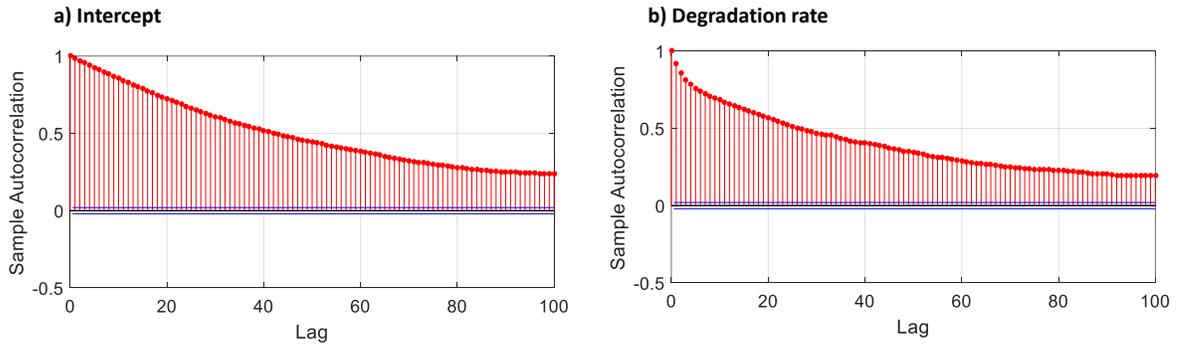


Figure 3.9: Autocorrelation plot

3.6 Remarks

This chapter presented the definitions of Bayesian inference and how it differs from frequentist approaches for parameter estimation. This chapter discussed three different variations of Markov chain Monte Carlo (MCMC) techniques named Metropolis-Hastings algorithm, Gibbs sampling algorithm, and evolutionary MCMC.

Bayesian analysis seems to have potential for handling complex degradation models. As presented in this chapter, few contributions in the literature regarding parameters estimation for track geometry degradation models have been made, so it is still an open field that deserves more attention for capturing uncertainty of the model parameters.

REFERENCES

- A. Andrade and P. F. Teixeira. Hierarchical Bayesian modelling of rail track geometry degradation. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(4):364–375, may 2013. ISSN 0954-4097. doi: 10.1177/0954409713486619.
- Antonio Andrade and Paulo Teixeira. A Bayesian model to assess rail track geometry degradation through its life-cycle. *Research in Transportation Economics*, 36(1):1–8, sep 2012. ISSN 07398859. doi: 10.1016/j.retrec.2012.03.011.
- N. Attoh-Okine. *Big data and differential privacy: analysis strategies for railway track engineering*. John Wiley & Sons, 2017. ISBN 1119229057, 9781119229056.
- Siddhartha Chib and Edward Greenberg. Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4):327–335, nov 1995. doi: 10.1080/00031305.1995.10476177.
- D.R. Cox. Frequentist and Bayesian statistics: a critique. In *Proceedings of the Statistical Problems in Particle Physics, Astrophysics and Cosmology*, pages 3–6, 2005.
- D. B. Dunson. Commentary: practical advantages of Bayesian analysis of epidemiologic data. *American Journal of Epidemiology*, 153(12):1222–1226, jun 2001. doi: 10.1093/aje/153.12.1222.
- Dani Gamerman and Hedibert Freitas Lopes. *Markov Chain Monte Carlo-Stochastic Simulation for Bayesian Inference*. 2006. ISBN 978-1-58488-587-0. doi: 10.1002/1521-3773(20010316)40:6(9823::AID-ANIE9823)3.3.CO;2-C.
- Faming Liang and WH Wong. Evolutionary Monte Carlo: Applications to Cp Model Sampling and Change Point Problem. *Statistica sinica*, 10:317–342, 2000. ISSN 10170405. doi: 10.1.1.15.4052.
- Faming Liang and Wing Hung Wong. Real-Parameter Evolutionary Monte Carlo with Applications to Bayesian Mixture Models. *Journal of the American Statistical Association*, 96(454):653–666, 2001.
- Dan Lu, Ming Ye, and Mary C. Hill. Analysis of regression confidence intervals and Bayesian credible intervals for uncertainty quantification. *Water Resources Research*, 48(9):n/a–n/a, sep 2012. doi: 10.1029/2011WR011289.

- Scott M. Lynch. Modern model estimation part 1: gibbs sampling. In *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*, pages 77–105. Springer New York, New York, NY, 2007. doi: 10.1007/978-0-387-71265-9_4.
- Leslie N. O. Mills, Nii O. Attoh-Okine, and Sue McNeil. Hierarchical Markov chain Monte Carlo simulation for modeling transverse cracks in highway pavements. *Journal of Transportation Engineering*, 138(6):700–705, jun 2012. doi: 10.1061/(ASCE)TE.1943-5436.0000383.
- Payam Mokhtarian, Tin Kin Ho, and Thomas Suesse. Bayesian Nonparametric Reliability Analysis for a Railway System at Component Level. In *2013 IEEE International Conference on Intelligent Rail Transportation Proceedings*, pages 197–202, 2013. ISBN 9781467352772.
- Richard E. Neapolitan. *Probabilistic methods for bioinformatics: with an introduction to Bayesian networks*. Morgan Kaufmann/Elsevier, 2009. ISBN 9780080919362.
- Howard B. Stauffer. *Contemporary Bayesian and frequentist statistical research methods for natural resource scientists*. John Wiley & Sons, Inc., Hoboken, NJ, USA, nov 2007. ISBN 9780470185094. doi: 10.1002/9780470185094.
- Jochen Voss. *An Introduction to Statistical Computing*. John Wiley & Sons Ltd, Chichester, UK, oct 2013. ISBN 9781118728048. doi: 10.1002/9781118728048.
- Ilker Yildirim. Bayesian inference: Metropolis-Hastings sampling. 2012.
- B Zhang. System Identification Using Evolutionary Markov Chain Monte Carlo. *Journal of Systems Architecture*, 47(7):587–599, 2001. ISSN 13837621.

Chapter 4

WIENER PROCESS FOR DEGRADATION ANALYSIS

4.1 Introduction

Degradation processes have played an important role in different disciplines including biology, finance, and engineering, among others. Over the past decades, researchers have been interested in analyzing the evolution of degradation data to determine the system's failure rate, and from there, establish suitable maintenance policies (Abdel-Hameed, 2010). As presented in Chapter 2, there are important contributions in the literature in modeling track geometry degradation, in which stochastic processes seek to account for the uncertainty of the degradation process. These models include random field (Iyengar and Jaiswal, 1995), gamma process (Meier-Hirmer et al., 2006), Dagum process (Vale and Lurdes, 2013), Markov chains (Yousefikia et al., 2014), and Wiener process (Soleimanmeigouni et al., 2016). However, estimation of the model parameters using Bayesian approaches and the time at which the degradation process reaches a threshold for the first time have not been addressed.

This chapter introduces the concepts of stochastic processes, where a special type of this models that is suitable for non-strictly monotonic increasing functions, named the Wiener process. Also, the formulation of track geometry degradation as a Wiener process and its first hitting time are defined.

4.2 Degradation in Reliability Analysis

Reliability is the measure of a system's success in providing its function properly during its designed life (Elsayed, 2013). Meeker and Escobar (1998) classified reliability data in three categories: failure-time data with non-explanatory variables, failure-time data with explanatory variables, and degradation data. Although time-failure data

have been also utilized to the development of reliability models, it is shown that relying on this type of data does not provide a big picture of the whole degradation process over time (Sotiris, 2011). Therefore, the use of degradation data has gained more attention because it provides more reliable and richer information (Meeker and Escobar, 1998; Sotiris, 2011).

Degradation can be defined as the loss of functionality over time of a device or system. Also, it is said that the system fails when the degradation reaches a threshold level (Figure 4.1). The relationship between system's failure and degradation data plays an important role in reliability analysis because it helps the development of degradation models that can be used to make predictions about failure time (Meeker and Escobar, 1998). In order to perform a degradation analysis, it is necessary to establish the failure threshold, which is theoretically defined as the minimum allowable degradation value before the system fails. Also, collect degradation data at different times, which can be expressed in terms of the measure of a physical degradation in time or as a measure of the system's performance over time (Meeker and Escobar, 1998). Finally, determine the failure time, which is defined as the moment at which the degradation data exceeds the failure threshold.

There exist in literature a wide range of references that have focused on modeling degradation. These models can be classified in three groups: general path models, stochastic process models, and other models. General path model was first introduced by Lu and Meeker (1993). The idea is to incorporate mixed-effect models, considering the degradation as deterministic. On the other hand, there exist the stochastic process models which is a more realistic representation of the degradation process, because provide a more flexible environment to model degradation and failure time. This is done by incorporating the inherent uncertainty from the degradation process over time. Examples of stochastic processes include Wiener process (Kahle and Lehmann, 2010), Gamma process (Meier-Hirmer et al., 2009), Markov chains (Shafahi and Hakhamaneshi, 2009), and inverse Gaussian process (Pettit and Young, 1999), among others.

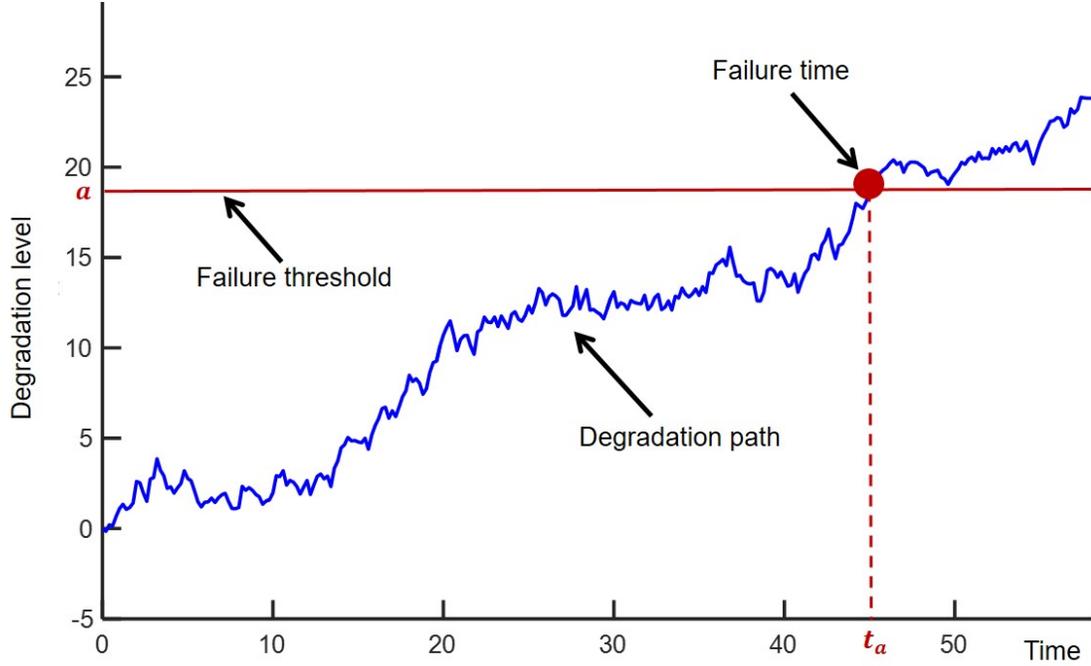


Figure 4.1: Representation of a degradation process with failure threshold

4.3 Review on Stochastic Processes

Consider a system that changes from state i to state j in time. Also, assume that the system changes over time is non deterministic, but it is influenced by a random phenomena instead. If X_t is referred as the system's state in time t and is expressed in terms of a random variable, then the collection of random variables is called a *stochastic process*.

4.3.1 Random Variable

A random variable (RV) is defined as follows (Mikosch, 2000; Ibe, 2013):

Definition 4.3.1. A random variable X is a real-valued function defined over a sample space S . The distribution function of X is

$$P\{X \leq x\} = F_X(x), \quad -\infty < x < \infty. \quad (4.1)$$

This function has the following properties:

1. The probability outcomes are real values $S \in \mathbb{R}$.

2. $F_X(x)$ is a nondecreasing function. That is

$$F_X(x_1) \leq F_X(x_2), \quad \text{for } x_1 < x_2. \quad (4.2)$$

3. $0 \leq F_X(x) \leq 1$.

4. $F_X(\infty) = 1$, $F_X(-\infty) = 0$

The collection of all possible outcomes is called sample space and it can be classified into discrete or continuous.

4.3.1.1 Discrete Random Variables

A discrete random variable X is defined as a random variable which has sample space (Figure 4.2) that is defined in a finite or countable infinite number of values. The probability mass function (PMF) of X is

$$P_X(x) = P[X = x], \quad (4.3)$$

with

$$p_X(x) = \begin{cases} p_X(x_i) \geq 0 & i = 1, \dots, n \\ p_X(x) = 0 & \text{otherwise.} \end{cases} \quad (4.4)$$

Additionally, the cumulative density function (CDF) $F_X(x)$ can be expressed as follows:

$$F_X(x) = \sum_{k \leq x} p_X(k). \quad (4.5)$$

4.3.1.2 Continuous Random Variables

A continuous random variable is a random variable defined for all real values $X \in (-\infty, \infty)$, where the probability density function (PDF) of X for any set of real

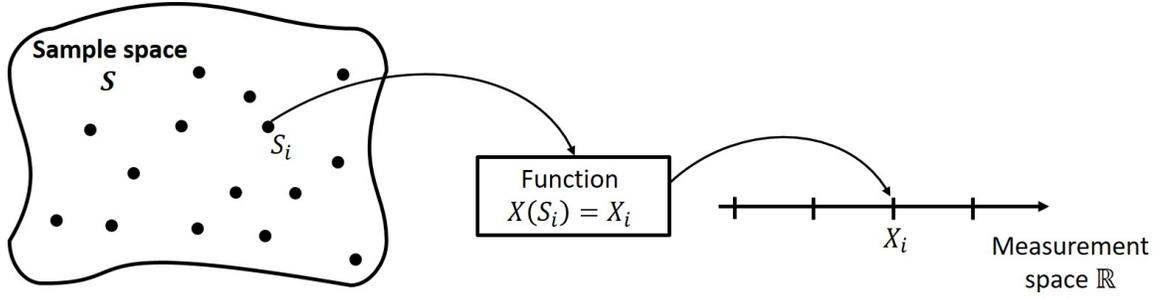


Figure 4.2: Illustration of a random variable defined as a function in the sample space S and takes values on \mathbb{R}

numbers A is

$$P[X \in A] = \int_A f_X(x) dx. \quad (4.6)$$

In railway track geometry degradation, there is uncertainty in terms of the geometry parameters and the track geometry parameters model. For example, the parameter surface right (62 ft) is a random variable, because the actual value at a specific location is only observed until inspection is performed. The ideal value for this parameter is zero; however, the measurement values can vary based on the condition of the track at that specific location or accuracy of the inspection device. Having said that, the sample space of surface right (62 ft) at a location i , can be the set of real values in which the mean is zero or near zero. Since this track geometry parameter can take any value in a defined interval, it is said that it is a continuous random variable.

4.3.2 Stochastic Process

A stochastic process X is a family of random variables X_t , where t is often represented as time, such that for any finite collection of values of $t \in T$, there is the joint probability distribution associated with the random variables $X_{t_1}, X_{t_2}, \dots, X_{t_n}$, defined in some space Ω (Mikosch, 2000). For the continuous case, time T can be expressed as an interval, for example $T = [a, b], [a, b)$ or $[a, \infty)$ for $a < b$. Therefore, X is a continuous-time process. In contrast, in a discrete-time processes, T is a finite or

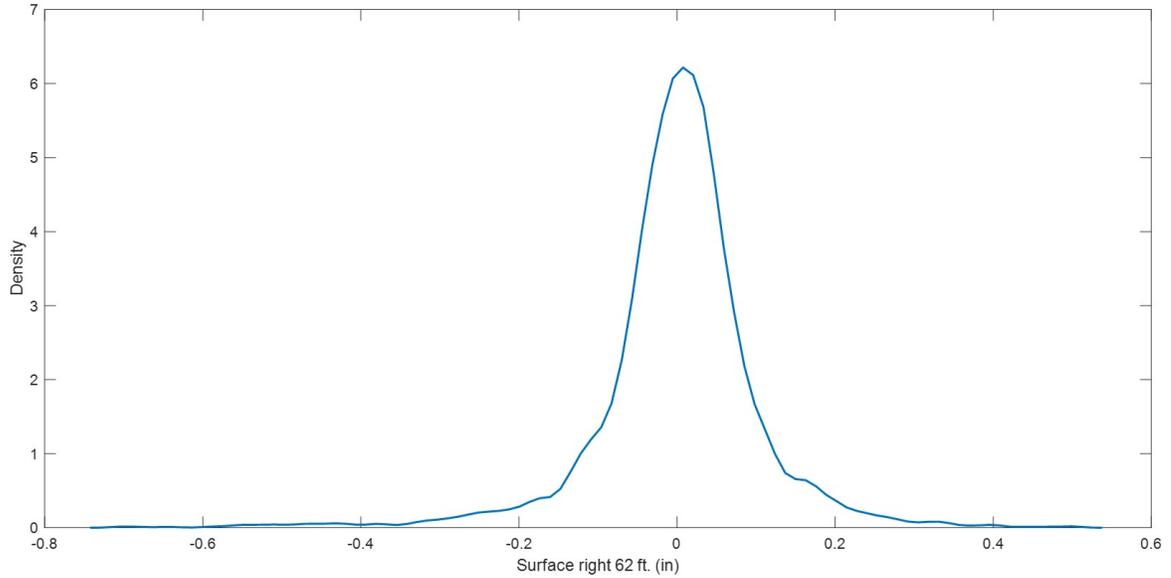


Figure 4.3: Kernel density for surface right 62 ft

countable infinite set.

A stochastic process X is a function of two variables (Mikosch, 2000):

1. For a fixed instant of time, t , it is a random variable:

$$X_t = X_t(w), w \in \Omega. \quad (4.7)$$

2. For a fixed random outcome $w \in \Omega$, it is a function of time:

$$X_t = X_t(w), t \in T. \quad (4.8)$$

Equation 4.8 is called a *sample path* of the process X .

4.4 Stochastic Differential Equations

4.4.1 Motivation

Ordinary differential equations (ODE) have played an important role in mathematics especially because the fundamental principles that govern physical phenomena have been developed using this approach. The resulting phenomena models include the heat equation and Black Scholes equation, among others (Calin, 2015). Equation

4.9 describes a one-dimensional dynamical system and it is called an ODE (Ditlevsen and Samson, 2013).

$$\frac{dx}{dt} = a(x, t). \quad (4.9)$$

If $a(\cdot)$ satisfies the condition that a unique solution exists, then $x(t) = x(t; x_0, t_0)$ is a system's solution with the initial condition $x(t_0) = x_0$. $x(t)$ can be interpreted as the location of a one-dimensional particle in space at time t and $\frac{dx}{dt}$ represents the change in location of the particle in a time interval $t, t + dt$. The trajectory of a ODE given the initial condition is smooth as presented in Figure 4.4.

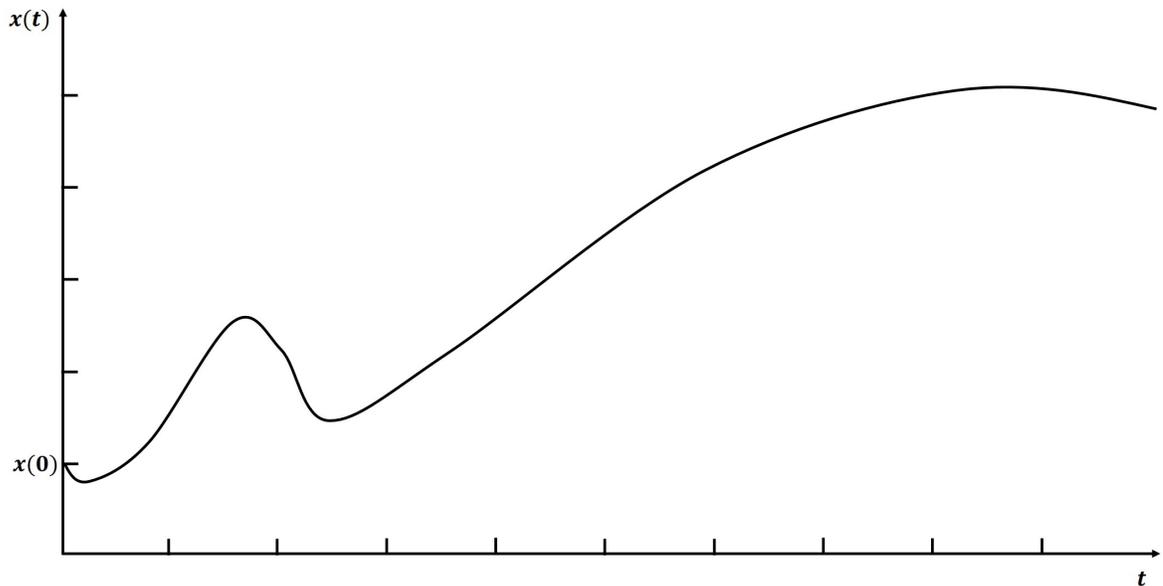


Figure 4.4: Trajectory of a ODE initial condition $x(0)$

There are some comments regarding ODEs (Mikosch, 2000; Ditlevsen and Samson, 2013):

- Solutions of ODEs are presented as functions. Each function represent the dynamics of a physical process over a period of time.
- ODE's unique solution is obtained if the initial condition $x(t_0) = x_0$. That is, once the system's condition at $t = 0$ is available, the function x_t is determined in the future for $t > 0$.

- Although closed-form solutions can be used, it is more common to rely on numerical solutions to ODE.

The smooth path presented in Figure 4.4 can be perturbed by the influence of random effects (also referred as white noise) (Figure 4.5). Equation 4.10 presents the random effects.

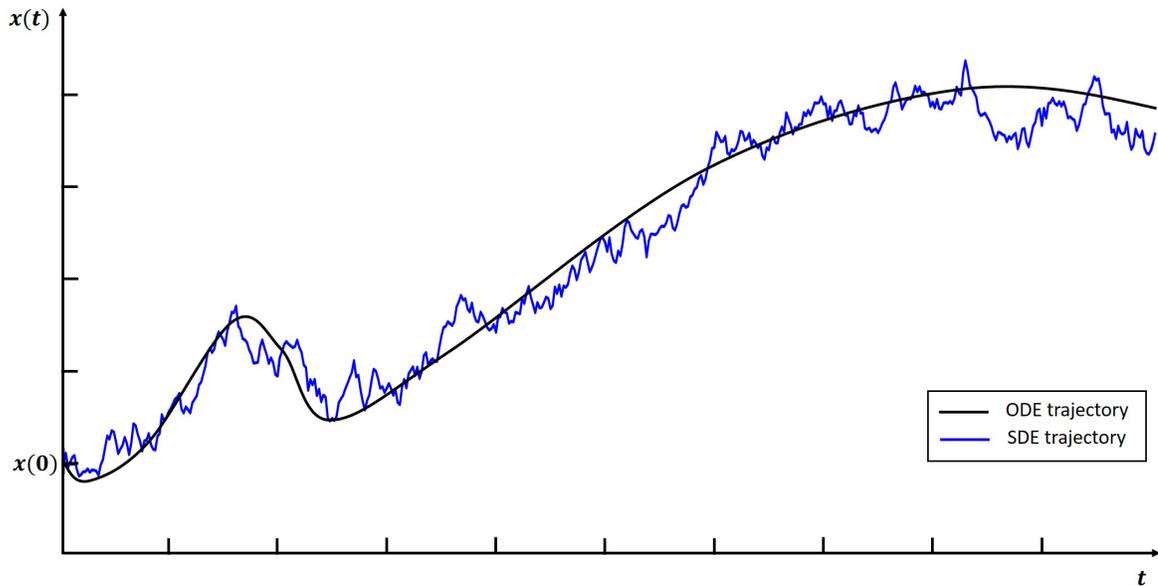


Figure 4.5: Schematic representation of a trajectory of a ODE and a SDE given initial condition $x(0)$

$$dX_t = m(t, X_t)dt + \sigma(t, X_t)dB_t, \quad (4.10)$$

where B_t is the standard Brownian motion.

The SDE differs from ODE in the sense that SDE includes stochastic process in one or more differential equations. The SDE has been applied widely in many areas including biology, medicine, engineering, finance, etc. (Kloeden and Platen, 1992).

4.4.2 Itô Integral

The Itô integral can be defined in a similar way to Riemann integral in differential calculus. While the Riemann integral considers the summation of deterministic

infinitesimal changes in time dt , the Itô integral is taken with respect to infinitesimal increments of the Brownian motion $B(t)$ which are random variables (Calin, 2015). The output of a Riemann integral, then, is a real number, whereas the output of a Itô integral is a random variable.

Despite the main differences between the deterministic and stochastic integrals, they have some common properties (Calin, 2015):

Consider $0 \leq a < b$ and let $F_t = f(W_t, t)$ be a nonanticipating process satisfying the “non-explosive” condition

$$\mathbb{E} \left[\int_a^b F_t^2 dt \right] < \infty. \quad (4.11)$$

Theorem 4.4.1. *Let f be a C^2 function, that is, a function with 2 continuous derivatives, and B_t be the standard Brownian motion. Then, for every t*

$$f(B_t) = f(B_0) + \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds. \quad (4.12)$$

Equation 4.12 can be expressed in a differential form

$$df(B_t) = f'(B_t) d(B_t) + \frac{1}{2} f''(B_t) dt. \quad (4.13)$$

Equation 4.13 states that the stochastic process $X_t = f(B_t)$ at time t evolves as a Brownian motion with drift $f''(B(t))/2$ and variance $f'(B_t)^2$.

For $t = 1$ and expanding f in a second Taylor approximation

$$f(x + y) = f(x) + f'(x)y + \frac{1}{2} f''(x)y^2 + o(y^2),$$

where $o(y^2)/y^2 \rightarrow 0$ as $y \rightarrow 0$. Similarly, for the example

$$f(B_1) - f(B_0) = \sum_{j=1}^n [f(B_{j/n}) - f(B_{(j-1)/n})].$$

Using the Taylor approximation

$$f(B_{j/n}) - f(B_{(j-1)/n}) = f'(B_{(j-1)/n})\Delta_{j,n} + \frac{1}{2}f''(B_{(j-1)/n})\Delta_{j,n}^2 + o(\Delta_{j,n}^2),$$

where $\Delta_{j,n} = B_{j/n} - B_{(j-1)/n}$.

Therefore, $f(B_1) - f(B_0)$ is equal to the sum of the following limits

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n f'(B_{(j-1)/n}) [B_{j/n} - B_{(j-1)/n}], \quad (4.14)$$

$$\lim_{n \rightarrow \infty} \frac{1}{2} \sum_{j=1}^n f''(B_{(j-1)/n}) [B_{j/n} - B_{(j-1)/n}]^2, \quad (4.15)$$

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n o([B_{j/n} - B_{(j-1)/n}]). \quad (4.16)$$

The Brownian motion increment satisfies the following expression

$$[B_{j/n} - B_{(j-1)/n}]^2 \approx 1/n. \quad (4.17)$$

Also, the limit in equation 4.14 is a simple process approximation to a stochastic integral. Therefore, the limit is

$$\int_0^1 f'(B_t) dB_t. \quad (4.18)$$

4.4.3 Stratonovich Integral

An alternative method to the Itô integral is the Stratonovich integral. In general a function $f(t, \omega)$ can be approximated using the following expression ([Øksendal, 2003](#)):

$$\sum_j f(t_j^*, \omega) \cdot \chi_{t_j, t_{j+1}}(t), \quad (4.19)$$

where t_j^* belong to the intervals $[t_j, t_{j+1}]$, and define the limit

$$\int_S^T f(t, \omega) dB_t(\omega) = \lim_{n \rightarrow \infty} \sum_j^S f(t_j^*, \omega) [B_{t_{j+1}}(\omega) - B_{t_j}(\omega)]. \quad (4.20)$$

In stochastic integrals, the selection of t_j^* changes the output, unlike the deterministic counterpart, the Riemann-Stieltjes integral. Two choices can be made in order to solve those integrals. The first option is to choose $t_j^* = t_j$, that is, as the left end point, leading to the Itô integral defined in subsection 4.4.2. The second option is to consider t_j^* as the mid-point of the interval $[t_j, t_{j+1}]$, that is, $t_j^* = (t_j + t_{j+1})/2$, leading to the Stratonovich integral

$$\int_S^T f(t, \omega) \circ dB_t(\omega). \quad (4.21)$$

4.4.4 Strong Law of Large Numbers and Central Limit Theorem

Given a probability space and a random variable $X \in \mathbb{R}$, which returns the outcome of a random experiment, repetitions of the random experiment can be done by introducing a sequence of random variables X_1, \dots, X_n . Each random variable contains the same probabilistic information as X . That is, the sequence X_1, \dots, X_n is identically distributed if (Evans, 2012)

$$F_{X_1}(x) = F_{X_2}(x) = \dots = F_{X_n}(x), \quad \forall x. \quad (4.22)$$

If it is also assumed that X_1, \dots, X_n are independent, then that sequence of random variables are said to be independent, identically distributed random variables.

4.4.4.1 Strong Law of Large Numbers

Theorem 4.4.2. *Let X_1, \dots, X_n be a sequence of independent, identically distributed random variables defined on the same probability space. If $m := E(X_i)$ for $i = 1, \dots,$*

then

$$P\left(\lim_{n \rightarrow \infty} \frac{X_1 + \cdots + X_n}{n} = m\right) = 1. \quad (4.23)$$

4.4.4.2 Central Limit Theorem

An important characteristic of the Brownian motion (to be discussed in section 4.5) is that the increments of a sample path are normally distributed. The justification is made by using the *central limit theorem*. In general, the limit distribution of the independent, identically distributed random variables X_1, X_2, \dots, X_n with mean μ and variance $\sigma^2 < \infty$. Let

$$Z_n = \frac{(X_1 + \cdots + X_n) - n\mu}{\sigma\sqrt{n}}, \quad (4.24)$$

and

$$\Phi(b) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \quad (4.25)$$

denotes the standard normal function. With the previous equations, the central limit theorem can be formulated as follows (Lawler, 2016):

Theorem 4.4.3. *The distribution of Z_n approaches a standard normal distribution when $n \rightarrow \infty$, that is, if $a < b$, then*

$$\lim_{n \rightarrow \infty} P\{a \leq Z_n \leq b\} = \Phi(b) - \Phi(a). \quad (4.26)$$

Equation 4.26 states that any distribution of X with finite variance, the scaled random variable is approximately a normal distribution.

4.5 Wiener Process

The Wiener process is a stochastic process introduced by the biologist Robert Brown in the 1820s (Mikosch, 2000). In his model, known as Brownian motion, Brown

studied that the motion of microscopic particles in liquid are random depend on different perturbations, noise, etc. These perturbations are caused by the influence of the molecules present in the liquid. The first mathematical representation of Brownian motion was presented by Wiener (1923). The model represents the displacement of the particle over a time interval as the sum of all independent effects on that particle that are normally distributed with mean zero and variance proportional to the length of the time interval.

The Wiener process has been applied in finance (Meyer and Saley, 2003; Bianchi, 2005), biology (Cohen and Moerner, 2006; Kadloor et al., 2012), engineering (Pandey et al., 2005; Guerin et al., 2010; Yea and Xie, 2015; Soleimanmeigouni et al., 2016), among other fields. The advantage of the Wiener process in degradation models is that it provides a representation for non-strictly monotonically increasing functions, allowing the incorporation of uncertainty into the model due to random noise, measurement errors and other random perturbations to the degradation data.

The Brownian motion or standard Wiener process is defined as follows:

Definition 4.5.1. $B_t = (B_t, t \in [0, \infty))$ is called a standard Brownian motion or a Wiener process if (Mikosch, 2000):

- It starts at zero: $W_0 = 0$
- It has stationary, independent increments
- For every $t > 0$, B_t is normally distributed ($N(0, t)$)
- Does not have “jumps”, that is, it has continuous sample paths

Definition 4.5.1 is also illustrated in Figure 4.6, in which a single realization of the standard Brownian motion, also known as the sample path, is presented.

The sample path presented in Figure 4.6 shows the variable time in the x axis and the Brownian motion values as a dependent variable in the y axis. The path starts at time zero (0), and its corresponding value for the Brownian motion is zero (0), that is, $\omega_0 = 0$.

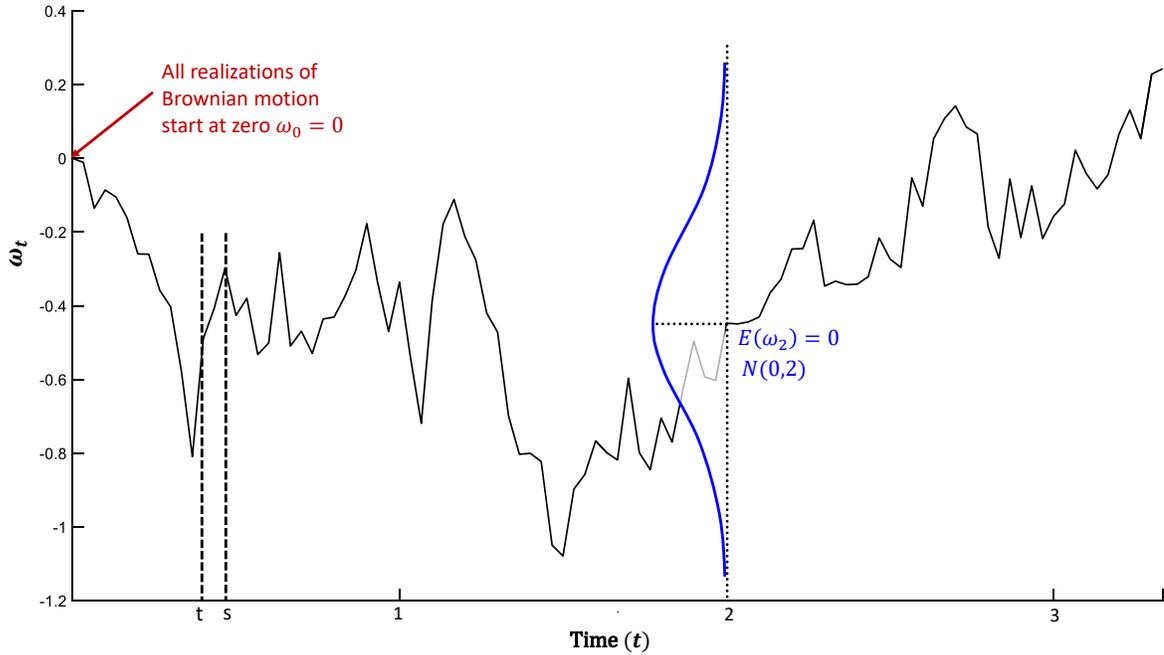


Figure 4.6: Realization of the Brownian motion

The increments are also characterized for being stationary. In general, a stationary process refers to a process that does not change over time. Hence, the Brownian motion increments are stationary if the expected function is constant and the covariance function depends only on the distance $|t - s|$. However, that concept cannot be extended to Brownian motion, which is not stationary because the variance increases with time t .

The Brownian motion is said to have stationary increments if

$$\omega_t - \omega_s \stackrel{d}{=} \omega_{t+h} - \omega_{s+h} \quad \forall t, s \in T \text{ and } h \text{ with } t+h, s+h \in T. \quad (4.27)$$

In addition, the Brownian motion is continuous in time with with independent Gaussian increments, that is, that the random variables $\omega_t - \omega_s$ and ω_{t-s} are normally distributed $N(0, t - s)$ for $s < t$. That is, if for every time selection $t_i \in T$ with $t_1 < \dots < t_n$ and $n \geq 1$ (Mikosch, 2000)

$$\omega_{t_2} - \omega_{t_1}, \dots, \omega_{t_n} - \omega_{t_{n-1}} \quad (4.28)$$

are independent random variables. As a result, the Wiener process exhibits the Markov property due to the independent increments.

Another characteristic presented in Definition 4.5.1 is that for any $t > 0$, all the possible values of $\omega(t)$ are normally distributed with mean $E(w(t)) = 0$ and variance $\sigma^2 = t$. As shown in Figure 4.6, for $t = 2$, the distribution of the Brownian motion or Wiener process evaluated at $t = 2$, $\omega(2)$ is $N(0, 2)$.

The finite-dimensional distributions (fidis) of the standard Brownian motion are defined as multivariate Gaussian distributions. In general, the Brownian motion is the accumulation of a set of normally distributed random variables. Between each time interval, the increment of the Brownian motion is normally distributed with mean 0 and variance $\sigma^2 \Delta t$. Also, Definition 4.5.1 states that Brownian motion, unlike Poisson processes, has continuous sample paths, being more suitable to describe a process with independent, stationary, and continuous increments.

Figure 4.7 presents different sample paths at various numbers of steps $N = 10$, $N = 100$, $N = 1000$, and $N = 10000$ in the time interval $[0, 1]$. The figure shows that the higher the number of iterations or the smaller the time step size, the more uncertainty that is captured by the process.

4.5.1 First Hitting Time

The first hitting time (FHT) is a natural event for stochastic processes (Lee and Whitmore, 2006). Examples of stochastic processes and their FHT include: in the Markov chain context, the first hitting time is expressed in terms of the number of steps required to the system to reach the absorbing state; the absorbing state is the analogous to the threshold for Wiener process. For a Bernoulli process, the FHT is the number of trials needed until success. For the gamma process, the FHT is the inverse gamma distribution. For the Poisson process, the FHT has an Erlang distribution. Finally, for the Wiener process, the FHT is expressed in terms of the inverse Gaussian distribution. Lee and Whitmore (2006) presented an extensive summary stochastic processes and first hitting time. The applications include engineering (Bian and Gebraeel, 2012; Hao

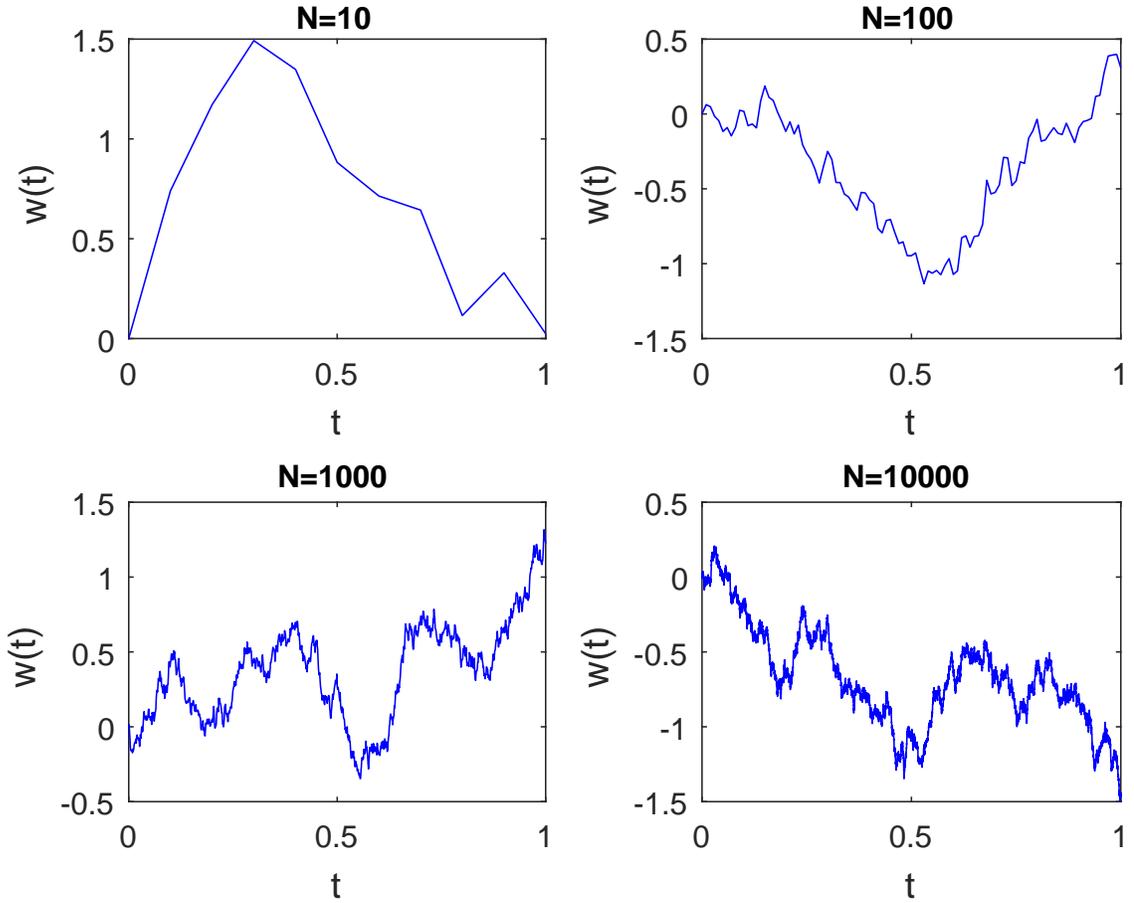


Figure 4.7: Wiener process sample paths at different number of steps (N)

and Su, 2014), medicine (Pennell et al., 2010), biology (Sæbø et al., 2005), and social sciences (Smith and Ratcliff, 2004), among others.

The FHT can be interpreted in degradation analysis as the moment at which the system reaches a specified threshold (Figure 4.8).

In general, the first hitting time model is composed by (Lee and Whitmore, 2006): a) a parent stochastic process $\{W(t), t \in T, w \in \Omega\}$, where T is the time space and χ is the state space of the process (Wiener process in this research) and b) a threshold set H expressed as

$$T = \inf\{t \mid W(t) \geq a\}, \quad (4.29)$$

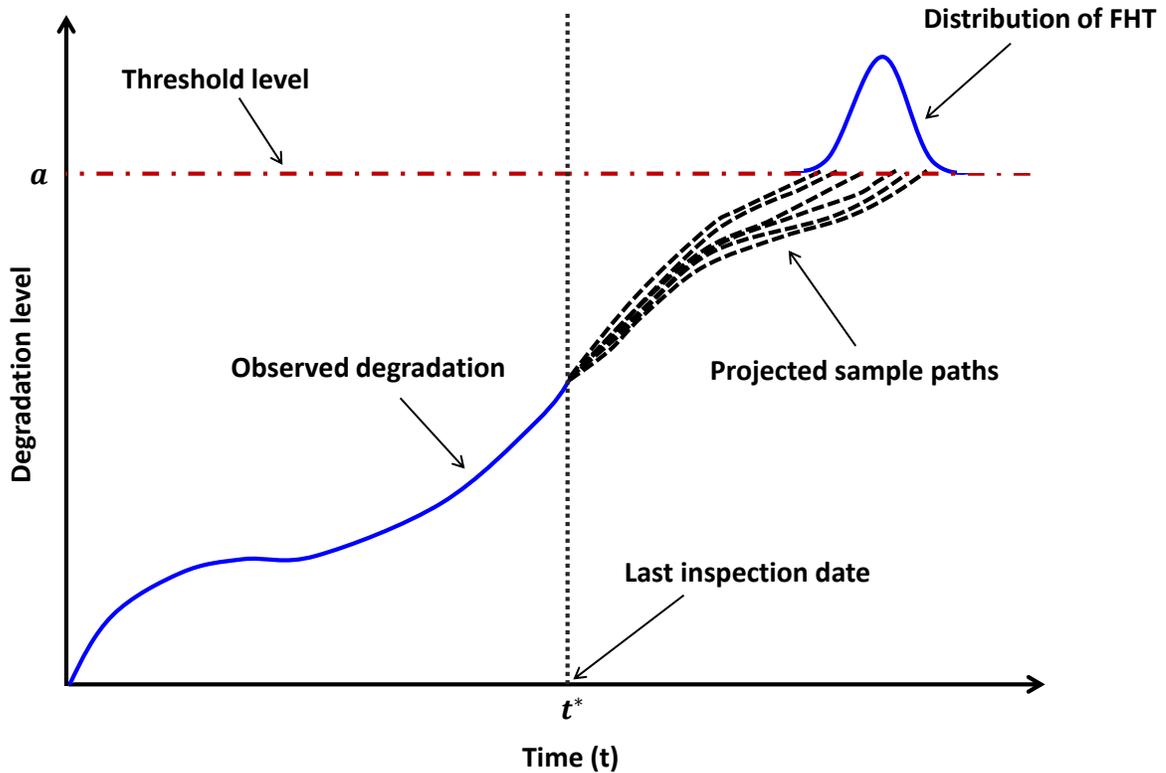


Figure 4.8: Schematic representation of the threshold-regression model

with $a \in H$.

Lemma 4.5.1. *Let t_a be the first time the degradation model (represented by a Brownian motion) W_t hits a predefined threshold level a . The distribution of t_a is formulated as (Calin, 2015)*

$$P(t_a \leq t) = \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^{\infty} e^{-y^2/2} dy. \quad (4.30)$$

Proof. Consider A and B as two events, where A represents the threshold level A , and B represents the first hitting time t_a . The probability of A can be represented by the following expression

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap \bar{B}) \\ &= P(A | B)P(B) + P(A | \bar{B})P(\bar{B}). \end{aligned}$$

If $a > 0$, $A = \{w; W_t(w) \geq a\}$, and $B = \{w; t_a(w) \leq t\}$, then

$$\begin{aligned} P(W_t \geq a) &= P(W_t \geq a \mid t_a \leq t) \\ &P(t_a \leq t) + P(W_t \geq a \mid t_a > t)P(t_a > t). \end{aligned} \tag{4.31}$$

If $t_a \geq t$, it can be concluded that the Brownian motion did not reach the threshold level a , that is, $W_t < a$. Therefore:

$$P(W_t \geq a \mid t_a > 0) = 0. \tag{4.32}$$

Likewise, if $t_a \leq t$, then $W_{t_a} = a$.

Since the standard Wiener process is a Markov process, the process starts at t_a . In addition, because of the symmetry of the probability density function of a normal distribution, the Brownian motion W_t has equal chances to go up or go down after the time interval $t - t_a$, therefore

$$P(W_t \geq a \mid t_a \leq t) = \frac{1}{2}. \tag{4.33}$$

By substituting equations 4.32 and 4.33 into 4.31:

$$P(t_a \leq t) = \frac{P(W_t \geq a) - P(W_t \geq a \mid t_a > t)P(t_a > t)}{P(W_t \geq a \mid t_a \leq t)}.$$

$$P(t_a \leq t) = 2P(W_t \geq a)$$

$$\begin{aligned} P(t_a \leq t) &= \frac{2}{\sqrt{2\pi}} \int_a^\infty e^{-x^2/(2t)} dx \\ &= \frac{2}{\sqrt{2\pi}} \int_{a/\sqrt{t}}^\infty e^{-y^2/2} dy. \end{aligned} \tag{4.34}$$

If $a < 0$, then the distribution of t_a is the same as t_{-a} :

$$P(t_a \leq t) = P(t_{-a} \leq t) \tag{4.35}$$



4.6 Wiener Process in Track Geometry Degradation

As shown in Chapter 2, railway track geometry degradation has a monotonic, but not strictly increasing behavior. There are situations in which the degradation parameter can slightly reduce its value due to measurement errors or minor interventions in the track. Stochastic processes, in general, offer a more flexible and robust environment that represents the increase in the degradation process, allowing a better adaptation of the data in their behavior. Although the Wiener process has been applied in degradation models, it is a new topic in rail track degradation analysis. Chapter 2 also concluded that stochastic rail track degradation has been modeled using the gamma process (Meier-Hirmer et al., 2009) and Markov chains (Shafahi and Hakhamaneshi, 2009). There is one implementation using the Wiener process presented by Soleimanmeigouni et al. (2016) (Figure 4.9). In addition to the Wiener process, shock events, that is, sudden changes (jumps) in the degradation values were considered. The main limitations of this paper is that track geometry degradation model parameters were estimated using the maximum likelihood method which is a frequentist method that provides a single point estimate, instead of the probability distribution of all the possible values for the parameters. Also, time to failure, a natural process in a stochastic process, was not addressed.

Based on what it was presented above, track geometry degradation is modeled as a Wiener process with drift.

Let Y_t be the stochastic component of the degradation for a specific track geometry parameter. The Wiener process with drift is presented as follows

$$Y_t = Y_0 + \mu t + \sigma W_t, \quad (4.36)$$

where:

- Y_t : degradation level at time t
- Y_0 : initial degradation

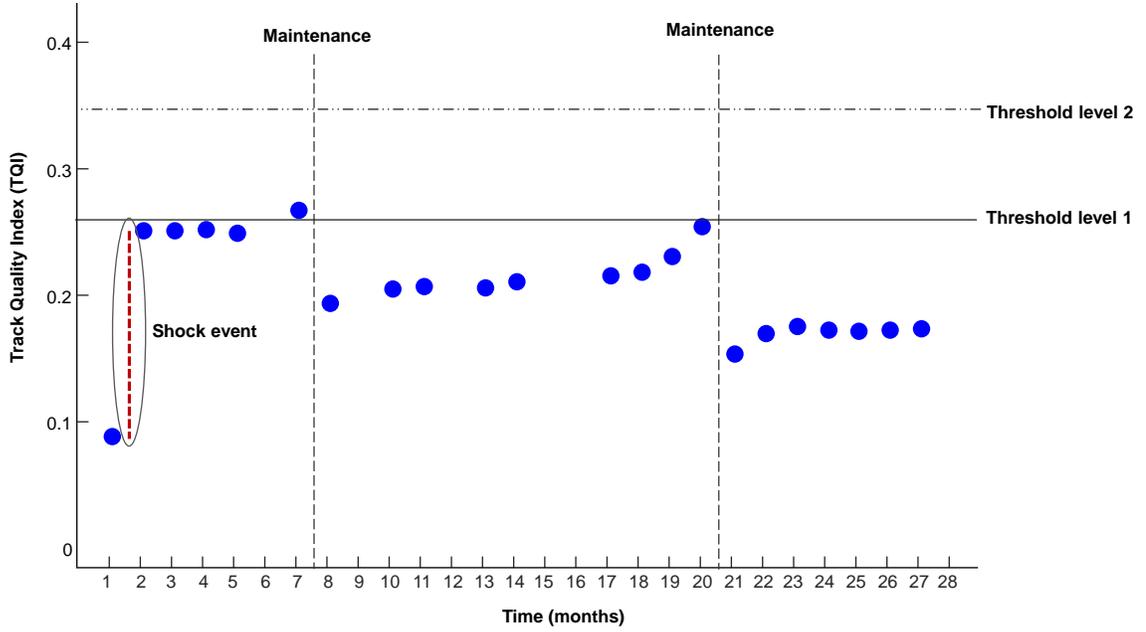


Figure 4.9: Representation of track geometry with shock events

- μ : drift parameter (degradation rate) (in)
- σ : diffusion parameter
- W_t : standard Wiener process (or Brownian motion)
- t : time

As shown in equation 4.36, the Wiener process has two main components: the standard Brownian motion (W_t) and the drift and diffusion parameters that are time dependent and can be estimated using frequentist and Bayesian methods. In this research, the Markov chain Monte Carlo (MCMC) approaches are used to estimate the parameters' posterior distributions.

In terms of time to failure in track geometry, this problem has been traditionally addressed using defect data. Important contributions in this approach include He et al. (2013); Zarembski et al. (2016); Alemazkoor et al. (2017). However, unless it is an accident due to extreme events, track geometry failure is the result of some underlying degradation process.

Therefore, this research considers the Wiener process to achieve two goals:

1. To predict track condition for each maintenance cycle for each section of the track, and
2. To determine the first hitting time, that is, the time at which the system “fails” for the first time.

4.7 Remarks

This chapter introduced the concepts of stochastic process. These concepts included stochastic differential equations (SDE), its differences from ordinary stochastic equations (ODE), the Itô and Stratonovich integrals, and the concept of a sample path. The chapter formally defined track geometry degradation as a Wiener process with drift. Also, the time at which the degradation path first reaches a threshold, called the first hitting time was defined. Based on the literature review performed, time to failure has been traditionally estimated using defect data and the stochastic processes utilized to model track geometry degradation have not been extended to the estimation of the first hitting time. This dissertation defined two goals in terms of the use of Wiener process in track geometry degradation. Chapters 6 and 7 will present the results regarding the implementation of this method.

REFERENCES

- Mohamed Abdel-Hameed. Degradation processes: an overview. In *Advances in Degradation Modeling: Applications to Reliability, Survival Analysis, and Finance*, pages 17–25. 2010. ISBN 978-0-8176-4923-4. doi: 10.1007/978-0-8176-4924-1.
- Negin Alemazkour, Conrad J Ruppert, and Hadi Meidani. Survival Analysis at Multiple Scales for the Modeling of Track Geometry Deterioration. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 0(0):095440971769565, mar 2017. doi: 10.1177/0954409717695650.
- Linkan Bian and Nagi Gebraeel. Computing and updating the first-passage time distribution for randomly evolving degradation signals. *IIE Transactions*, 44(11):974–987, nov 2012. doi: 10.1080/0740817X.2011.649661.
- Sergio Bianchi. Pathwise identification of the memory function of multifractional Brownian motion with application to finance. *International Journal of Theoretical and Applied Finance*, 08(02):255–281, mar 2005. doi: 10.1142/S0219024905002937.
- Ovidiu Calin. *An informal introduction to stochastic calculus with applications*. World Scientific, aug 2015. ISBN 978-981-4678-93-3. doi: 10.1142/9620.
- Adam E Cohen and W E Moerner. Suppressing Brownian Motion of Individual Biomolecules in Solution. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4362–5, mar 2006. ISSN 0027-8424. doi: 10.1073/pnas.0509976103.
- Susanne Ditlevsen and Adeline Samson. *Stochastic biomathematical models*, volume 2058. 2013. ISBN 978-3-642-32156-6. doi: 10.1007/978-3-642-32157-3.
- A. Elsayed. *Reliability engineering*. Wiley, 2013. ISBN 1118309545.
- Lawrence C. Evans. *An introduction to stochastic differential equations*. American Mathematical Society, 2012. ISBN 1470410540.
- Fabrice Guerin, M. Barreau, S. Cloupet, J. Hersant, and R. Hambli. Bayesian Estimation of Degradation Model Defined by a Wiener Process - Application on Disc Brake Wear. *IFAC Proceedings Volumes*, 43(3):74–79, 2010. ISSN 14746670. doi: 10.3182/20100701-2-PT-4012.00014.

- Huiping Hao and Chun Su. A Bayesian Framework for Reliability Assessment via Wiener Process and MCMC. *Mathematical Problems in Engineering*, 2014:1–8, 2014.
- Qing He, Hongfei Li, Debarun Bhattacharjya, Dhairat P. Parikh, and Arun Hampapur. Railway Track Geometry Defect Modeling: Deterioration, Derailment Risk, and Optimal Repair. *Transportation Research Board 92st Annual Meeting*, pages 1–20, 2013.
- Oliver C. Ibe. *Elements of Random Walk and Diffusion Processes*. John Wiley & Sons, Inc, Hoboken, NJ, sep 2013. ISBN 9781118618059. doi: 10.1002/9781118618059.
- R. N. Iyengar and O. R. Jaiswal. Random Field Modeling of Railway Track Irregularities. *Journal of Transportation Engineering*, 121(4):303–308, may 1995. ISSN 0733-947X. doi: 10.1061/(ASCE)0733-947X(1997)123:3(245.2).
- S. Kadloor, R. S. Adve, and A. W. Eckford. Molecular communication using Brownian motion with drift. *IEEE Transactions on NanoBioscience*, 11(2):89–99, jun 2012. doi: 10.1109/TNB.2012.2190546.
- Waltraud Kahle and Axel Lehmann. The Wiener process as a degradation model: modeling and parameter estimation. In *Advances in Degradation Modeling: Applications to Reliability, Survival Analysis, and Finance*, pages 127–146. 2010. ISBN 978-0-8176-4923-4. doi: 10.1007/978-0-8176-4924-1.
- Peter E. Kloeden and Eckhard. Platen. *Numerical solution of stochastic differential equations*. Springer Berlin Heidelberg, 1992. ISBN 9783662126165. doi: 10.1007/978-3-662-12616-5.
- Gregory F. Lawler. *Introduction to Stochastic Calculus with Applications*. Chapman & Hall, 2016. ISBN 9781466570801.
- Mei-Ling Ting Lee and G. a. Whitmore. Threshold Regression for Survival Analysis: Modeling Event Times by a Stochastic Process Reaching a Boundary. *Statistical Science*, 21(4):501–513, nov 2006. ISSN 0883-4237. doi: 10.1214/088342306000000330.
- C. Joseph Lu and William Q. Meeker. Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 35(2):161, may 1993. doi: 10.2307/1269661.
- William Q. Meeker and Luis A. Escobar. *Statistical Methods for Reliability Data*. Wiley, 1998. ISBN 9780471143284.
- C. Meier-Hirmer, A. Senée, G. Riboulet, F. Sourget, and M. Roussignol. A decision support system for track maintenance. In *Computers in Railways X*, volume 1 of *WIT Transactions on The Built Environment, Vol 88*, pages 217–226, Southampton, UK, jun 2006. WIT Press. ISBN 1845641779. doi: 10.2495/CR060221.

- C Meier-Hirmer, G Riboulet, F Sourget, and M Roussinol. Maintenance optimization for a system with a gamma deterioration process and intervention delay: application to track maintenance. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 223(3):189–198, sep 2009. ISSN 1748-006X. doi: 10.1243/1748006XJRR234.
- Bernard De Meyer and Hadiza Moussa Saley. On the strategic origin of Brownian motion in finance. *International Journal of Game Theory*, 31(2):285–319, jan 2003. ISSN 0020-7276. doi: 10.1007/s001820200120.
- T Mikosch. *Elementary Stochastic Calculus*. World Scientific Publ, 2000. ISBN 9810235437. doi: 10.1142/3856.
- B. K. Øksendal. *Stochastic differential equations: an introduction with applications*. Springer, 2003. ISBN 3540047581. doi: 10.1007/978-3-662-03620-4.
- M.D. Pandey, X.-X. Yuan, and Jan M. van Noortwijk. Gamma Process Model for Reliability Analysis and Replacement of Aging Structural Components. In *Ninth International Conference on Structural Safety and Reliability*, pages 2439–2444, 2005. ISBN 90-5966-040-4.
- Michael L. Pennell, G. A. Whitmore, and Mei Ling Ting Lee. Bayesian Random-Effects Threshold Regression with Application to Survival Data with Nonproportional Hazards. *Biostatistics*, 11(1):111–126, 2010. ISSN 14654644. doi: 10.1093/biostatistics/kxp041.
- L.I. Pettit and K.D.S. Young. Bayesian Analysis for Inverse Gaussian Lifetime Data with Measures of Degradation. *Journal of Statistical Computation and Simulation*, 63(3):217–234, 1999. ISSN 0094-9655. doi: 10.1080/00949659908811954.
- S. Sæbø, T. Almøy, B. Heringstad, G. Klemetsdal, and A.H. Aastveit. Genetic evaluation of mastitis resistance using a first-passage time model for Wiener processes for analysis of time to first treatment. *Journal of Dairy Science*, 88(2):834–841, 2005. doi: 10.3168/jds.S0022-0302(05)72748-X.
- Y Shafahi and R Hakhamaneshi. Application of a Maintenance Management Model for Iranian Railways Based on the Markov Chain and Probabilistic Dynamic Programming. *Transaction A: Civil Engineering*, 16(1):87–97, 2009.
- Philip L Smith and Roger Ratcliff. Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3):161–168, 2004. ISSN 01662236. doi: 10.1016/j.tins.2004.01.006.
- Iman Soleimanmeigouni, Alireza Ahmadi, Christophe Letot, Arne Nilsson, and Uday Kumar. Cost-based optimization of track geometry inspection. In *11th World Congress on Railway Research*, pages 1–7, 2016.

- Vasilis A. Sotiris. *Statistical degradation models for electronics. Ph.D. Thesis.* PhD thesis, University of Maryland, 2011.
- Cecília Vale and Simões Lurdes. Stochastic model for the geometrical rail track degradation process in the Portuguese railway Northern Line. *Reliability Engineering and System Safety*, 116:91–98, 2013. ISSN 09518320. doi: 10.1016/j.ress.2013.02.010.
- Norbert Wiener. Differential-Space. *Journal of Mathematics and Physics*, 2(1-4):131–174, oct 1923. doi: 10.1002/sapm192321131.
- Zhi Sheng Yea and Min Xie. Stochastic modelling and analysis of degradation for highly reliable products. *Applied Stochastic Models in Business and Industry*, 31(1): 16–32, 2015. ISSN 15264025. doi: 10.1002/asmb.2063.
- Mohammad Yousefikia, Sara Moridpour, Sujeeva Setunge, and Ehsan Mazloumi. Modeling Degradation of Tracks for Maintenance Planning on a Tram Line. *Journal of Traffic and Logistics Engineering*, 2(2):86–91, 2014. ISSN 23013680. doi: 10.12720/jtle.2.2.86-91.
- Allan M. Zarembski, Daniel Einbinder, and Nii Attoh-Okine. Using multiple adaptive regression to address the impact of track geometry on development of rail defects. *Construction and Building Materials*, 127:546–555, 2016. ISSN 09500618. doi: 10.1016/j.conbuildmat.2016.10.012.

Chapter 5

EXPLORATORY DATA ANALYSIS

5.1 Introduction

Before the implementation of the MCMC and its variants, an exploratory data analysis (EDA) was carried out. EDA seeks to explore data for patterns and relationships without requiring prior hypotheses (Velleman and Hoaglin, 2012). In general, the EDA is an approach that by means of a variety of techniques (most of them graphical) allows a first insight about how the data is characterized, whether there are anomalies in the variables (outliers), whether there exist relationships within the variables, etc.

In this chapter, a detailed explanation about the data set utilized in this research is presented, as well as the exploratory data analysis. The goal of this chapter is to provide a better understanding of the track geometry parameters and the input data that is used to the implementation of the hybrid Bayesian approaches and Wiener process for predicting track geometry degradation and the first hitting time.

5.2 Data Set Description

In this research, an inspection data set for a U.S railroad was utilized. The data set contains information regarding the track geometry parameters for one mile of track measured foot-by-foot, between years 2013-2016, and it was collected using automated inspection with the track geometry car. Each track foot is characterized in terms of 32 different attributes, where 20 of them correspond to variants of geometry parameters. Examples of these variants include:

- Gage
- Crosslevel

- Surface right (62 ft)
- Surface left (62 ft)
- Surface right (124 ft)
- Surface left (124 ft)
- Alignment right (62 ft)
- Alignment left (62 ft)
- Alignment right (124 ft)
- Alignment left (124 ft)
- Warp (62 ft)

Figure 5.1 presents a sample of the data at a specific inspection date for parameters crosslevel, surface left (124 ft), alignment left (124 ft), and warp.

During the period of interest, maintenance activities were performed, of which tamping was one of the most common practices. As presented in Chapter 2, tamping is a maintenance technique with the main goal of correcting the alignment of the rails so they are parallel and level by using the tamping machine, that is located on top of the tie. Then the tamping tines are placed into the ballast, squeezing the ballast under the tie (Audley and Andrews, 2013). Apart from tamping, a major track reconstruction activity was performed. A geocell layer was placed in the track substructure, separating the subgrade and the ballast.

The geometry data is presented in terms of signals, and it is reported in tabular form; there were no missing values, so it can be said that the data set for preprocessing is structured. However, data availability is an important point to make, because only a mile of track was available for analysis, in which a major reconstruction was performed as presented in Figure 5.2. Having said that, the training data used to perform the forecasting model considered only the inspection dates before the geocell layer was placed. This is because since this practice is not frequent, it may significantly bias the results because it is a considerably adjustment of the geometry parameters that

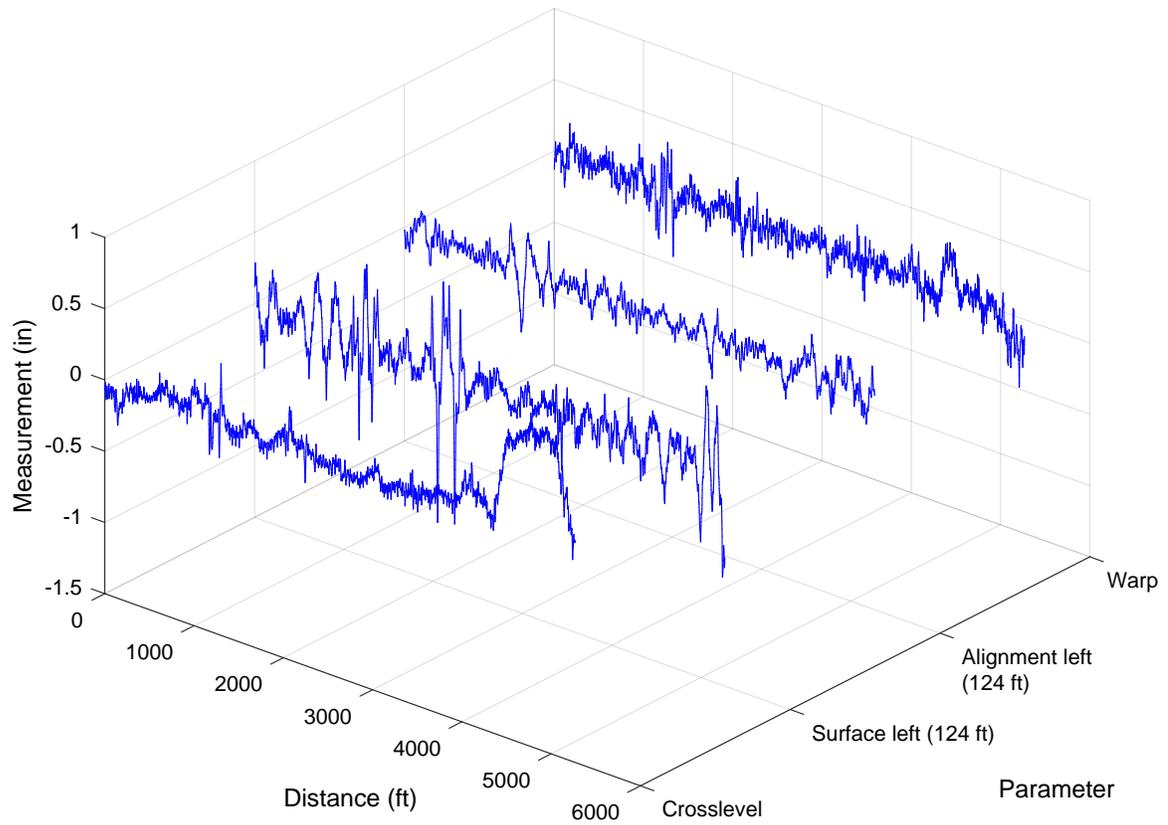


Figure 5.1: Illustration of foot-by-foot measurements for some track geometry parameters

are not common considering tamping activities. The goal then, is to extract as much information as possible that better describes the track geometry degradation and helps predict its behavior from the available data set.

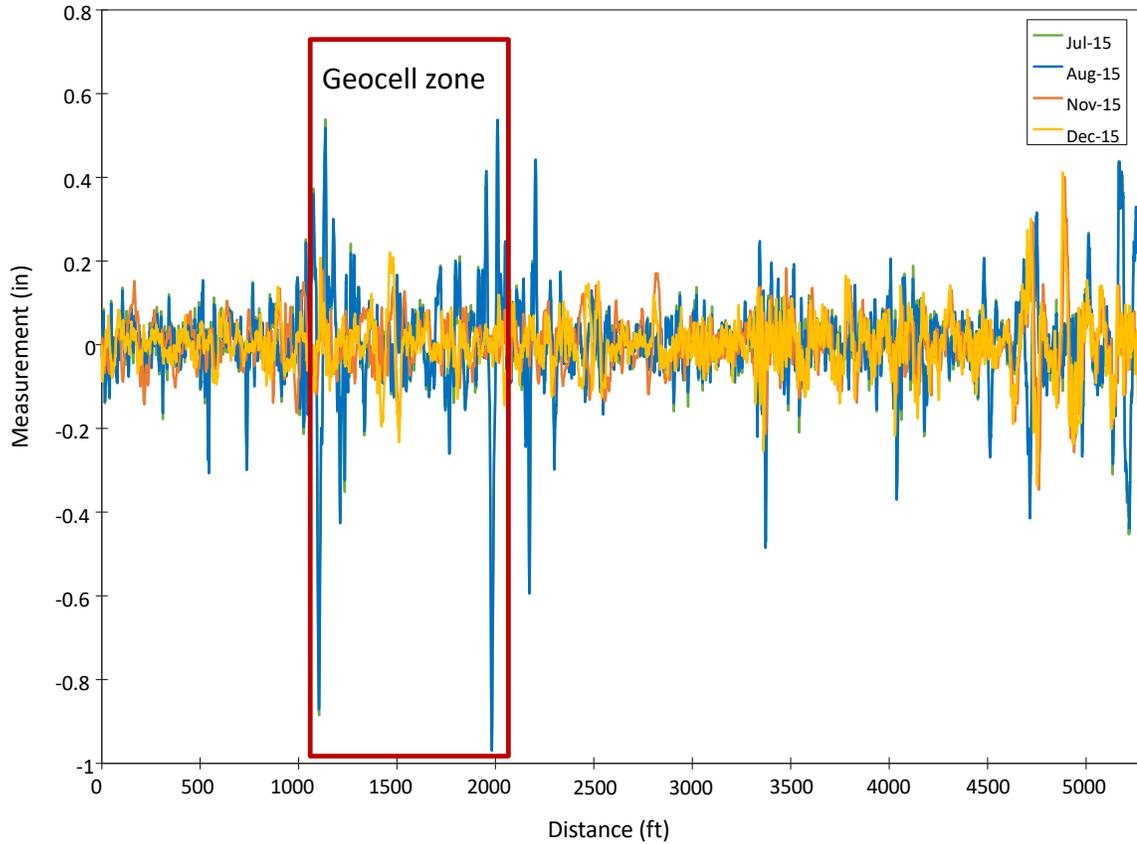


Figure 5.2: Illustration of surface right (62 ft) data

5.2.1 Types of Data Set of EDA

The histograms and normal quantile plots and box plots are presented considering two different subgroups of the data set. One is considering longitudinal data, that is, variables that are fixed at a specific track location but varies over time. Table 5.1 presents a sample of longitudinal data.

Table 5.1: Sample of longitudinal data

Location	Inspection Date						
	Jun-13	Jul-13	Aug-13	...	Sep-13	Mar-16	Apr-16
1	-0.28992	0.0293	0.00916	...	0.05035	0.06561	0.03784

On the other hand, the track geometry parameters can be also analyzed based

on the condition of the whole track at a specific inspection time. In this case, cross-sectional data is utilized. Table 5.2 presents a sample of cross-sectional data.

Table 5.2: Sample of cross-sectional data

Location	Inspection date t
1	-0.289916992
2	-0.250447591
3	-0.196411133
4	-0.135559082
\vdots	\vdots
$n - 3$	0.036010742
$n - 2$	0.022705078
$n - 1$	0.012003581
n	-0.003356934

Both types of data sets are important because they provide insight of the characteristics of the track, as well as similarities of differences between a single parameter that is analyzed over time at a specific location, versus the same parameter at a fixed time but at different locations. This also provides an idea of the contribution of a specific location of the whole available track length.

5.3 Graphical Methods for EDA

5.3.1 Histogram and Quantile-Quantile Plot

The histogram is a graphical method that shows the underlying distribution based on frequency of a continuous random variable, as well as showing if there is an indication of skewness, among others. On the other hand, the QQ plot presents the data in such way that shows whether the data set comes from a normally distributed population. This graphical method compares, based on quantiles, the observed data points with data points predicted from a standard normal distribution. If there is a systematic deviation from the line in the QQ plot, then it can be assumed that the data are not normal (Thode, 2002).

5.3.1.1 Histograms and QQ Plots for Longitudinal Data

Figures 5.3 and 5.4 show an example of histograms and QQ plot for six different geometry parameters; all of them measured at a specific location of the track in the 2013-2016-year time frame. Figure 5.3 shows that data points for crosslevel at a specific track location, has an S-shape distribution. For surface right (62 ft), it can be said that most of the data points are located over the theoretical line. However, it has a long left tail, as shown in both the histogram and QQ plot. For alignment right (62 ft), the histogram does not show the typical Gaussian shape. Instead, the QQ plot shows that the data points follow an S-shape, in which center data and the tails are not located over the theoretical line.

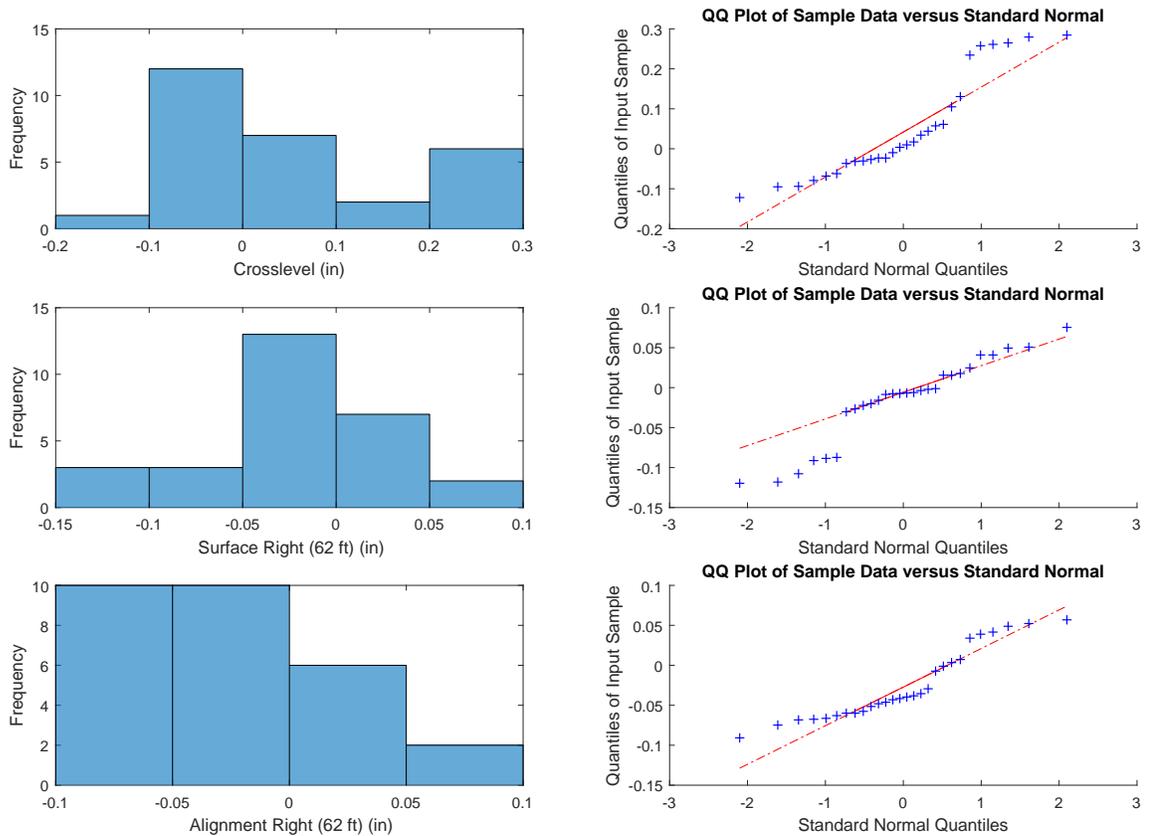


Figure 5.3: Illustration of histogram and QQ plot for crosslevel, surface right (62 ft), and alignment right (62 ft) at a specific track location

Figure 5.4 shows that data points for surface left (62 ft) rely over the theoretical

line in the QQ plot. This parameter shows a more symmetric behavior compared to surface right (62 ft) at the same location. For alignment left (62 ft), most of the data points are located over the theoretical line, except from the points located at the right most side, as opposite to alignment right (62 ft). Finally, for warp (62 ft), most of the data points are concentrated in the interval $[-0.1, 0.1]$ inches and located over the theoretical line; however, there is an indication of a potential outlier as shown on the left side of the QQ plot.

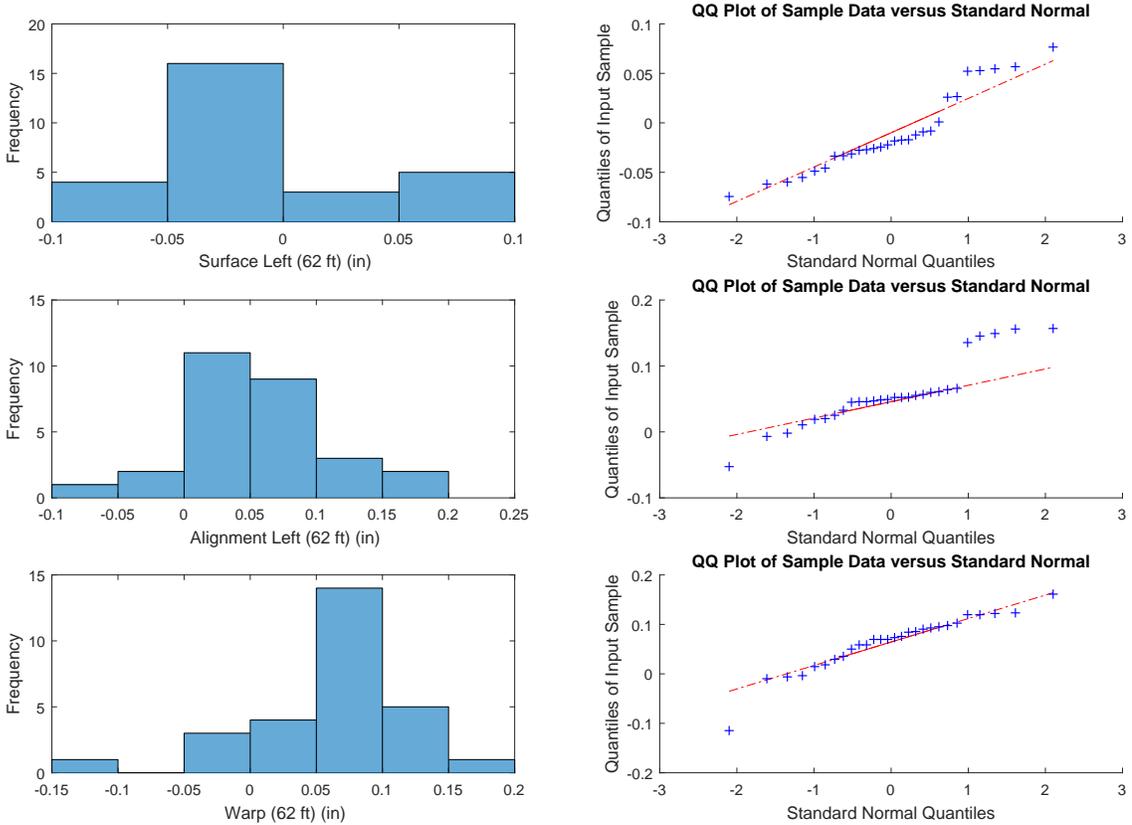


Figure 5.4: Illustration of histogram and QQ plot for surface left (62 ft), alignment left (62 ft), and warp (62 ft) at a specific track location

5.3.1.2 Histograms and QQ Plots for Cross-Sectional Data

Figure 5.5 and 5.6 present the histograms and QQ plots for track geometry parameters crosslevel, surface right (62 ft), alignment right (62 ft), surface left (62 ft), alignment left (62 ft), and warp (62 ft).

Figure 5.5 shows that parameter crosslevel has a Gaussian shape in the left most side of the histogram and QQ plot, and a lower density in the right most side of the plots. This suggests that there is a subgroup in the measurements for this parameter of the track that are different from the majority of the data points. As a result, a bi-modal type distribution can be observed at a specific inspection date. This is explained by a shallow curve in the track under study. For surface right (62 ft) the center data points, which corresponds to the highest concentration of points for this parameter, are located over the theoretical line, and the remaining points correspond to long left and right tails. For alignment right (62 ft), the data points have a Gaussian shape as presented in the histogram and the points rely on the theoretical line in the QQ plot, except from a tail in the left side of the plot.

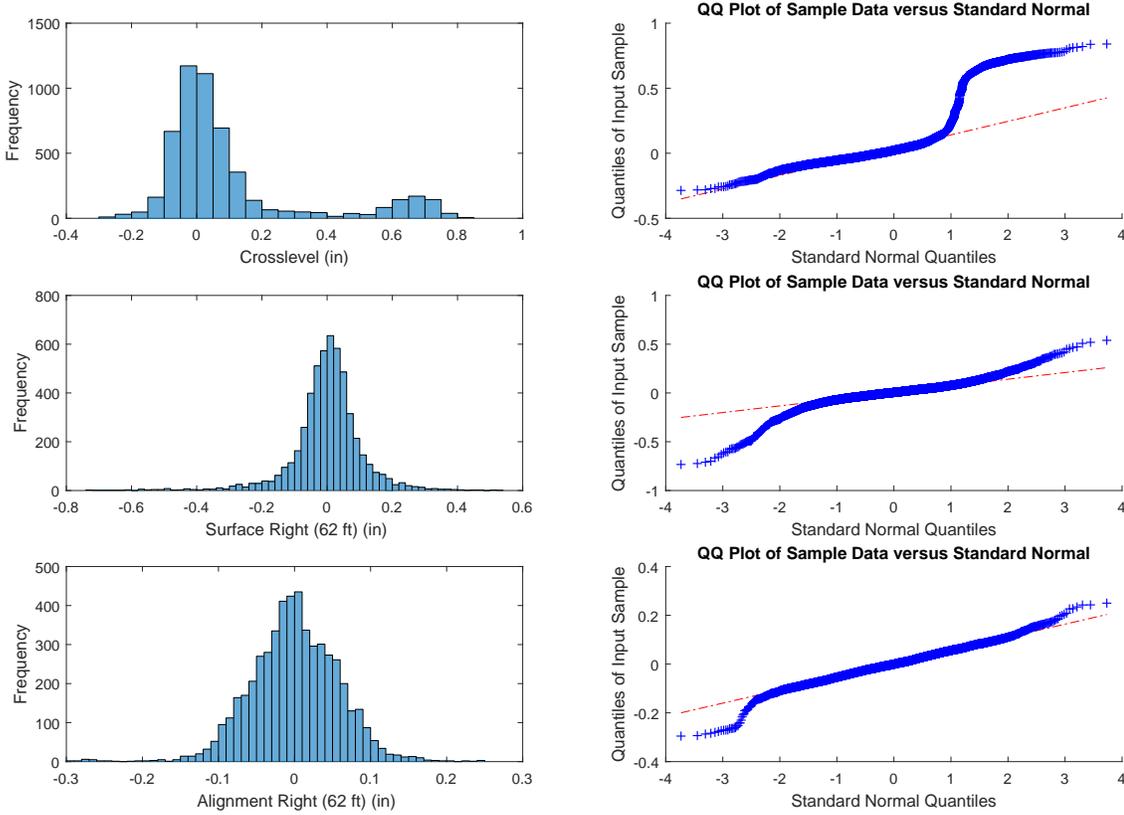


Figure 5.5: Illustration of histogram and QQ plot for crosslevel, surface right (62 ft), and alignment right (62 ft) at a specific inspection date

Figure 5.6 shows that surface left (62 ft) data points are concentrated in the

interval $[-0.2, 0.2]$ as shown in the histogram and is heavy tailed in both the left and right sides of the distribution. The same pattern is presented in surface right (62 ft), in which an S-shape is presented and even though the histogram shows a symmetric shape pattern, a further analysis would be required to determine the underlying distribution for this parameter. Similar to alignment right (62 ft), alignment left (62 ft) also follows a normal distribution based on the histogram and QQ plot. Finally, warp (62 ft) presents a slight S-shape pattern in the QQ plot, skewed to the left side of the distribution.

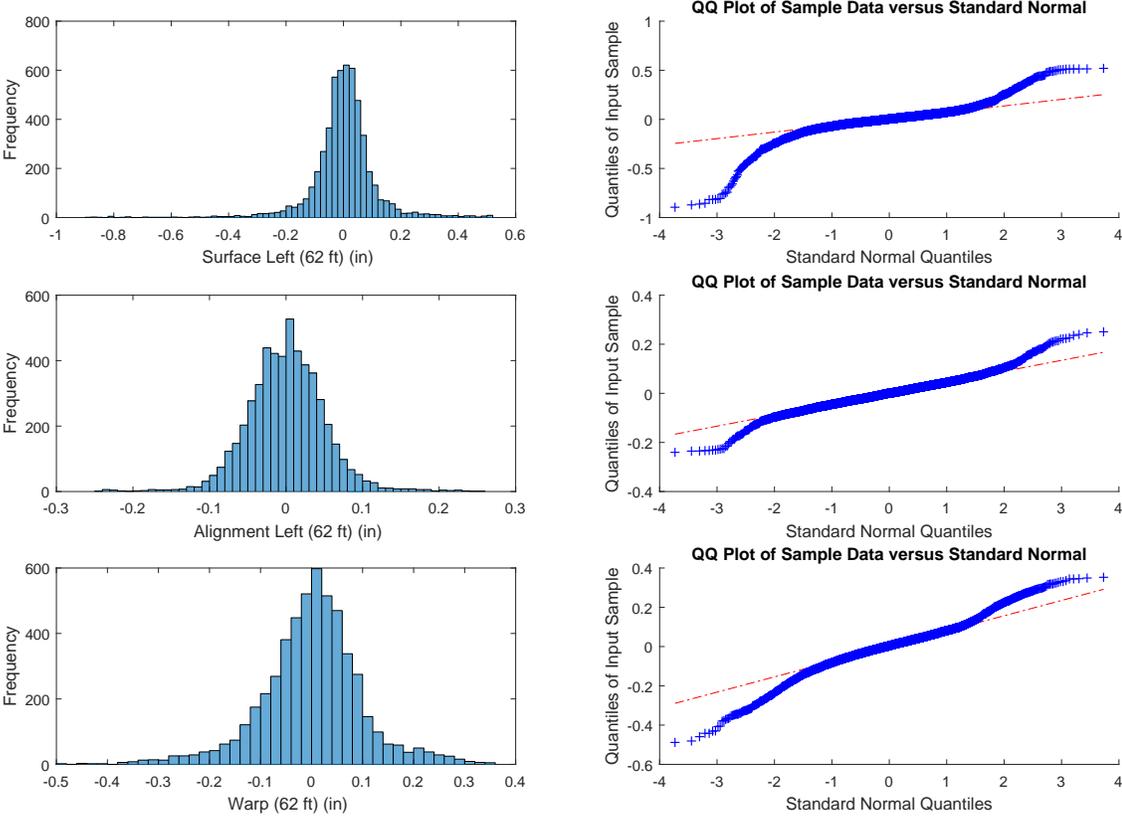


Figure 5.6: Illustration of histogram and qq plot for surface left (62 ft), alignment left (62 ft), and warp (62 ft) at a specific inspection date

5.3.2 Box Plot

Another important behavior to examine in a exploratory data analysis, is whether there is presence of any potential outliers in the data set. This can be done by using

the box plot. The box plot is a graphical method that summarizes the data by using the median and the upper and lower quartiles, that is, the 25th and 75th percentiles. The box plot is very useful because it helps describing the behavior of the data in the middle and at the ends of the distributions. If the lower quartile is $Q1$ and the upper quartile is $Q3$, then the difference ($Q3 - Q1$) is called the interquartile range (IQ). Any values that are outside the box plot can be interpreted as potential outliers.

Similar to the histogram and QQ plots, the box plots are presented by viewing the data set as longitudinal data (that is, a fixed track location in different time) and cross-sectional data, which is done by observing the entire population (the mile of track in this case) at a fixed inspection date.

5.3.2.1 Box Plot for Longitudinal Data

Figures 5.7-5.9 present the box plot for longitudinal data for parameters crosslevel, surface right (62 ft), and alignment right (62 ft) for selected track locations. For the three parameters, the box plots show that adjacent locations have the same or very close median values. For the data subset presented, parameter crosslevel does not present any outliers (Figure 5.7). However, for surface right (62 ft) (Figure 5.8), it is shown that there is high variability in the distributions at different locations. For alignment right (62 ft) there is a slight variation on the distribution for each location. The lower whiskers in the box plots mean that geometry parameters values vary among the least positive quartile group, this behavior is the most frequent for the three parameters. Finally, the box plots allow visualization any potential outliers in the data. The potential outliers in Figures 5.8-5.9 are represented as red crosses, in which for surface right (62 ft) presents a high number of those values that required an in-depth analysis.

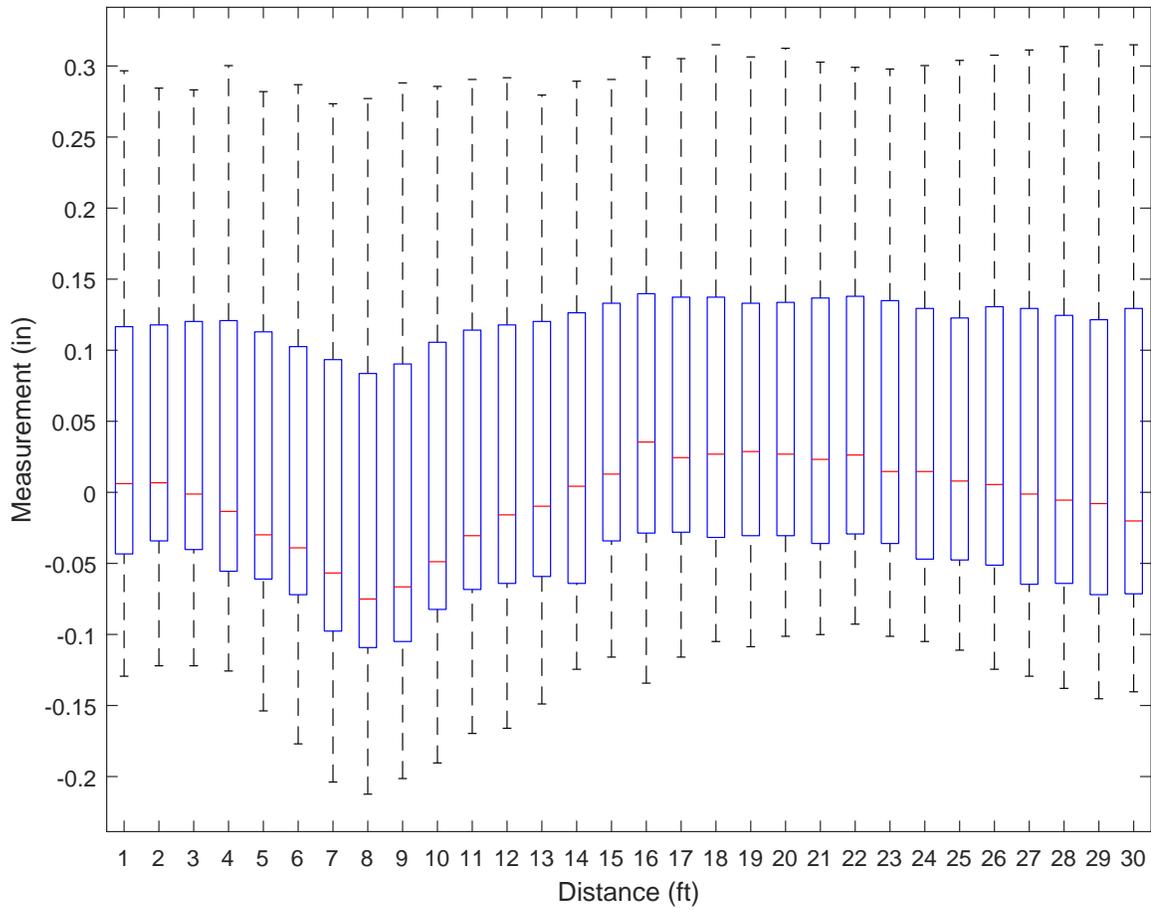


Figure 5.7: Illustration of box plot for crosslevel

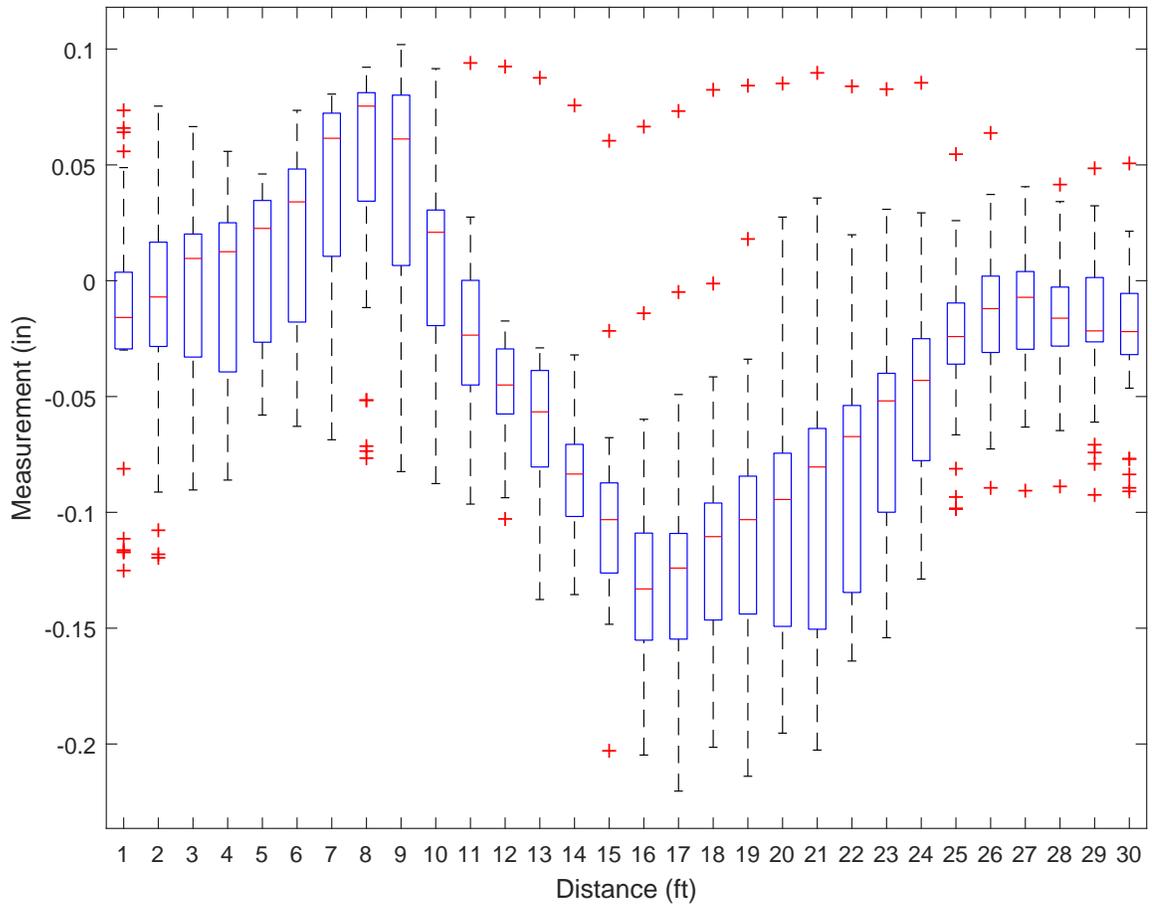


Figure 5.8: Illustration of box plot for surface right (62 ft)

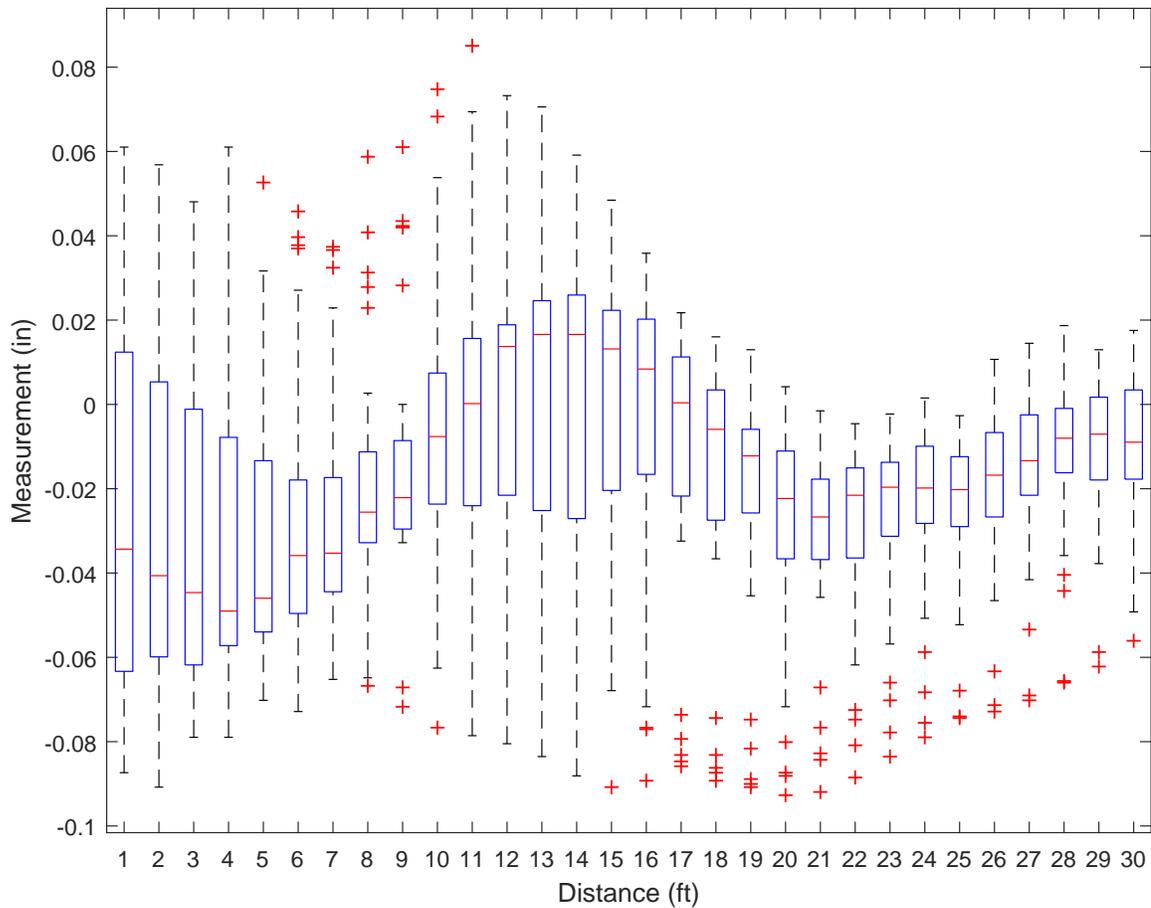


Figure 5.9: Illustration of box plot for alignment right (62 ft)

5.3.2.2 Box Plot for Cross-Sectional Data

By using cross-sectional data, it can be seen that each location of the track have different characteristics and cannot be easily placed in a single category. This is shown in the high number of potential outliers presented as red crosses on each plot. Figure 5.10 presents the box plot for multiple track geometry parameters. The figure shows that the median values for all parameters are around zero and the presence of values that are out of the upper and lower whiskers is very high compared to longitudinal data. This case might be explained by the condition of the track which varies depending on the location. Figures regarding box plots for other inspection dates are presented in Appendix A.

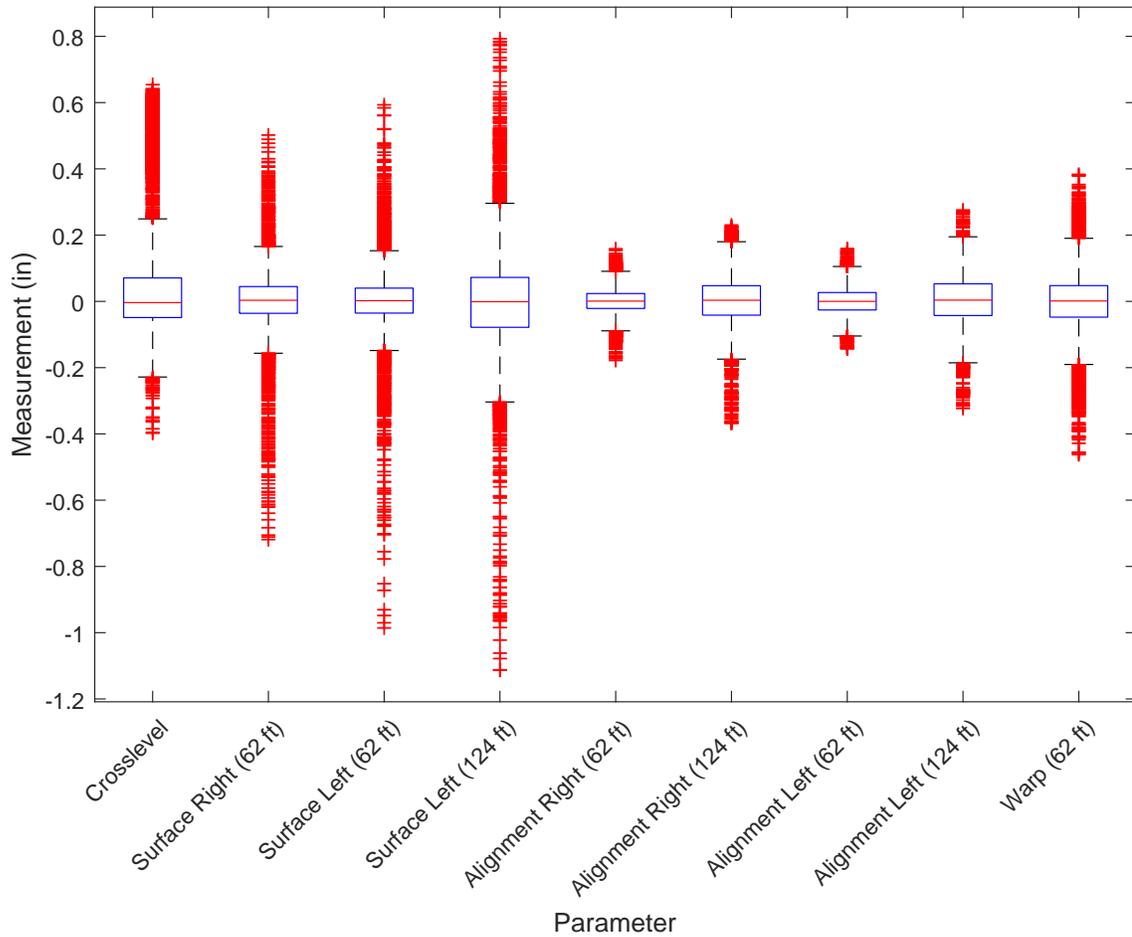


Figure 5.10: Illustration of box plot for selected geometry parameters at a specific inspection time

5.3.3 Correlation Plot

Figure 5.11 presents the correlation matrix plot between pairs of geometry parameters. This plot displays, in the main diagonal the histograms for individual track geometry parameters, and the upper and lower cells show the scatter plot and the Pearson's correlation coefficients for each pair of parameters. Each histogram shows the underlying distribution of the geometry parameters. Symmetric distributions appear to describe most of the parameters except crosslevel, where a bimodal distribution is observed due to the presence of a shallow curve in the track under study. On the other hand, no strong linear correlation was found for any of the variable pairs in the

matrix.

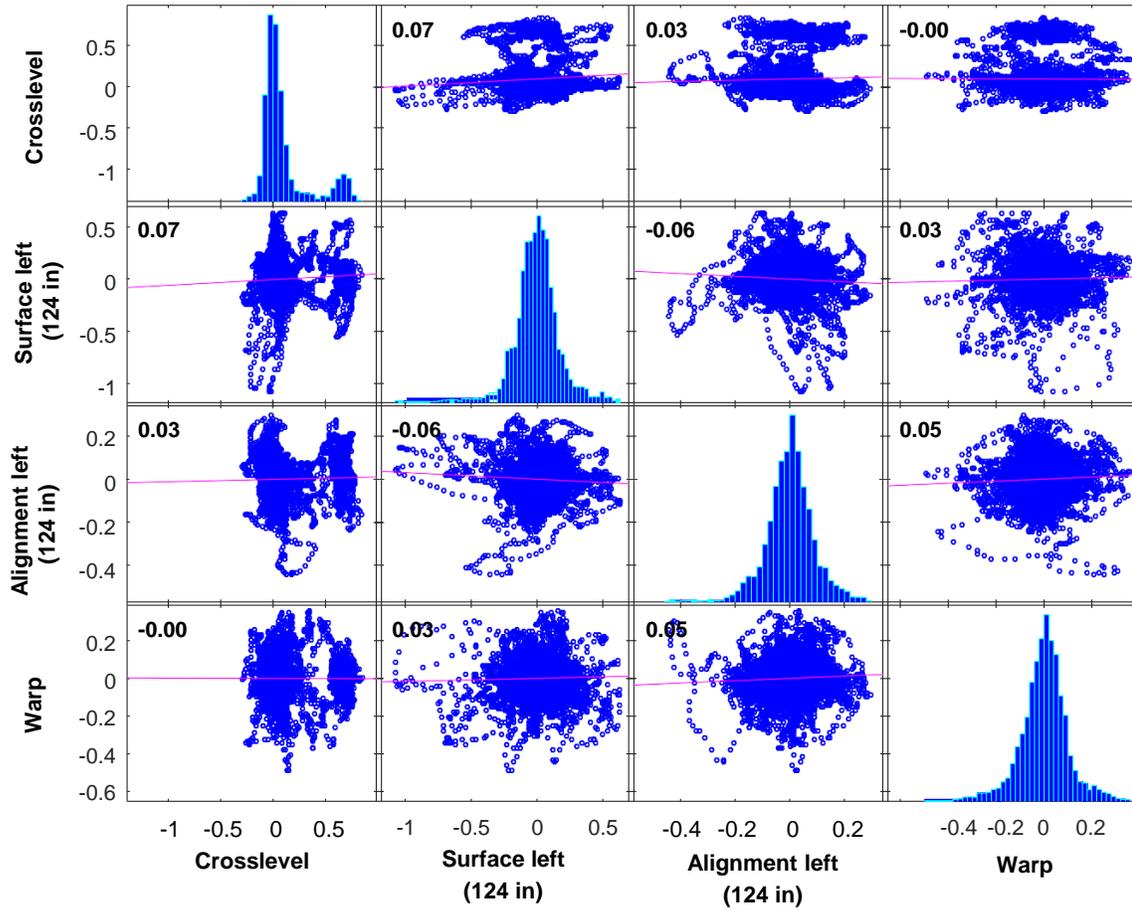


Figure 5.11: Correlation matrix and scatter plots between selected track geometry parameters at a specific inspection date

5.4 Remarks

Based on the exploratory data analysis presented in this chapter, the following conclusions can be made:

1. Due to high variability on the track geometry parameters and the non-strictly monotonic increasing behavior of track geometry degradation, a stochastic approach to modeling track geometry degradation appears to be appropriate.
2. By dividing the data set into longitudinal and cross-sectional data, and performing graphical normality test, it was possible to observe that track locations

have different distributions for different geometry parameters. Parameters surface right and left appear to be the parameters with the highest variability compared to the rest of the parameters. This provide insights about the dominant parameters in track geometry degradation.

3. From the box plots it was possible to observe that there are similarities in terms of the distribution of the parameter values for adjacent locations of the track.
4. Surface right and left are the track geometry parameters with the highest variation compared to the other parameters. This is shown in the box plots, in which they do not have consistent median values along the locations.

REFERENCES

M. Audley and J. Andrews. The effects of tamping on railway track geometry degradation. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(4):376–391, mar 2013. ISSN 0954-4097. doi: 10.1177/0954409713480439.

Henry C. Thode. *Testing for normality*. Marcel Dekker, 2002. ISBN 0203910893.

Paul F. Velleman and David C. Hoaglin. Exploratory data analysis. In *APA handbook of research methods in psychology, Vol 3: Data analysis and research publication*, pages 51–70. 2012. ISBN 1-4338-1006-9. doi: 10.1037/13621-000.

Chapter 6

HYBRID MARKOV CHAIN MONTE CARLO AND WIENER PROCESS IN RAILWAY TRACK GEOMETRY DEGRADATION ANALYSIS

6.1 Introduction

This chapter implements the Metropolis-Hastings algorithm for parameter estimation of the Wiener process model parameters for track geometry degradation. Also, the simulation of the sample paths for multiple track geometry parameters is presented. Based on the characteristics of the track geometry data, the Wiener process appears to be suitable for modeling the degradation process. The results show the potential of using this approach for the prediction of track geometry degradation. To implement the hybrid Bayesian-Wiener process in track geometry degradation, the following methodology was conducted (Figure 6.1).

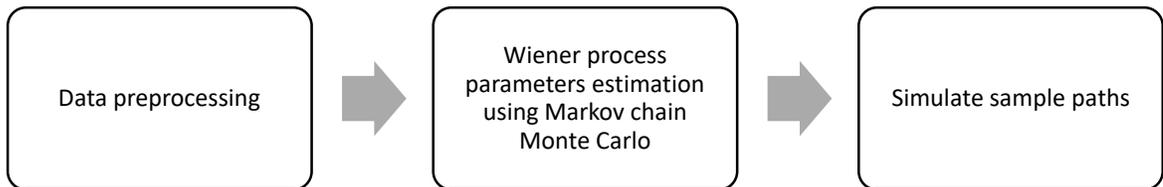


Figure 6.1: General overview for the implementation of the hybrid Bayesian-Wiener process for a single track geometry parameter

6.2 Data Preprocessing

This stage consisted of arranging the data set into the format required for running the MCMC algorithm. The data set described in Chapter 5 was used for this case study. Both cross-sectional and longitudinal data for each track geometry parameter were utilized. The cross-sectional data, that is, taking all data points for each parameter at a specific inspection date, were used to divide the track into a specified number

of sections. In this chapter, degradation is expressed in terms of the measure of the track performance. For that reason, 150-foot and 500-foot sections length were considered to divide the track. After the sections were defined, a track quality index (TQI), based on the standard deviation of each track geometry parameters were calculated. As presented in Chapter 2, there are different approaches to measure the quality of track sections, in which the standard deviation is the one that is most commonly used. Once the TQI for each track section and track geometry parameter were calculated, the data was arranged as longitudinal data, that is, the data was presented in such way that the TQI over time for each track section was observed. Figures 6.2 and 6.3 present a sample of the degradation plots for multiple track geometry parameters at a specific section.

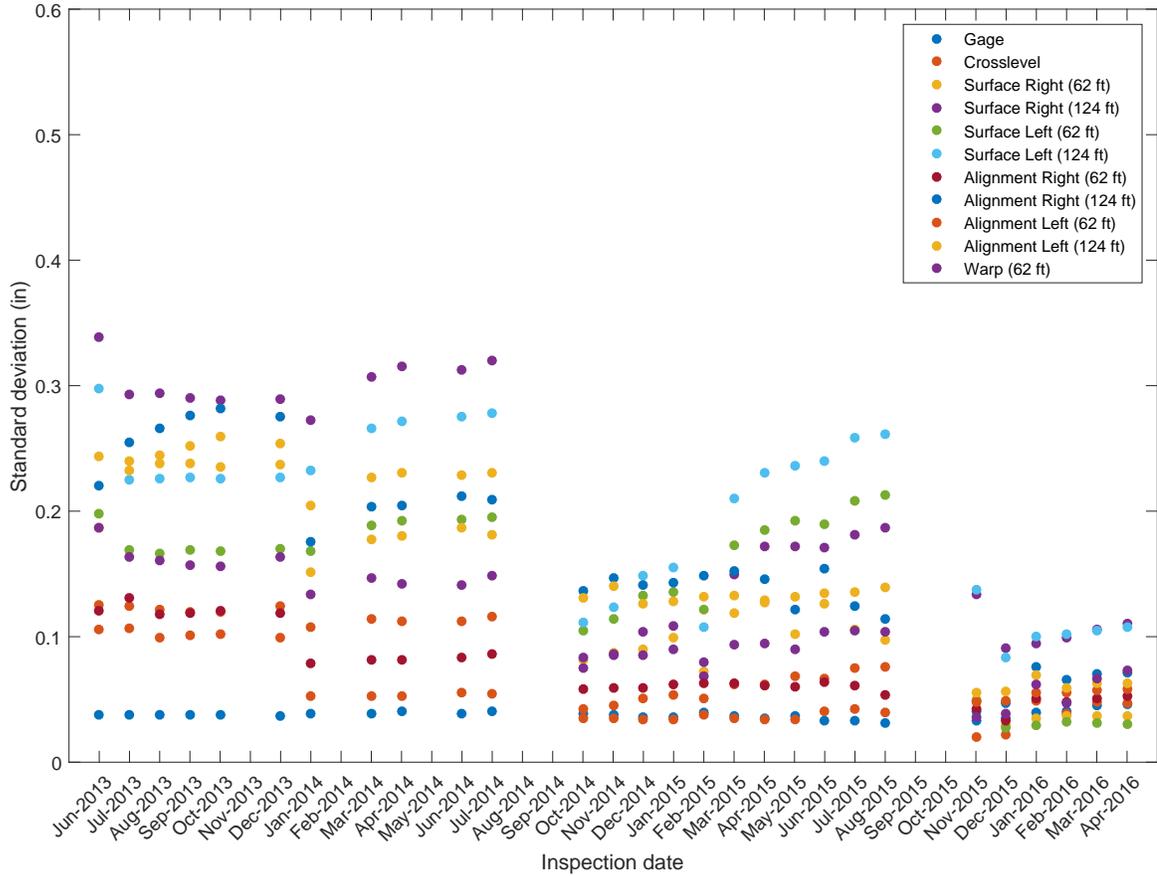


Figure 6.2: Degradation plot for multiple track geometry parameters at a specific 150-foot track section

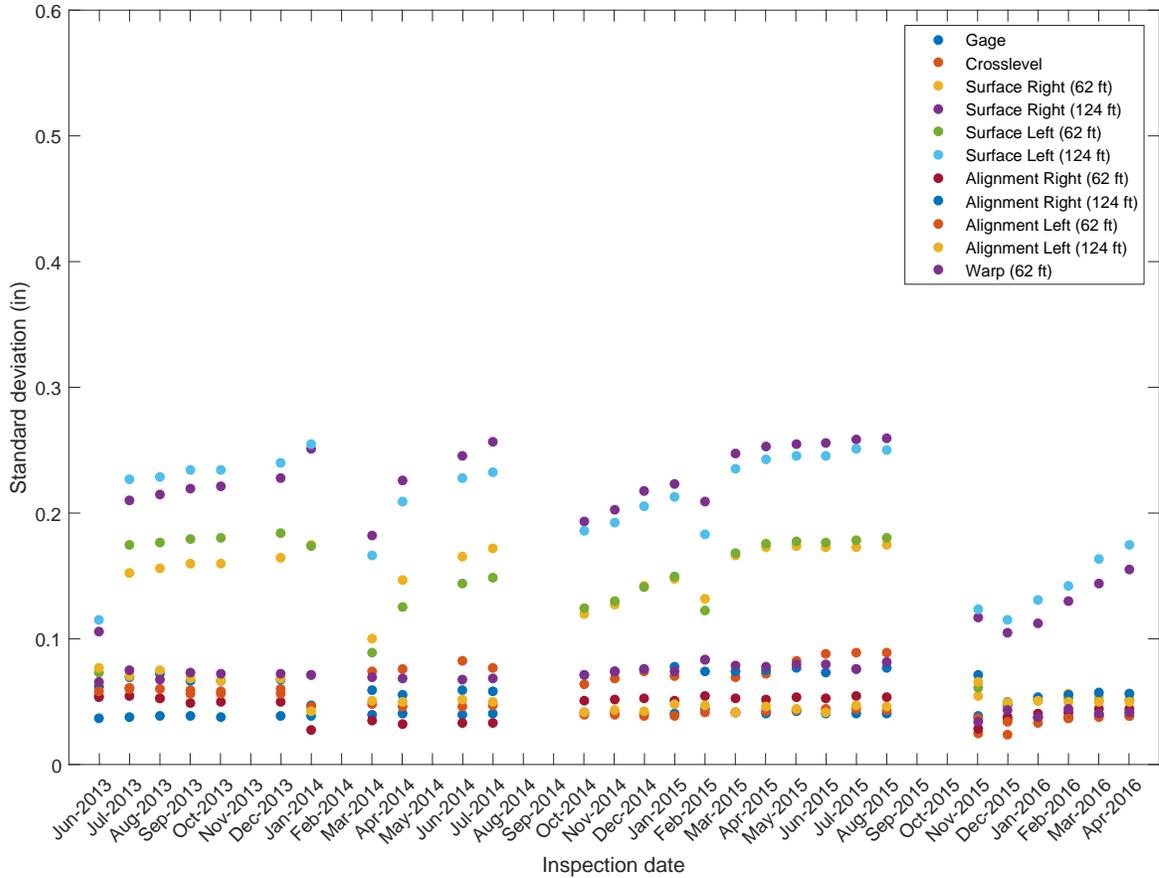


Figure 6.3: Degradation plot for multiple track geometry parameters at a specific 500-foot track section

After comparing the degradation values for each section it was observed that parameter gage reported stable TQI values over time, so in terms of degradation, an increasing or decreasing pattern was not observed for all track sections. Therefore, gage was not considered for further analysis. For the remaining parameters, however, an increasing degradation values in some of the sections were observed. Figures 6.2 and 6.3 show examples of sections with clear degradation and restoration patterns, in which it was observed that the TQI values increased over time. The data also showed the times at which maintenance activities were performed. Parameter surface left measured at a 62 and 124 foot-chord appear be the dominant parameter in terms of higher TQI compared with the rest of the parameters.

Since one of the goals of this dissertation is to predict track degradation and to

predict the first hitting time only, the data points considered as input for estimating the Wiener process parameters were the degradation points before a maintenance activity was performed. That is, only data points for one maintenance cycle.

6.3 Metropolis-Hastings Algorithm Implementation and Output Analysis

To estimate the Wiener process model parameter vector $\theta = [\mu, \sigma]$, a Markov chain Monte Carlo technique based on the Metropolis-Hastings (MH) algorithm was considered. As presented in Chapter 3, the MH algorithm simulates a Markov chain in such a way that the stationary distribution of the Markov chain is the posterior distribution of the model parameters. For the MH implementation, three elements are defined: a) proposal density, b) prior function, and c) likelihood function.

Proposal Function

The MH algorithm can be defined as an extension of the accept-reject method, in which parameter's values are proposed and then accepted or rejected based on the MH ratio. The proposal function, also known as the envelope function, should be spread enough to cover the target density, but not too spread, making the acceptance rate too high. In this case, the algorithm would accept samples that are not a close representation of the model parameters. There is not an exact procedure that formulates this function, it is, in a high number of cases, a result of a trial and error procedure. This procedure is tested based on the acceptance rate obtained. The literature recommends an asymptotically acceptance rate of 0.234 (Roberts et al., 1997). In this research, the proposal function is symmetric and is normally distributed $\sim N(chain_i, sd)$, where $chain_i$ corresponds to the parameter's value at iteration i and sd is the standard deviation tuned by a trial and error procedure.

Prior Function

The prior function incorporates the subjective component of the MH algorithm (and Bayesian inference in general). It allows incorporation of prior knowledge to the calculation of the posterior distribution. In this research, an adaptive prior function

was defined. Adaptive means that the prior distribution would change its shape as the number of iterations increase, that is, as new data are available. This distribution starts as a non-informative prior and the log normal density was defined. In each iteration, the log-normal distribution parameters are updated with the current accepted values of the chain, so this function changes from non-informative to informative prior. Equation 6.1 presents the general formulation of the probability density function for the prior distribution.

$$f(\theta | \nu, s) = \frac{1}{\theta s \sqrt{(2\pi)}} \exp \left[\frac{-(\ln \theta - \nu)^2}{2s^2} \right], \quad (6.1)$$

where ν and s are the log mean and log standard deviation, respectively.

Likelihood Function

The likelihood function represents the probability of observing the data given the parameters. Since the track geometry degradation is modeled based on a Wiener process, degradation increments dy for p sample paths and time j are normally distributed with mean $\mu\Delta t$ and standard deviation $\sigma\Delta t$, the likelihood function is presented in equation 6.2.

$$f(dY | \mu, \sigma) = \prod_{i=1}^p \prod_{j=1}^{q_i} \frac{1}{\sigma \sqrt{2\pi \Delta t_{ij}}} \exp \left[\frac{(dY - \mu \Delta t_{ij})^2}{2\sigma^2 \Delta t_{ij}} \right], \quad (6.2)$$

where p and q are the number of sample paths and number of measures on each sample path, respectively.

MH Output Analysis

As presented in Chapter 3, output analysis plays an important role in Markov chain Monte Carlo routines, because it allows verification of the algorithms' convergence. In this chapter, the output analysis for the Metropolis-Hastings algorithm is presented for 150-foot track sections for parameters crosslevel, surface left (124 ft), alignment left (124 ft), and warp (62 ft). This analysis is presented in terms of plots

(Kernel density, trace, and autocorrelation plot). The output plots for 500-foot track sections are presented in Appendix B.

In addition to the graphic output analysis for the Metropolis-Hastings algorithm, the point estimates were calculated. Tables 6.1 and 6.2 present the summary statistics for track geometry parameters crosslevel and surface left (124 ft). The point estimates were calculated for each track section for parameters μ and σ , which include the mean value, the standard Monte Carlo error, the mode of the distribution, the median, and the mean square error (MSE), which was compared with the frequentist method maximum likelihood estimation (MLE).

Crosslevel

The plots for both drift and diffusion parameters show that the values for each chain are not dependent. Also, the point estimates provide different interpretations of the degradation values increments. As shown in Table 6.1, sections 1, 5, and 7 reported positive drift values. Also, for all the track sections, the standard Monte Carlo error reported small values < 0.65 for drift parameter and < 0.85 for diffusion parameter.

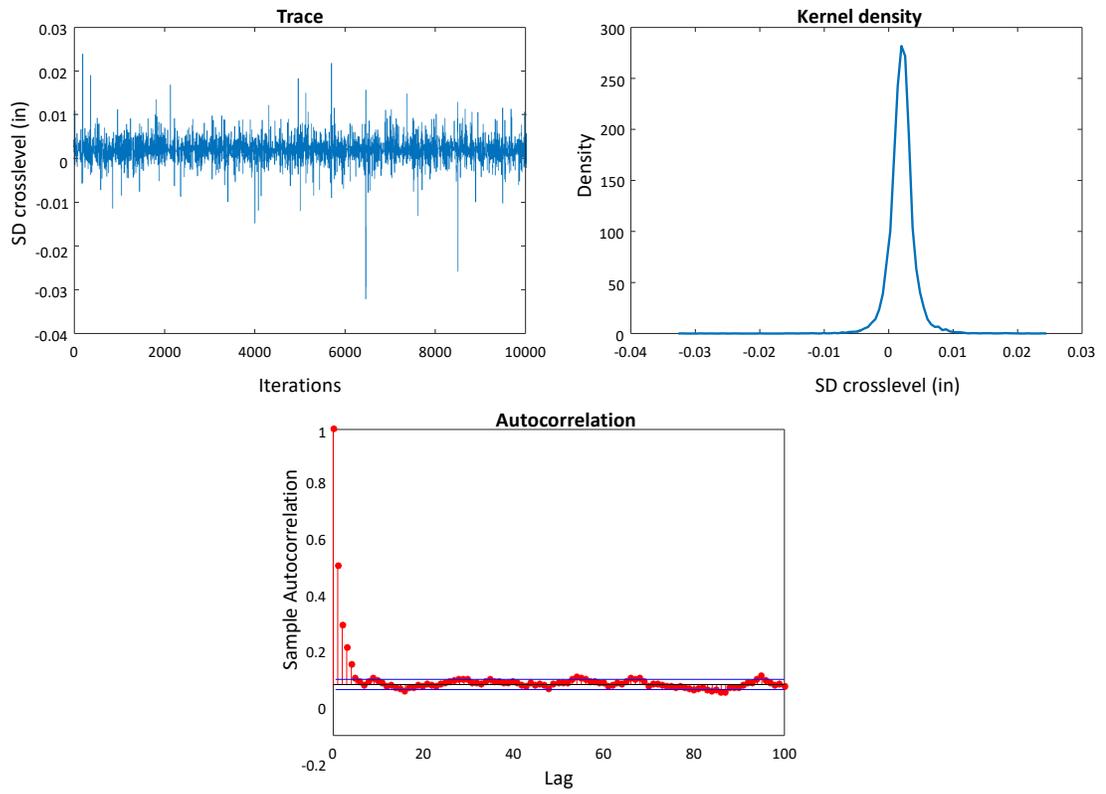


Figure 6.4: Crosslevel: MCMC posterior plots for drift parameter

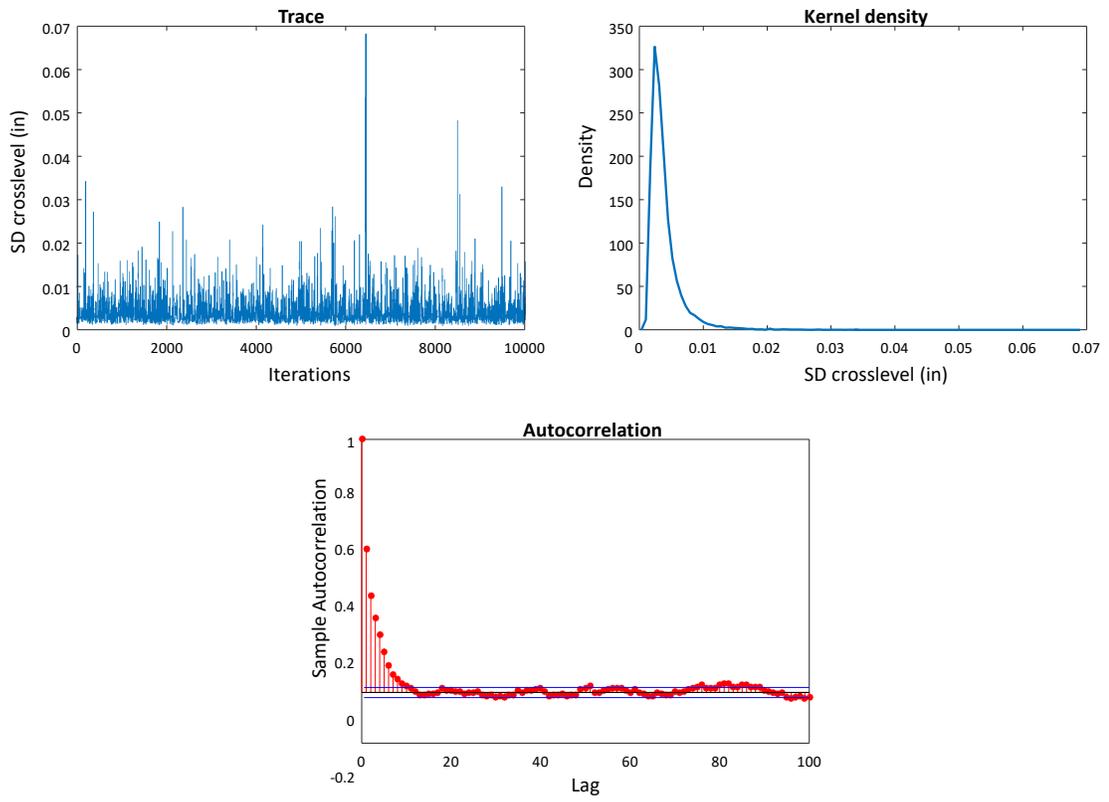


Figure 6.5: Crosslevel: MCMC posterior plots for diffusion parameter

Table 6.1: Output summary for crosslevel track sections in the geocell zone

Section	Mean		Standard Monte Carlo		Mode		Median		MSE	
	μ	σ	error	μ	μ	σ	μ	σ	MCMC	MLE
1	0.00181	0.00909	0.00300	0.00270	0.00188	0.00793	0.00179	0.00858	0.00021	0.00021
2	-0.00304	0.01046	0.00470	0.00584	-0.00383	0.00593	-0.00308	0.00902	0.00006	0.00006
3	-0.00426	0.01227	0.00646	0.00843	-0.00492	0.00569	-0.00434	0.01014	0.00036	0.00036
4	-0.00011	0.00631	0.00139	0.00109	-0.00075	0.00545	-0.00014	0.00617	0.00015	0.00015
5	0.00263	0.00768	0.00342	0.00365	0.00385	0.00618	0.00270	0.00675	0.00004	0.00004
6	0.00082	0.00721	0.00336	0.00381	0.00077	0.00384	0.00079	0.00626	0.00004	0.00004
7	-0.00089	0.00242	0.00131	0.00172	-0.00089	0.00136	-0.00089	0.00198	0.00000	0.00000

Surface Left (124 ft)

For surface left (124 ft), the MCMC output plots show that the chain values reached the stationary distribution faster compared to crosslevel for drift parameter. However, the autocorrelation plot for diffusion parameter shows that it took longer to reach the stationary distribution compared to the drift parameter and to the diffusion parameter for crosslevel. The Kernel density shows that the drift parameter follows a symmetric density, whereas a skewed distribution appears to describe the diffusion parameter. The point estimates reveal that the standard Monte Carlo error for both drift and diffusion parameters are < 0.019 and < 0.02 respectively. Finally, sections 1, 3, 4, 5, and 6 reported positive drift values.

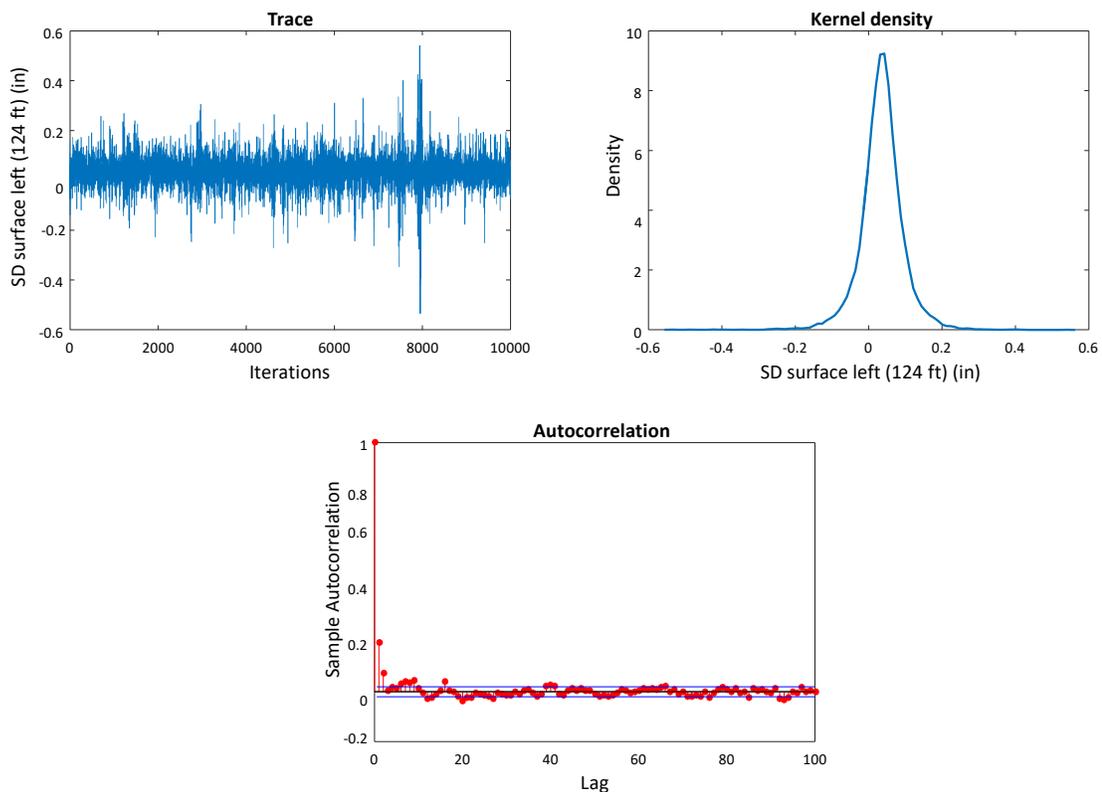


Figure 6.6: Surface left (124 ft): MCMC posterior plots for drift parameter

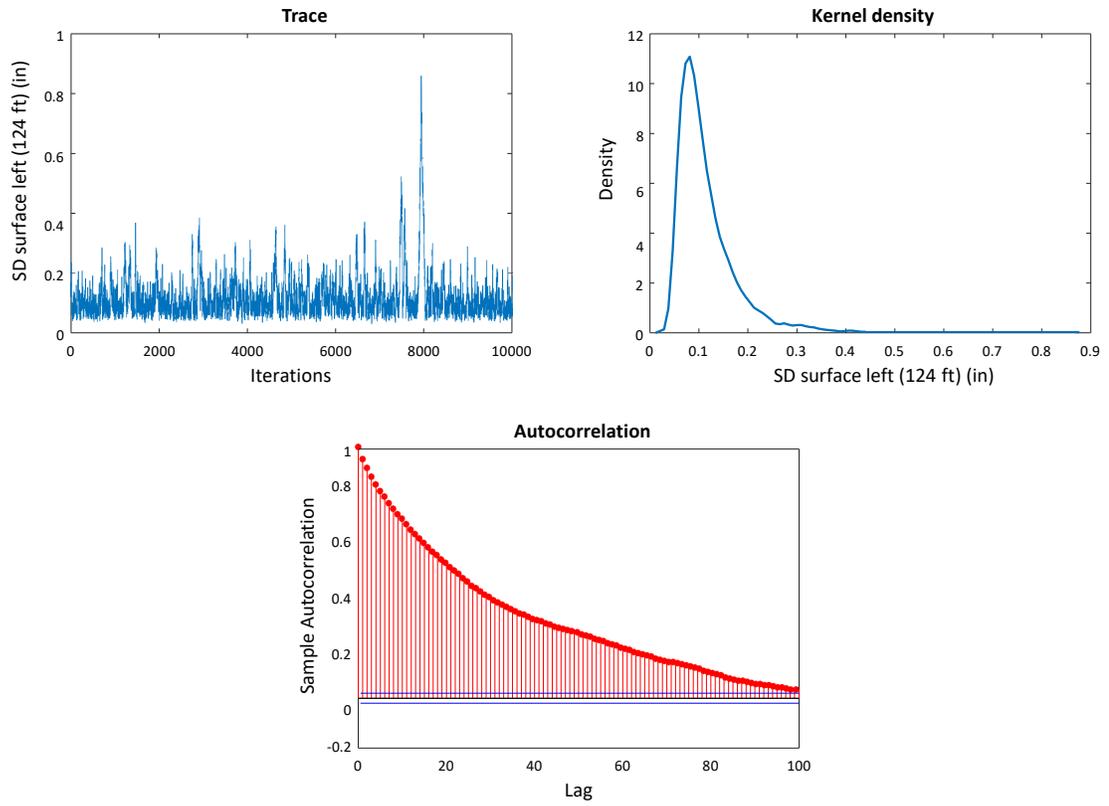


Figure 6.7: Surface left (124 ft): MCMC posterior plots for diffusion parameter

Table 6.2: Output summary for surface left (124 ft) track sections in the geocell zone

Section	Mean		Standard Monte Carlo		Mode		Median		MSE	
	μ	σ	error	μ	μ	σ	μ	σ	MCMC	MLE
1	0.02604	0.04737	0.01574	0.01382	0.02627	0.02913	0.02600	0.04462	0.03430	0.03446
2	-0.01086	0.04214	0.01884	0.01976	-0.01322	0.02427	-0.01093	0.03710	0.00423	0.00425
3	0.00647	0.01993	0.01084	0.01581	0.00518	0.00831	0.00639	0.01591	0.00073	0.00072
4	0.00105	0.01635	0.00372	0.00291	0.00246	0.01620	0.00103	0.01596	0.00089	0.00089
5	0.01429	0.02980	0.01349	0.01439	0.01249	0.01650	0.01445	0.02601	0.00302	0.00303
6	0.02739	0.05438	0.02494	0.02720	0.02536	0.03313	0.02761	0.04761	0.01334	0.01337
7	-0.00321	0.00941	0.00493	0.00617	-0.00262	0.00616	-0.00317	0.00768	0.00012	0.00013

Alignment Left (124 ft)

The output plots for this parameter show that the chain values for both drift and diffusion parameters reached the stationary distribution faster than surface left (124 ft). Also, the Kernel density has a symmetric shape for drift parameter and a skewed shape for diffusion parameter. In terms of the standard Monte Carlo error, the values are < 0.092 for drift parameter and < 0.011 for diffusion parameter. Finally, sections 1, 5, and 6 reported positive drift values.

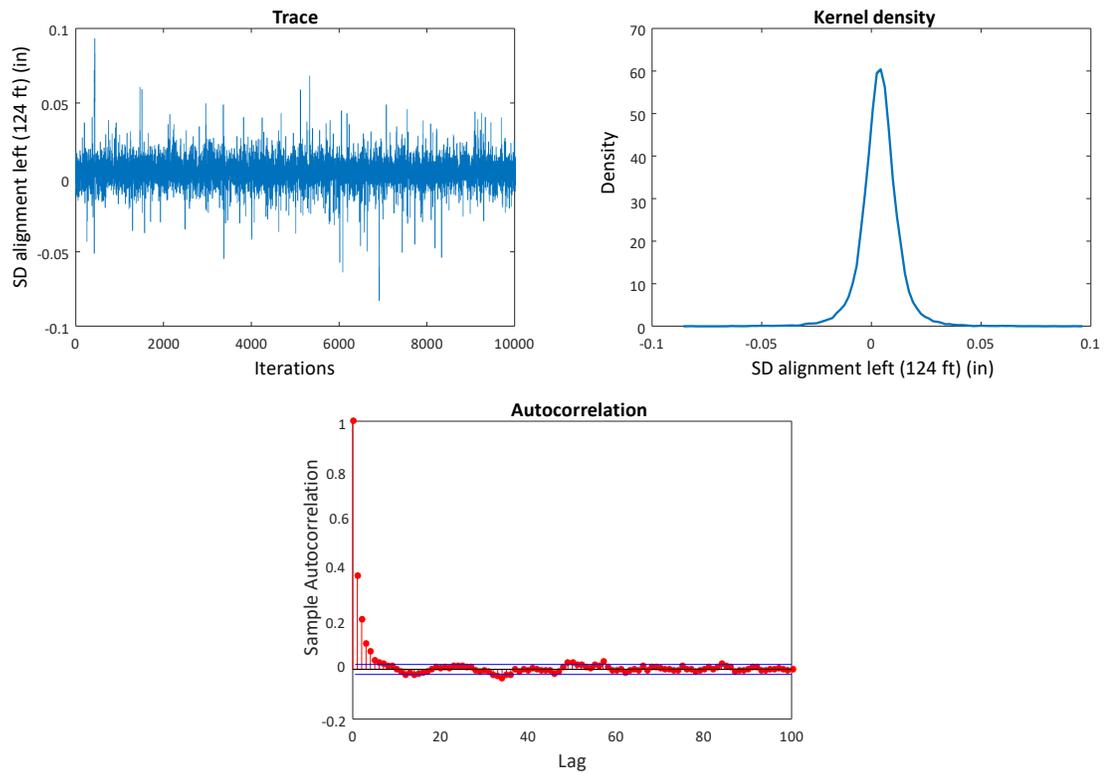


Figure 6.8: Alignment Left (124 ft): MCMC posterior plots for drift parameter

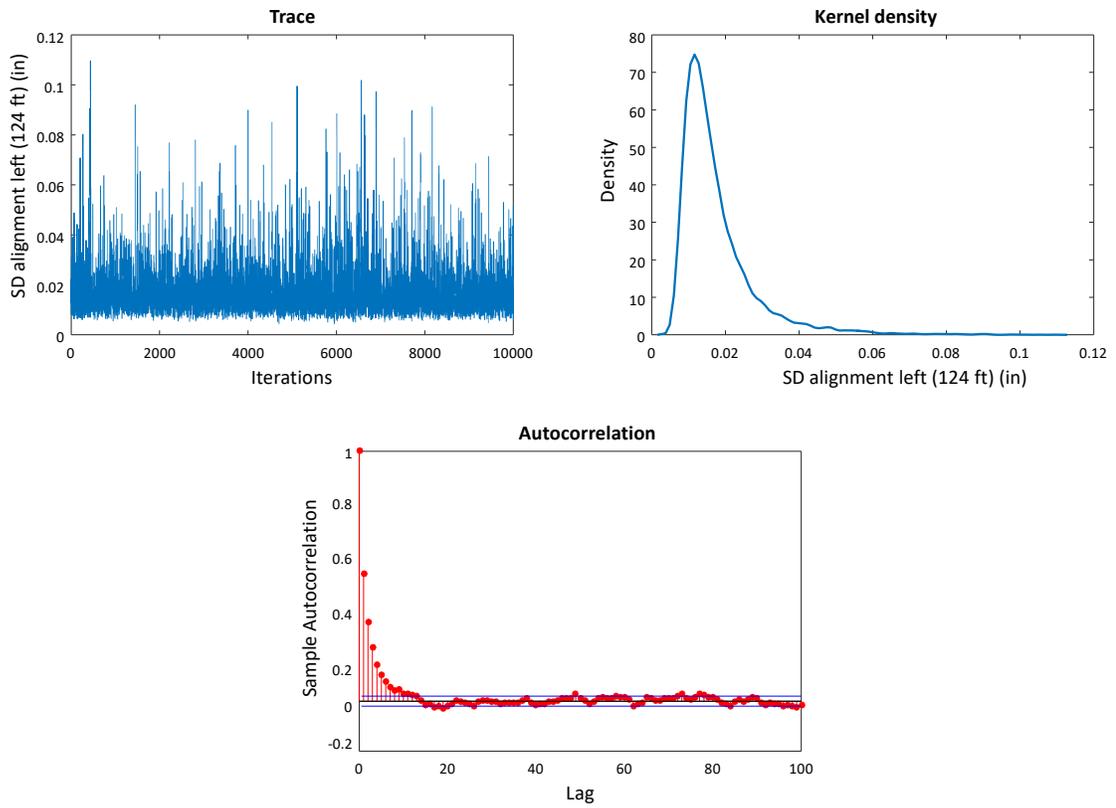


Figure 6.9: Alignment Left (124 ft): MCMC posterior plots for diffusion parameter

Table 6.3: Output summary for alignment left (124 ft) track sections in the geocell zone

Section	Mean		Standard Monte Carlo		Mode		Median		MSE	
	μ	σ	error	μ	μ	σ	μ	σ	MCMC	MLE
1	0.00367	0.00843	0.00281	0.00248	0.00279	0.00588	0.00361	0.00797	0.00067	0.00067
2	-0.00628	0.07061	0.03096	0.03348	-0.00410	0.03357	-0.00611	0.06208	0.00298	0.00299
3	-0.00281	0.01162	0.00609	0.00733	-0.00190	0.00687	-0.00280	0.00957	0.00013	0.00013
4	-0.00232	0.01095	0.00240	0.00192	-0.00115	0.01121	-0.00230	0.01065	0.00135	0.00136
5	-0.00032	0.01074	0.00474	0.00503	0.00091	0.00741	-0.00033	0.00949	0.00005	0.00005
6	-0.00301	0.01356	0.00613	0.00693	-0.00229	0.00920	-0.00300	0.01182	0.00007	0.00007
7	-0.00761	0.01951	0.01081	0.01389	-0.00656	0.01067	-0.00779	0.01589	0.00118	0.00119

Warp (62 ft)

The output analysis for plots in warp (62 ft) is similar to crosslevel in the sense that both reached the stationary distribution quicker than the rest of the parameters. Also, it is shown that the drift parameter has a symmetric Kernel density and the diffusion parameter has a skewed density. The Monte Carlo error is < 0.0092 for drift parameter and < 0.01 for diffusion parameter. Finally, sections 1, 5, and 6 reported positive drift values.

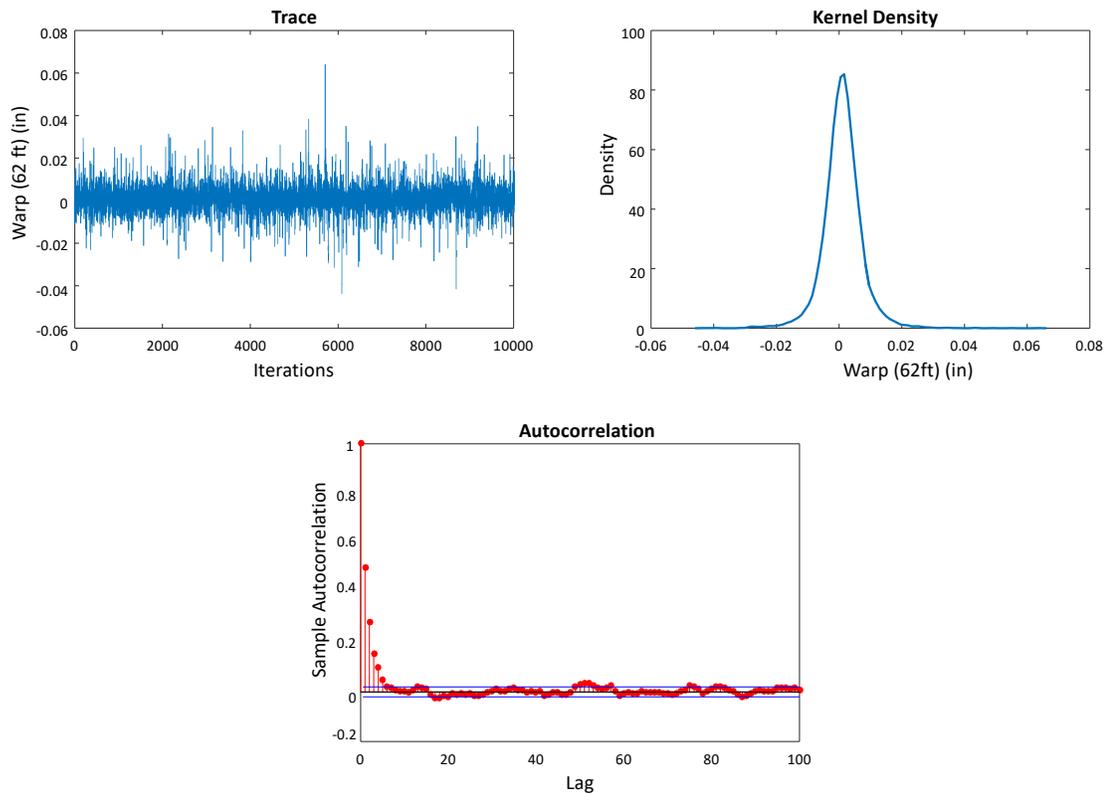


Figure 6.10: Warp (62 ft): MCMC posterior plots for drift parameter

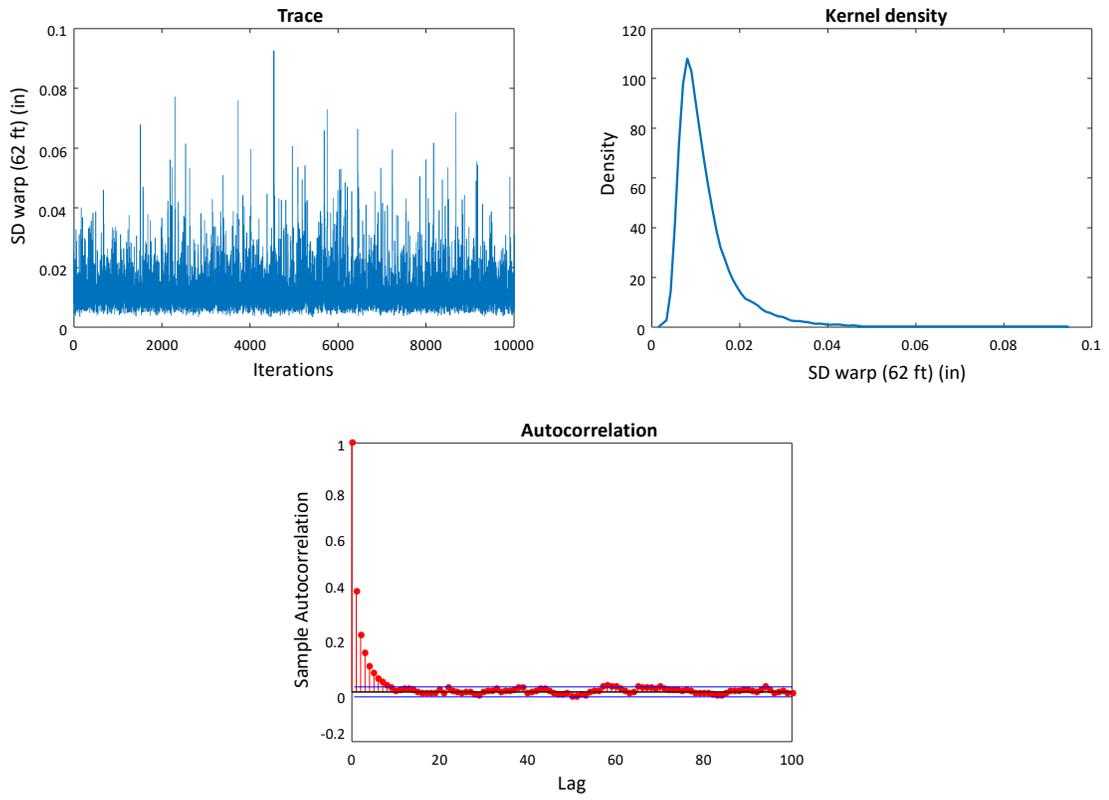


Figure 6.11: Warp (62 ft): MCMC posterior plots for diffusion parameter

Table 6.4: Output summary for warp (62 ft) track sections in the geocell zone

Section	Mean		Standard Monte Carlo		Mode		Median		MSE	
	μ	σ	error	μ	μ	σ	μ	σ	MCMC	MLE
1	0.00717	0.02472	0.00815	0.00689	0.01103	0.01925	0.00708	0.02340	0.00359	0.00357
2	-0.00886	0.02011	0.00911	0.01002	-0.00913	0.01295	-0.00885	0.01758	0.00091	0.00091
3	-0.00122	0.01120	0.00592	0.00796	-0.00185	0.00538	-0.00132	0.00922	0.00007	0.00006
4	-0.00044	0.00496	0.00110	0.00087	-0.00080	0.00473	-0.00044	0.00485	0.00009	0.00009
5	0.00319	0.01187	0.00534	0.00572	0.00186	0.00698	0.00316	0.01042	0.00017	0.00016
6	0.00110	0.01905	0.00851	0.00959	0.00240	0.01104	0.00111	0.01657	0.00031	0.00031
7	-0.00013	0.00278	0.00166	0.00209	-0.00052	0.00169	-0.00012	0.00222	0.00000	0.00000

6.4 Wiener Process Sample Paths

After validating the convergence of the Metropolis-Hastings algorithm for the Wiener process parameters, the Wiener process sample paths were simulated for each maintenance cycle per track section. The total number of steps (N) was 10,000, and the total number of sample paths was 1000. Figures 6.12-6.15 present the sample paths and the observed degradation for a specific track section considering the parameters crosslevel, surface left (124 ft), alignment left (124 ft), and warp (62 ft).

To illustrate how the sample paths contained the observed degradation values, two subplots are presented for each geometry parameter. The bigger subplot presents a scaled version of the observed and simulated degradation points.

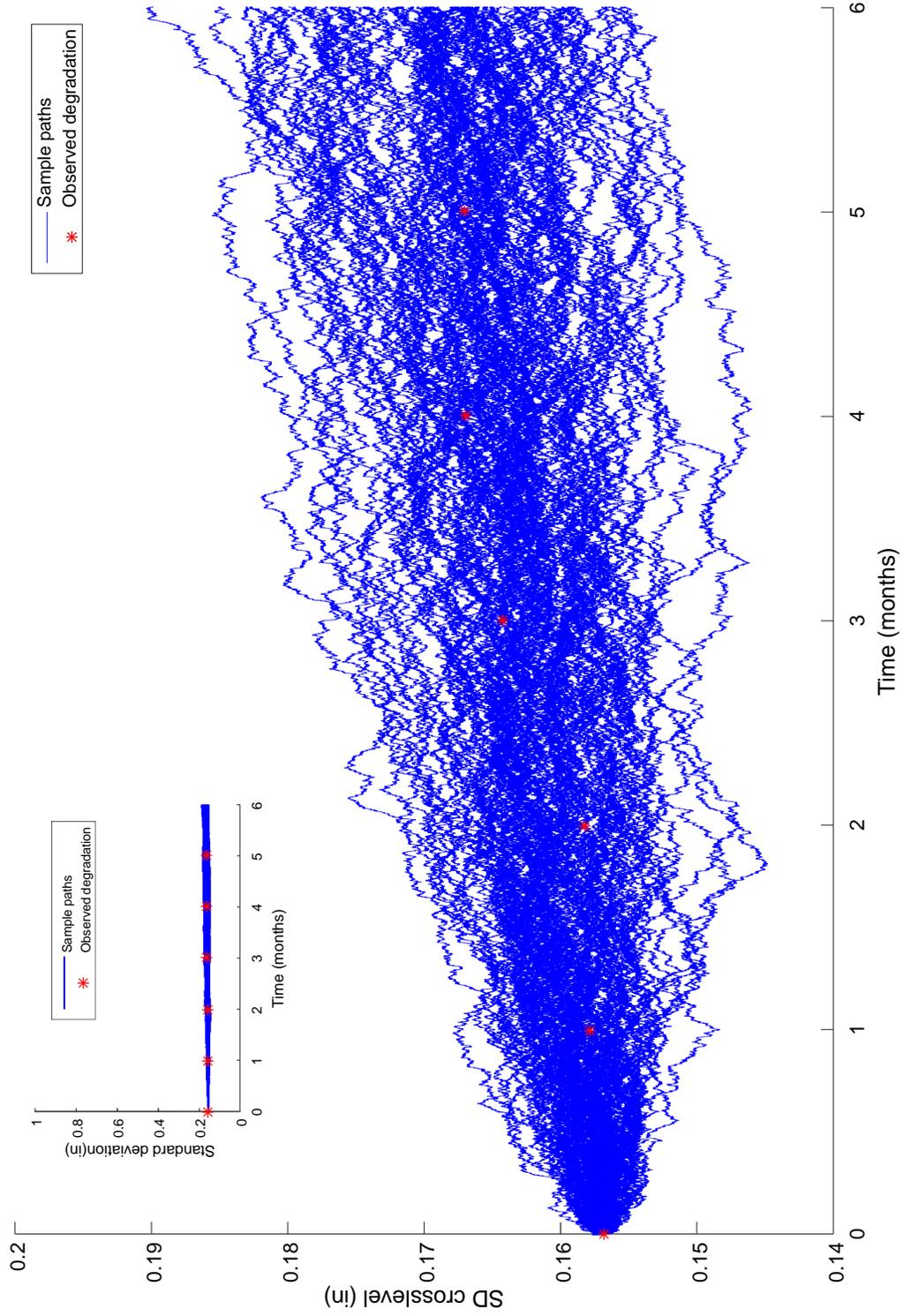


Figure 6.12: Sample paths and observed data for selected track section for crosslevel

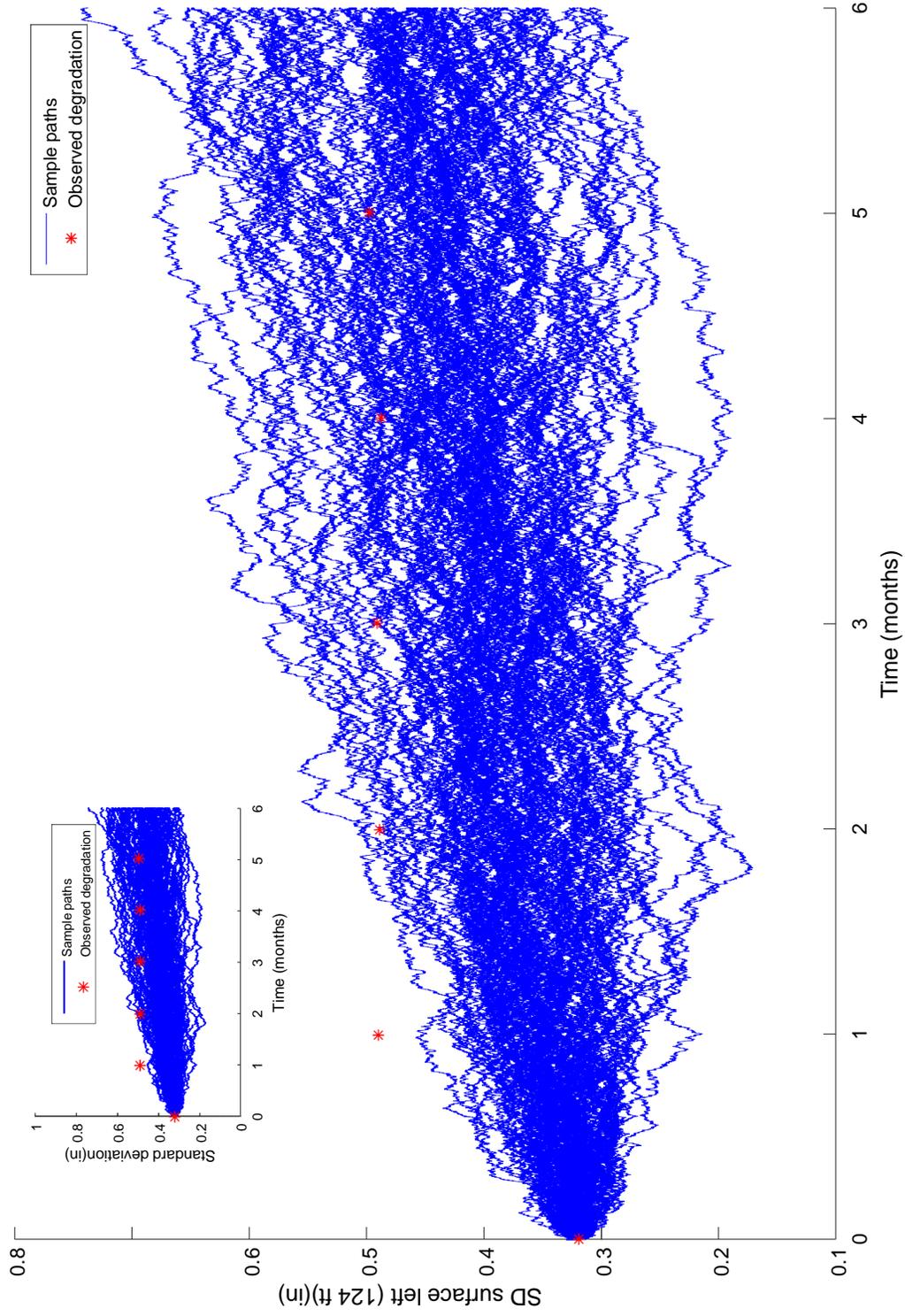


Figure 6.13: Sample paths and observed data for selected track section for standard deviation of surface left (124 ft)

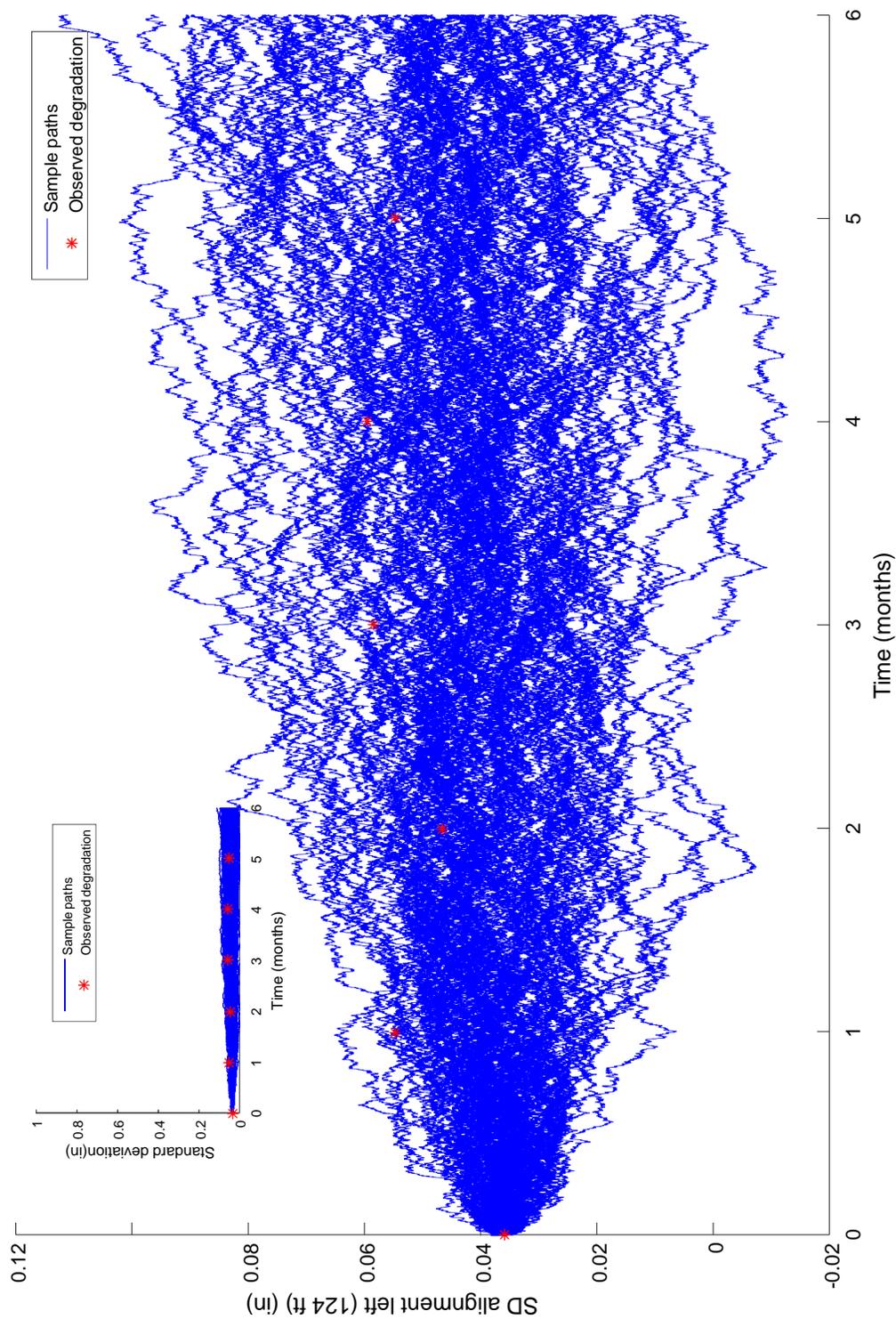


Figure 6.14: Sample paths and observed data for selected track section for standard deviation of alignment left (124 ft)

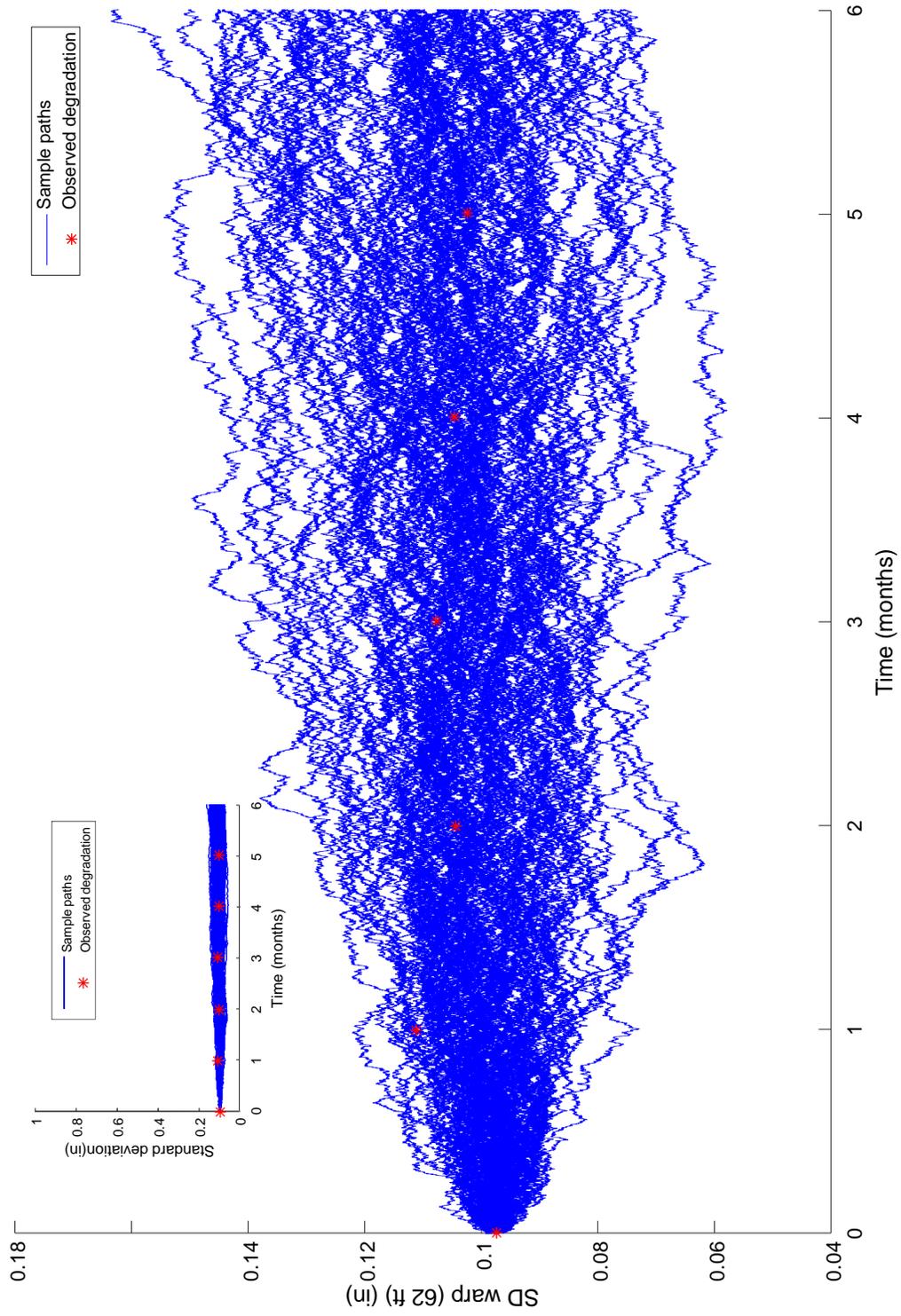


Figure 6.15: Sample paths and observed data for selected track section for standard deviation of warp (62 ft)

6.5 Model Validation

To validate the model convergence, data points for each maintenance cycle and for each track section were divided into two data sets: a training data set and a testing data set. The training data set consists of approximately 60% of the total data points and were utilized to estimate the Wiener process model parameters for each track geometry parameter. Once the parameters were estimated, the Wiener process sample paths were extrapolated from the last inspection date from the training set until the last reported inspection date for the maintenance cycle (last inspection date for the testing set) as presented in Figure 6.16.

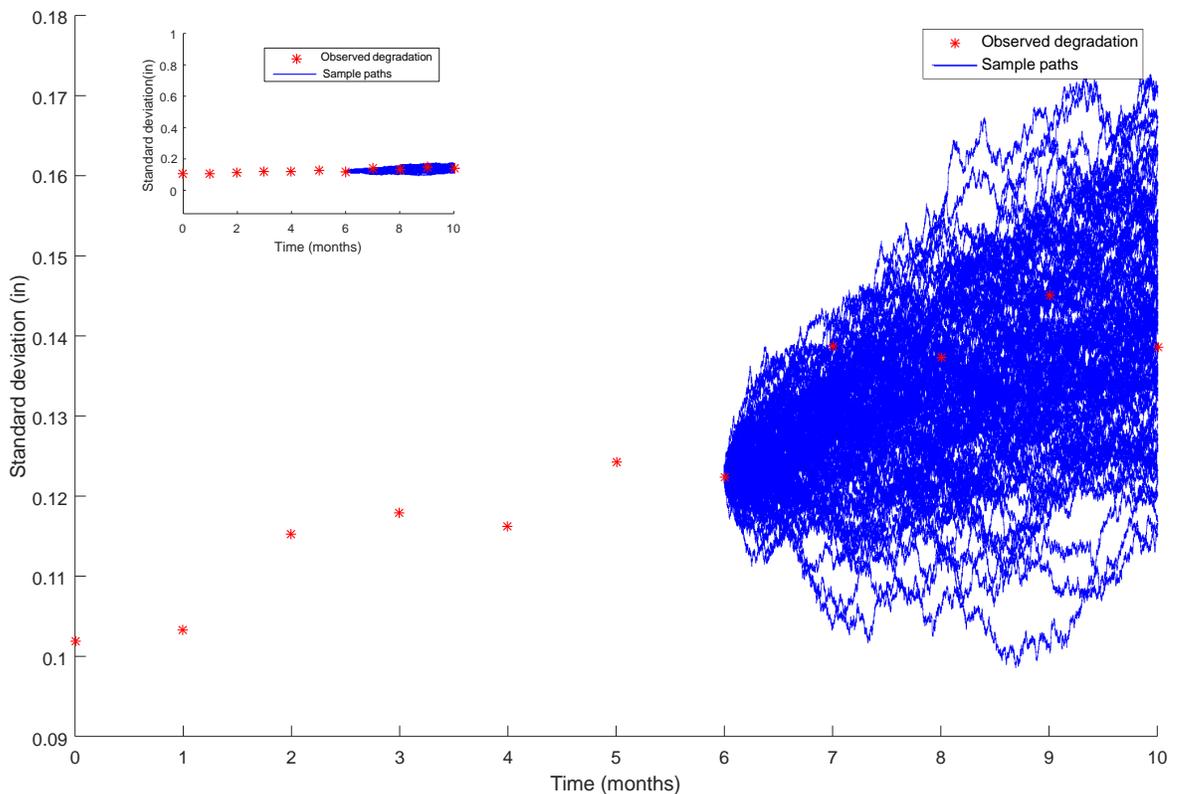


Figure 6.16: Illustration of training and testing data sets for a track section for parameter surface left (124 ft)

To measure the deviation of the observed data points from the predicted sample paths, the mean squared error (MSE) was calculated. Table 6.5 presents the MSE for 150-foot sections included in the geocell zone for parameters crosslevel, surface left

(124 ft), alignment left (124 ft), and warp (62 ft). As observed in the table, the MSE reported values near to zero, bounding the training data points. Results show how the Metropolis-Hastings algorithm is appropriate for estimation of the Wiener process model parameters when using a small number of training data points for all the experiments conducted.

Table 6.5: Mean squared error for testing data set

Section	Crosslevel	Surface left (124 ft)	Alignment left (124 ft)	Warp (62 ft)
1	7.32E-04	6.82E-04	5.26E-05	1.39E-03
2	4.02E-05	1.21E-02	5.79E-03	1.03E-03
3	1.26E-03	6.76E-04	2.09E-04	8.64E-04
4	8.05E-05	1.28E-04	1.30E-03	6.10E-05
5	3.34E-05	4.44E-04	1.23E-04	6.45E-05
6	4.83E-05	2.38E-03	3.01E-04	6.57E-05
7	9.60E-07	6.48E-04	3.11E-03	1.52E-05

6.6 Remarks

In this chapter, a hybrid Wiener process and Markov chain Monte Carlo (MCMC) for predicting track geometry degradation was proposed. A data set regarding track geometry inspection from a U.S. railroad was used. From this research the following conclusions were made:

1. From the data preprocessing, it was observed that degradation patterns present a non-strictly monotonic increasing behavior of track geometry degradation. Therefore, a stochastic approach to modeling track geometry degradation was appropriate.
2. The underlying degradation process was defined as a Wiener process with drift. The Wiener process parameters were estimated using an adaptive Metropolis-Hastings algorithm, where the lognormal distribution was updated during the simulation process. This approach allows movement from non-informative to informative prior as new data are available, allowing the prior to have an increasing influence over the posterior distribution as the number of iterations increases.
3. The Metropolis-Hastings output analysis was presented in order to analyze the algorithm convergence. The analysis shows that the generated Markov chains

successfully reached the steady state distribution and the posterior distribution of the parameters, given track geometry inspection data. This is concluded based on the low values of the standard Monte Carlo error, as well as the autocorrelation plots for both drift and diffusion parameters.

From the Kernel densities it is shown that drift parameter is represented by a symmetric distribution and a right skewed distribution represented the diffusion parameter for all track geometry parameters under study. In addition, the number of sections that reported a positive drift is equal to 3 for parameters crosslevel, alignment left (124 ft), and warp (62 ft). However, parameter surface left (124 ft) reported 5 sections with positive drift values. This is also a confirmatory analysis on the influence of surface parameter in track degradation compared to the rest of parameters. Based on the results presented, Bayesian analysis-based Markov chain Monte Carlo seems to have potential for handling complex degradation models. The results suggest that track maintenance interventions can be planned more accurately using this methodology.

4. The Wiener process was simulated in order to predict the track geometry degradation. From the results it was observed that the simulated 1000 paths bounded the observed degradation points. Parameter surface left (124 ft) reported sample paths that were more spread compared to the rest of the parameters. This behavior is expected since the exploratory data analysis showed that variations of surface parameter have higher variability.
5. To validate the Wiener process model, data points for each track section and each maintenance cycle were divided into training and testing data set. The mean squared error (MSE) was considered to measure deviations from the predicted sample paths and the test data set, in which small MSE values were obtained. These results show that the predicted sample paths can account for the uncertainty of the degradation data which can be used for maintenance scheduling purposes.

REFERENCES

- G. O. Roberts, A. Gelman, and W. R. Gilks. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1): 110–120, feb 1997. doi: 10.1214/aoap/1034625254.

Chapter 7

FIRST HITTING TIME IN RAILWAY TRACK GEOMETRY DEGRADATION ANALYSIS

7.1 Introduction

This chapter estimates the first hitting time for the first maintenance cycle of track geometry degradation sample paths. The underlying degradation paths are modeled as a Wiener process with drift and the FHT follows an inverse Gaussian distribution. The results provide a more robust representation of track geometry failure time using degradation data. The methodology followed to estimate the FHT is presented in Figure 7.1.

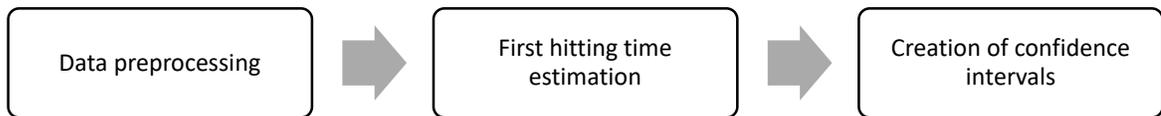


Figure 7.1: General overview for the estimation of the first hitting time for a single track geometry parameter

7.2 Data Preprocessing

In this case study two data sets were utilized. The first data set is described in Chapter 5 and the mile of track was divided into 150-foot and 500-foot sections, and their corresponding TQI based on the standard deviation was calculated as presented in Chapter 6. The second data set utilized consists of degradation data based on foot-by-foot measurements for each track geometry parameter. The input data used in this case study, therefore, correspond to the simulated sample paths from the Wiener process with drift for the first maintenance cycle and for each track geometry parameter. The schematic representation of the estimation of the FHT is presented in Figure 7.2.

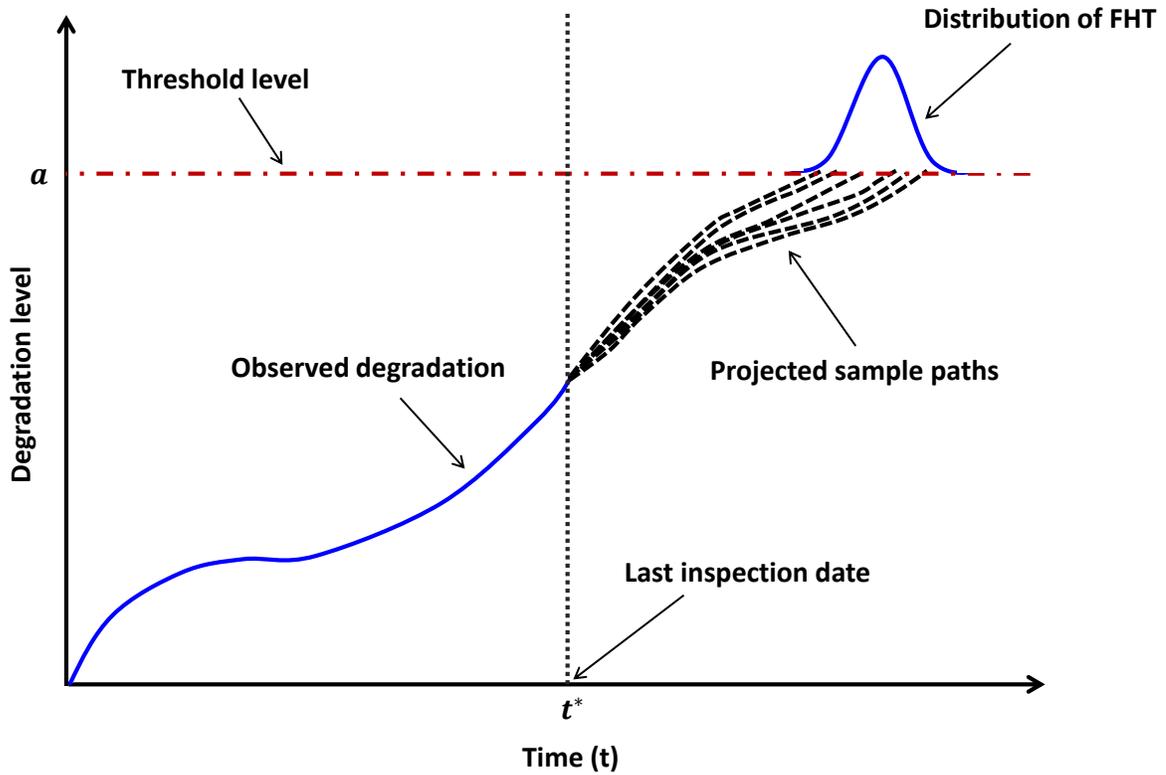


Figure 7.2: Schematic representation of the threshold-regression model

7.3 Estimation of the FHT for TQI

The probability density function (PDF) and the cumulative density function (CDF) of the FHT were estimated for various degradation patterns in their first maintenance cycle for each track geometry parameter. In this research, an arbitrary threshold of 0.4 inches was established and the predicted sample paths from the Wiener process with drift were extrapolated with a maximum number of months equal to 1000. This number was also arbitrary but it is large enough to determine which sections for each parameter reached the threshold, considering that tamping activities are carried out in a time frame less than 20 months. As presented in Chapter 4, the FHT for a Wiener process is the inverse Gaussian distribution.

To illustrate the comparison of the FHT at various parameters, a track section of 150 feet is presented. This section was chosen because it was the one that a FHT was able to estimate for all the track geometry parameters. Figures 7.3-7.6 show the

FHT for parameters crosslevel, surface left (124 ft), alignment left (124 ft), and warp (62 ft) respectively.

The PDF was estimated using the nonparametric approach Kernel density which plots the density of the time at which each sample path reached the threshold (continuous line in Figures 7.3-7.6). A theoretical PDF was also calculated and plotted (dashed line) in Figures 7.3-7.6 using the inverse Gaussian distribution. The scale parameter was equal to the drift parameter estimated for the Wiener process for each track geometry parameter and section. The shape parameter was a transformation of the diffusion parameter of the Wiener process.

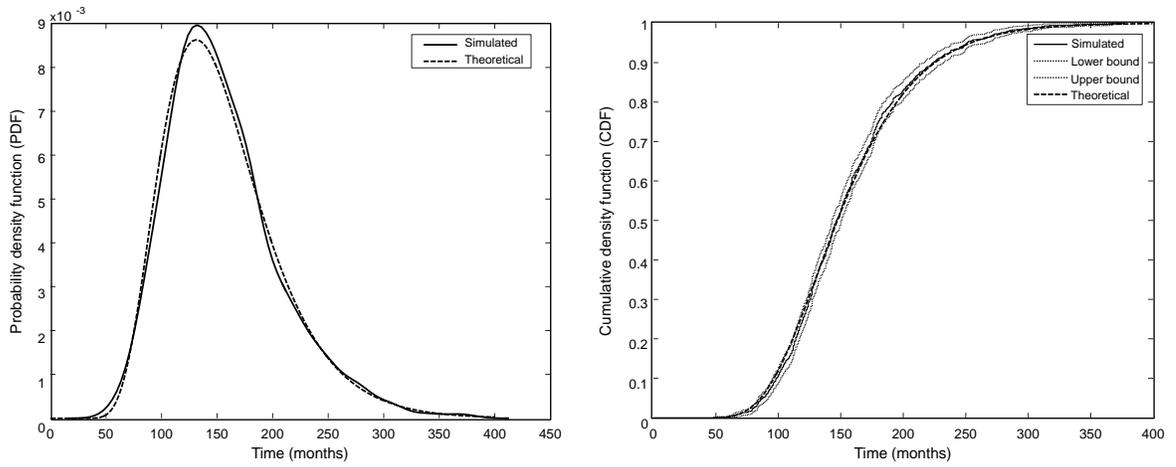


Figure 7.3: PDF and CDF of FHT for crosslevel

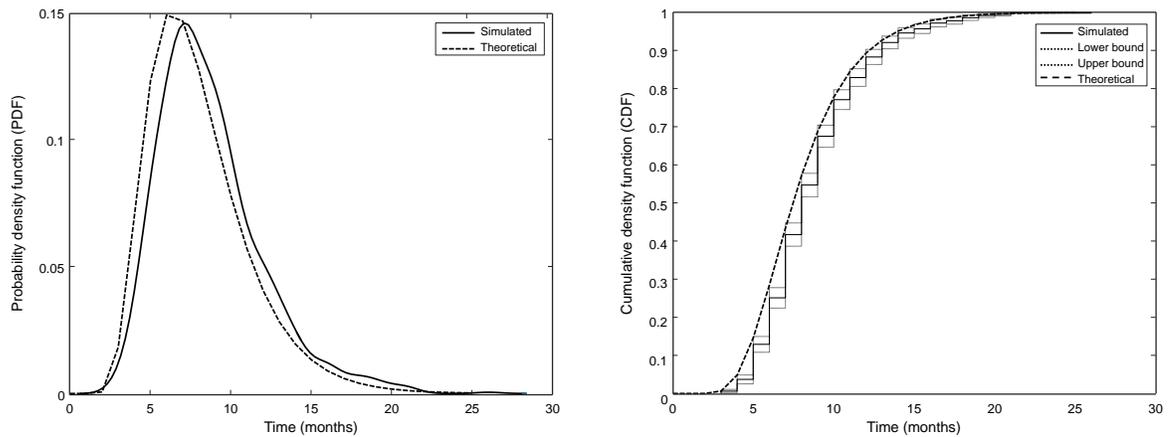


Figure 7.4: PDF and CDF of FHT for surface left (124 ft)

As shown in Figures 7.3-7.6, both simulated and theoretical PDF and CDF are similar, so the simulated sample paths follow the inverse Gaussian distribution since they are projected as a Wiener process. On the other hand, crosslevel and alignment right (124 ft) reported high values of the scale parameter of the inverse Gaussian distribution, that is, the specific section of track would take longer to reach the defined threshold compared to the rest of the parameters if the track continues degrading as a Wiener process.

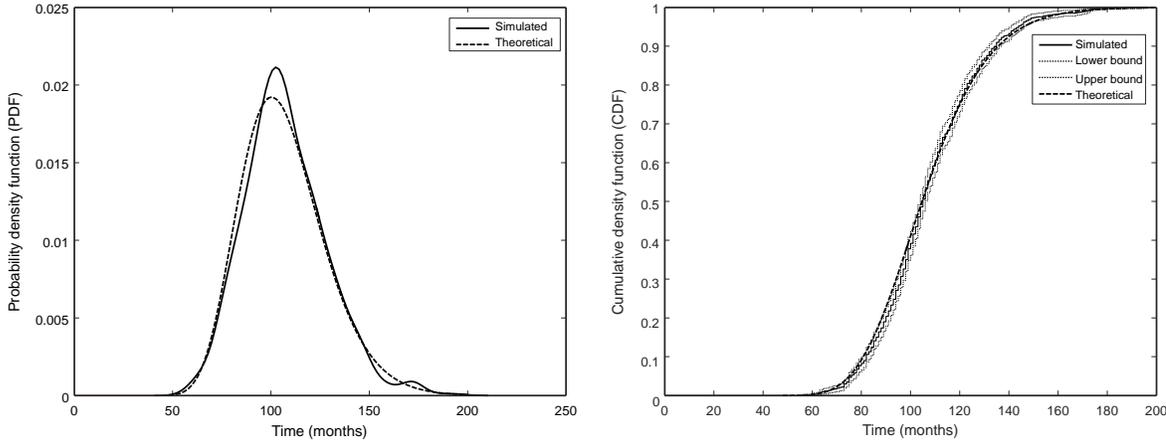


Figure 7.5: PDF and CDF of FHT for alignment left (124 ft)

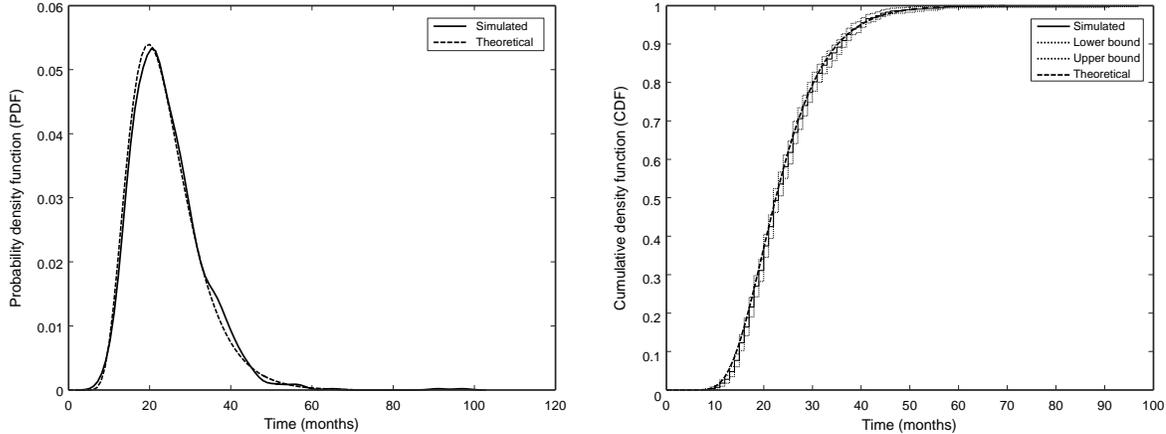


Figure 7.6: PDF and CDF of FHT for warp (62 ft)

Table 7.1 presents the skewness and kurtosis for the FHT of each track geometry parameter. It can be observed that for all track geometry parameters, the probability

density function is positive, so the PDF is skewed to the right. Also, the kurtosis shows how heavy tailed or outlier prone the FHT is. For all the track geometry parameters the kurtosis measure shows that the FHT is heavy tailed considering that the kurtosis for a Gaussian distribution is equal to 3.

Table 7.1: Skewness and kurtosis for FHT

Parameter	Section	Skewness	Kurtosis
Crosslevel	1	0.9488	4.2899
	5	0.4315	3.5525
	13	0.4662	3.7484
Surface left (124 ft)	1	1.1752	4.9987
	3	0.8736	4.2463
	4	1.6094	6.2097
	5	1.0117	6.1094
	6	2.0419	9.2104
Alignment left (124 ft)	1	0.5941	3.7454
Warp (62 ft)	1	1.7456	11.1826
	5	0.7746	3.8961

7.3.1 Confidence Limits for FHT

In addition to the estimation of the PDF and CDF of the FHT for each section and track geometry parameter, the FHT is also expressed in terms of confidence intervals. For this purpose, the Wald test suitable for constructing confidence limits for inverse Gaussian densities was utilized. In this research, three different confidence levels were defined as presented in Table 7.2.

Table 7.2: Illustration of the confidence intervals for the FHT for surface data

Section	99% confidence interval		95% confidence interval		90% confidence interval	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
1	21.6	22.6	21.7	22.5	21.8	22.4
2	296.5	308.1	297.9	306.7	298.6	306.0
3	58.7	61.9	59.1	61.5	59.3	61.4
4	57.6	62.1	58.2	61.6	58.5	61.3
5	16.7	17.3	16.7	17.3	16.8	17.2
6	72.0	77.5	72.7	76.8	73.0	76.5
7	76.7	80.2	77.1	79.7	77.3	79.5
8	82.4	89.7	83.2	88.8	83.7	88.4
9	116.9	122.5	117.6	121.8	117.9	121.5
10	73.7	75.1	73.9	75.0	74.0	74.9
11	4.7	5.2	4.8	5.2	4.8	5.1

7.4 FHT for Raw Data

In addition to estimate the FHT using the standard deviation of a 150-foot and 500-foot track sections and defining an arbitrary threshold, the FHT was also estimated considering FRA safety standards. Since these standards can only be compared with actual track geometry values, raw data was utilized in this application. Figure 7.7 presents the FHT for four adjacent locations for parameter surface left (62 ft). It can be observed from the figure that adjacent locations have the same distribution of failure time. Therefore, this approach can be used to predict the time at which both TQI and raw data reach a maintenance threshold and for predicting the time until track geometry parameters reach a safety threshold.

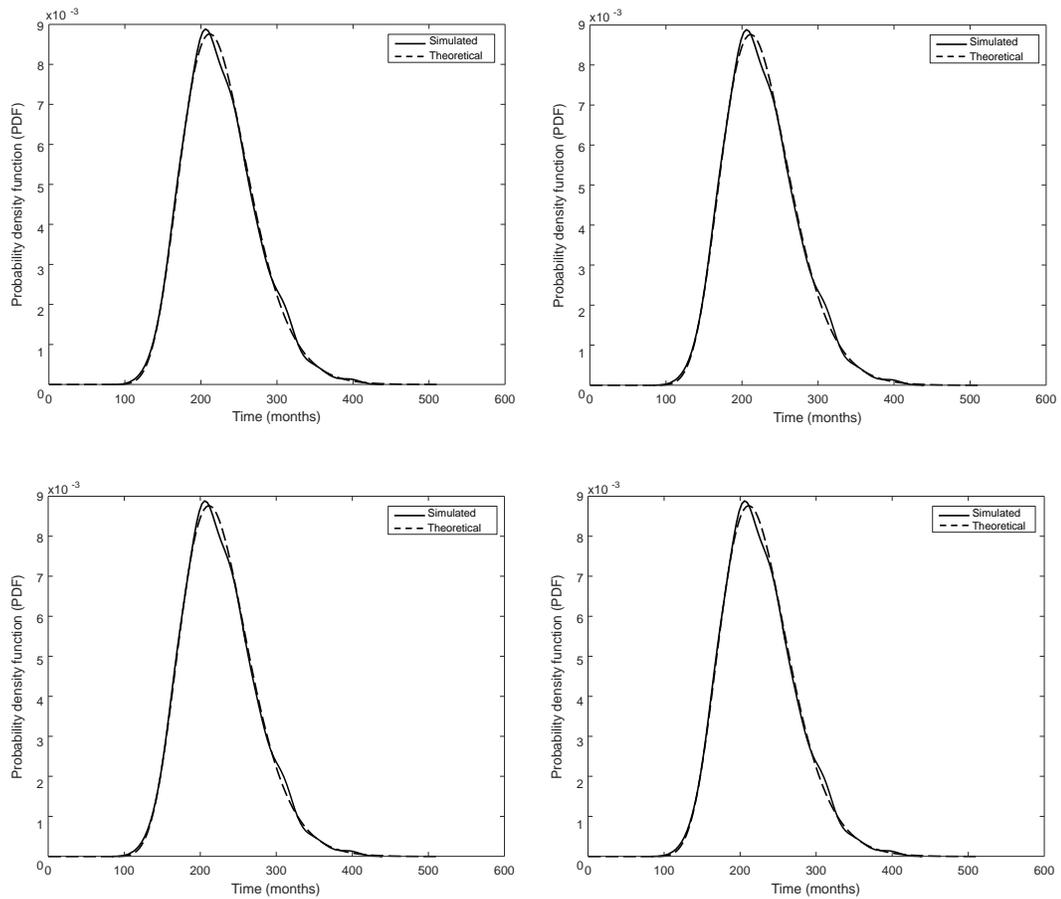


Figure 7.7: Theoretical and simulated FHT at four adjacent locations for surface left (62 ft)

7.5 Remarks

This chapter presented the formulation and implementation of the first hitting time (FHT) model in the context of track geometry degradation. For this case, analysis for various track geometry parameters were considered.

1. The first hitting time was estimated using two approaches: (i) drawing the Kernel density from the predicted sample paths and (ii) obtaining the theoretical density using an inverse Gaussian distribution. From the two densities it can be observed that both simulated and theoretical FHT are close to each other.
2. The estimation of the FHT for different track geometry parameters and different track sections allows observation of dominant parameter for specific sections. That is, given a specific threshold, it can be observed which geometry parameter first reaches the threshold and which locations of the track requires more attention. This would be useful as an alternative approach to allocate budget for maintenance purposes at different confidence levels.
3. The FHT was estimated using foot-by-foot measurements and the threshold level were established based on FRA safety standards. It was observed that adjacent track locations have the same FHT. These results show potential to use degradation data based on foot-by-foot measurements for the prediction of track geometry defects.

Chapter 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 Introduction

This dissertation presented an in-depth analysis of track geometry data for a U.S. Class I railroad. Extensive literature review and various quantitative techniques were re-evaluated. The merits of traditional techniques which have been previously used in the literature were discussed. The current work formulates the degradation process using Wiener process and develops first hitting time principles as a method to determine the time for maintenance of rail track based on time history.

8.2 Conclusions

1. This research conducted a literature review on track geometry degradation models. This review provided a better understanding of the state-of-the-art regarding the current practices for analyzing track degradation. Also, research gaps were identified, such as the use of Bayesian approaches to estimate track geometry degradation, the use of stochastic processes to predict not only the degradation, but to estimate the first hitting time. This research attempted to bridge those gaps by providing a more robust framework for predicting track geometry degradation model parameters and exploiting the advantages of stochastic processes for estimating the first hitting time.
2. This research conducted an exploratory data analysis (EDA) to provide initial insights of the data set available. The EDA results showed that variations of parameter surface appeared to be the parameter with the highest variability. Also, it was observed from the box plots using cross-sectional data that adjacent locations of the track have similar characteristics. This is an important conclusion because the definition of the track section lengths can be increased without losing information. The results of the EDA were used as a confirmatory data analysis tool for the case studies performed. That is, explanations regarding the high variability of the model parameters, degradation paths, and first hitting times were identified based on the previous knowledge from the EDA.

3. A case study regarding the hybrid Bayesian inference and Wiener process with drift was conducted. As presented in Chapter 3, there are contributions in the literature for using Bayesian inference based on Markov chain Monte Carlo (MCMC) to estimate track degradation model parameters. This research improved those studies by:
 - (a) Formulating a variation of the Metropolis-Hastings algorithm as an adaptive MCMC, in which the prior distribution, initially defined as noninformative, became informative as the number of iterations increased. This was done by including to the prior distribution parameters the current accepted values of the Markov chain. This is relevant because it allows incorporation of noninformative priors when there is not enough knowledge of the track degradation model parameters.
 - (b) Performing a detailed MCMC output analysis. Literature in track geometry degradation have not addressed the convergence of MCMC algorithms which is relevant in analyzing the model parameters. Graphical methods such as Kernel density, trace, and autocorrelation plots were presented to illustrate the independence of the samples over the iterations, as well as to indicate the shape of the posterior distribution. Also, point estimates and standard Monte Carlo errors were calculated, which allowed identification of track geometry parameters with positive drift.

The Wiener process sample paths were simulated using the output from the adaptive Metropolis-Hastings algorithm. It was observed that the predicted sample paths were able to capture the variability of the degradation process by bounding the observed degradation data points.

4. A second case study was performed to estimate the first hitting time (FHT) for the first maintenance cycle for each track geometry parameter. The case study presented the conducted methodology in which three main stages were defined. These stages include: (i) data preprocessing, (ii) first hitting time estimation, and (iii) creation of confidence intervals.

The FHT was estimated using two approaches:

- (a) Drawing the Kernel density by extrapolating the predicted sample paths obtained in the previous case study, and collecting the times at which each sample path reached the threshold
- (b) Drawing the FHT by using an inverse Gaussian distribution which is the analytic expression of the FHT in the Wiener process

Both densities were plotted together and it was observed that they were close to each other, so the simulated sample paths were able to predict the FHT. Also, it was concluded that parameter surface left measured in a 62-foot and 124-foot

chord were the dominant parameters compared to the remaining parameters, which required less time to reach the threshold.

5. From a policy making perspective, the results of this research provide information regarding the track sections that require more attention and can be used as an input for maintenance scheduling activities.

8.3 Recommendations

This dissertation provided major contributions for improving data analysis in railway track geometry degradation. However, there are opportunities for improvement as presented below.

1. The exploratory data analysis conducted in this research considered longitudinal and cross-sectional data. Further analysis can include panel data in which groups of locations of the track are observed over time.
2. In this research, the Metropolis-Hastings algorithm was utilized to estimate the Wiener process model parameters. Possible extension of this approach include:
 - (a) The use of population-based Markov chain Monte Carlo methods, in which metaheuristic methods such as genetic algorithms, particle swarm optimization, and ant colony optimization, among others can explore regions in the search space by exchanging information between Markov chains.
 - (b) The formulation and implementation of likelihood-free methods. The approximate Bayesian computation (ABC) is one of these methods and according to the literature, is suitable because it is not required to make assumptions about the likelihood function.
3. In terms of track geometry degradation, this research considered the Wiener process with drift, which was able to predict the degradation values for individual track geometry parameters. Extensions of this approach may include the use of an adaptive drift function. That is, the drift parameter can be recursively updated as new degradation points are available. Although this research predicted the degradation paths for each track geometry parameter, future work can include the formulation of the Wiener process with multiple hazards, in which more than one track geometry parameter can be incorporated in the degradation model. Also, the threshold level for estimating the first hitting time can be expressed as a random variable.
4. This research considered the standard deviation of individual track geometry parameters for a track section as a track quality index (TQI), and from there, perform the hybrid Bayesian estimation and Wiener process for predicting track

degradation and failure. Traditionally, failure time has been addressed using track geometry defect data instead of foot-by-foot measurements. Understanding that TQI values cannot be compared directly to FRA safety standards, massive amounts of foot-by-foot track geometry data can be used to predict the failure time. Therefore, it is recommended to develop innovative approaches to use degradation data that can be utilized, for example, to estimate track failure time.

Appendix A
EXPLORATORY DATA ANALYSIS PLOTS

A.1 Foot-By-Foot Measurements

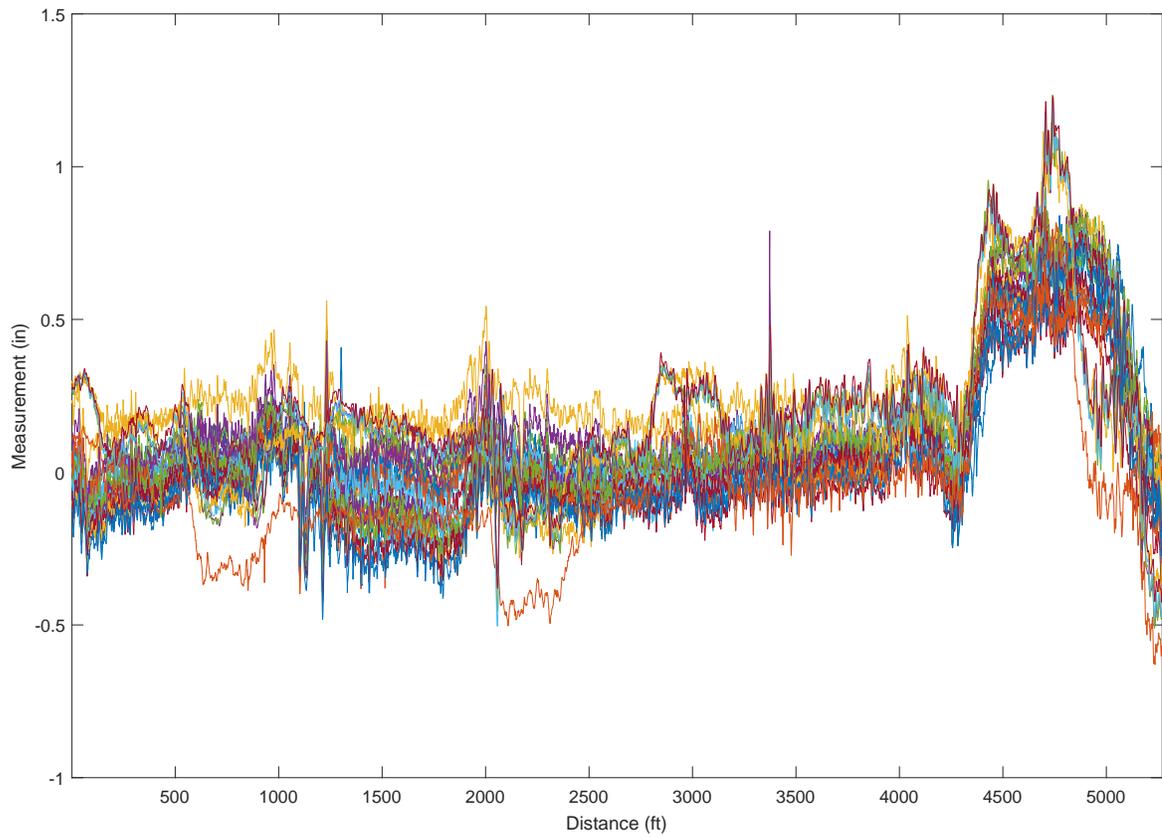


Figure A.1: Illustration of crosslevel data at multiple inspection dates

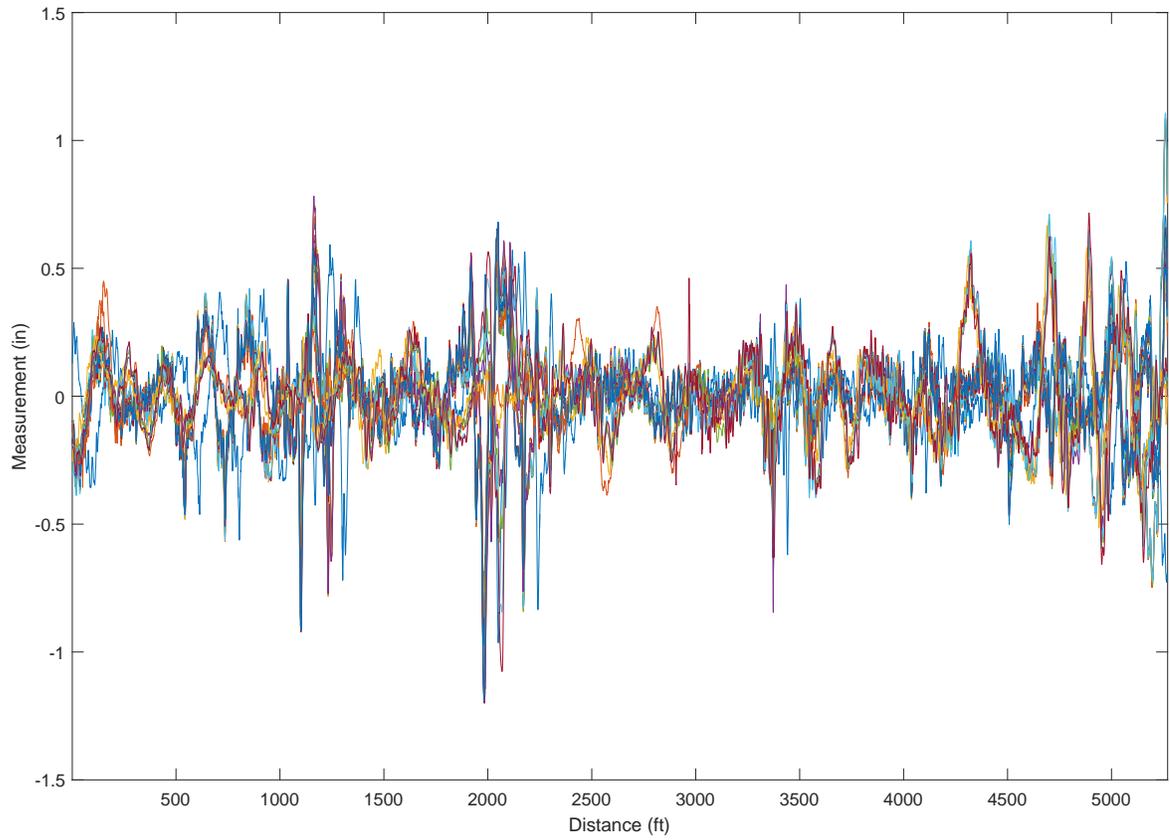


Figure A.2: Illustration of surface right (124 ft) data at multiple inspection dates

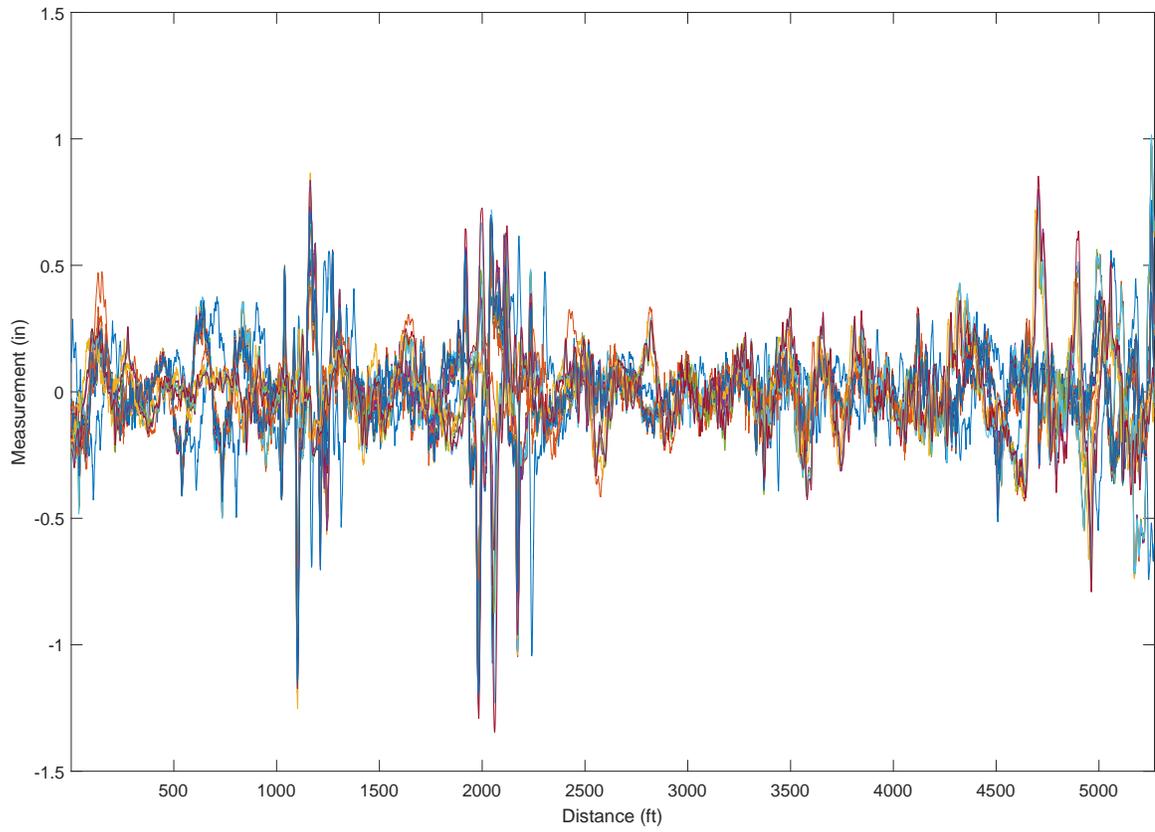


Figure A.3: Illustration of surface left (124 ft) data at multiple inspection dates

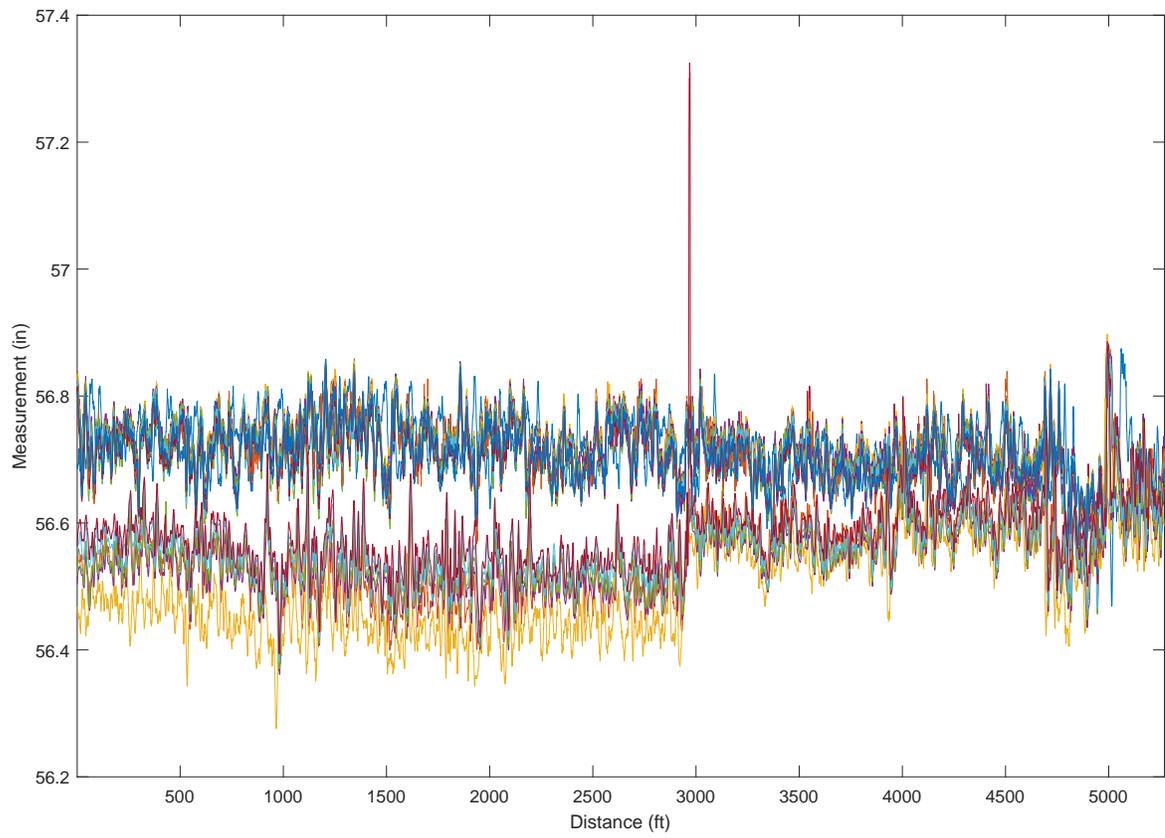


Figure A.4: Illustration of gage data at multiple inspection dates

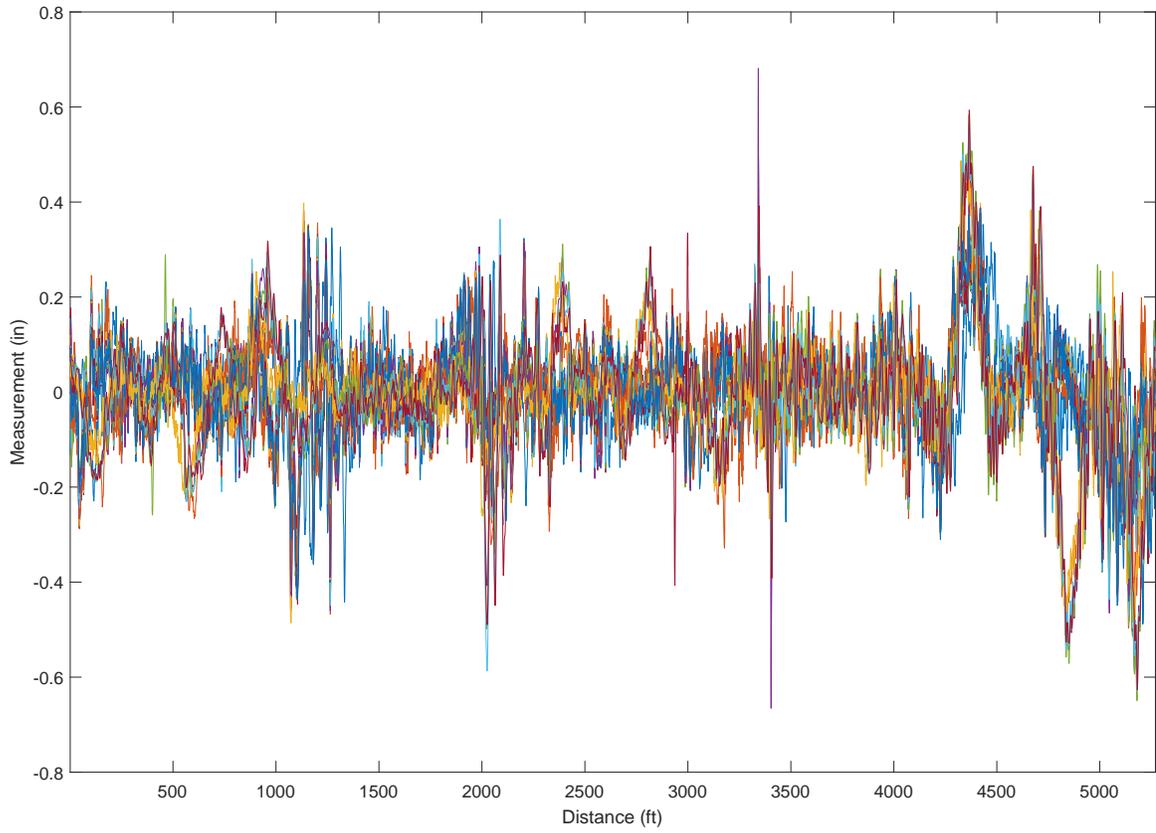


Figure A.5: Illustration of warp (62 ft) data at multiple inspection dates

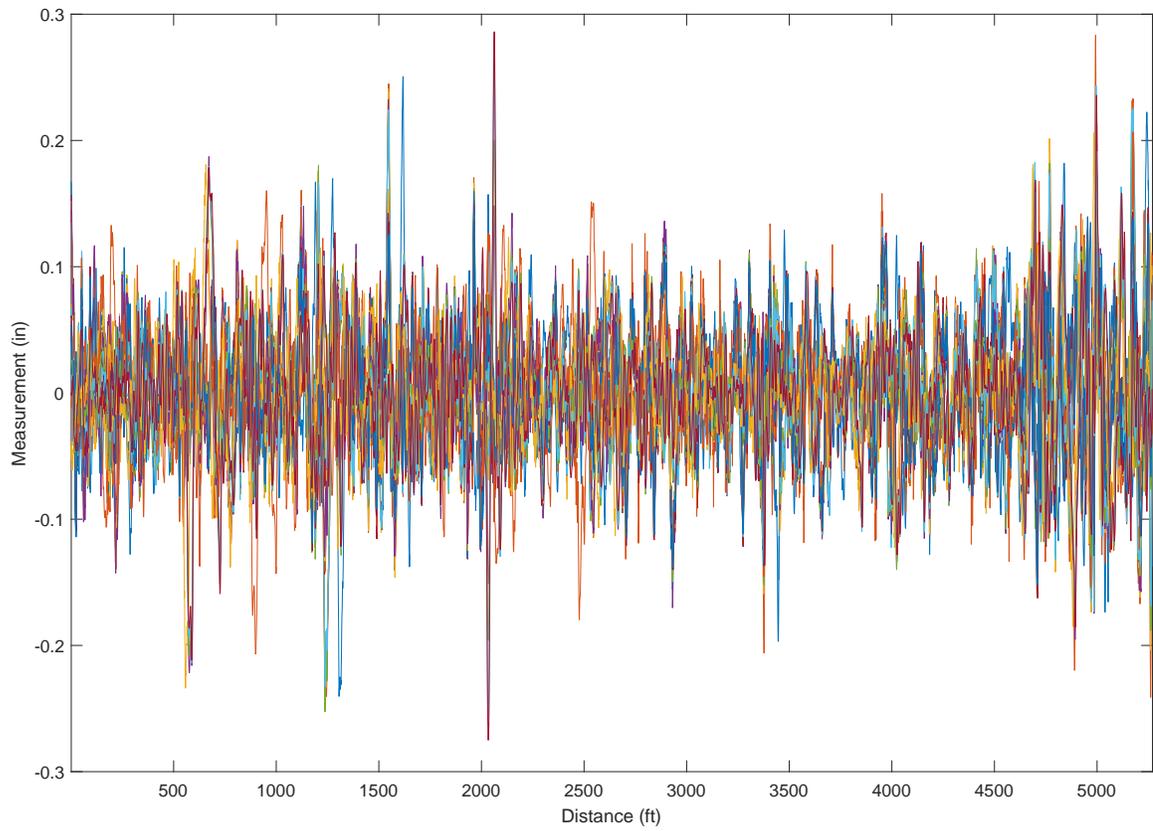


Figure A.6: Illustration of alignment left (62 ft) data at multiple inspection dates

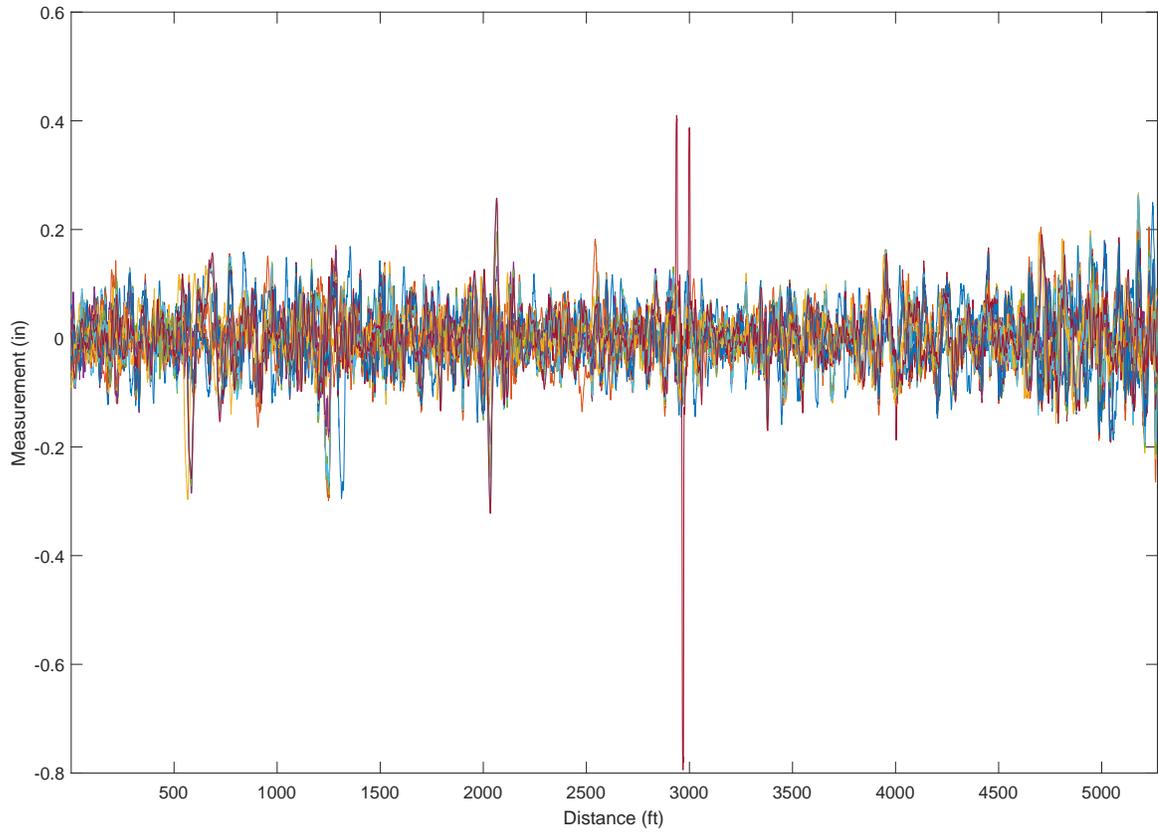


Figure A.7: Illustration of alignment right (62 ft) data at multiple inspection dates

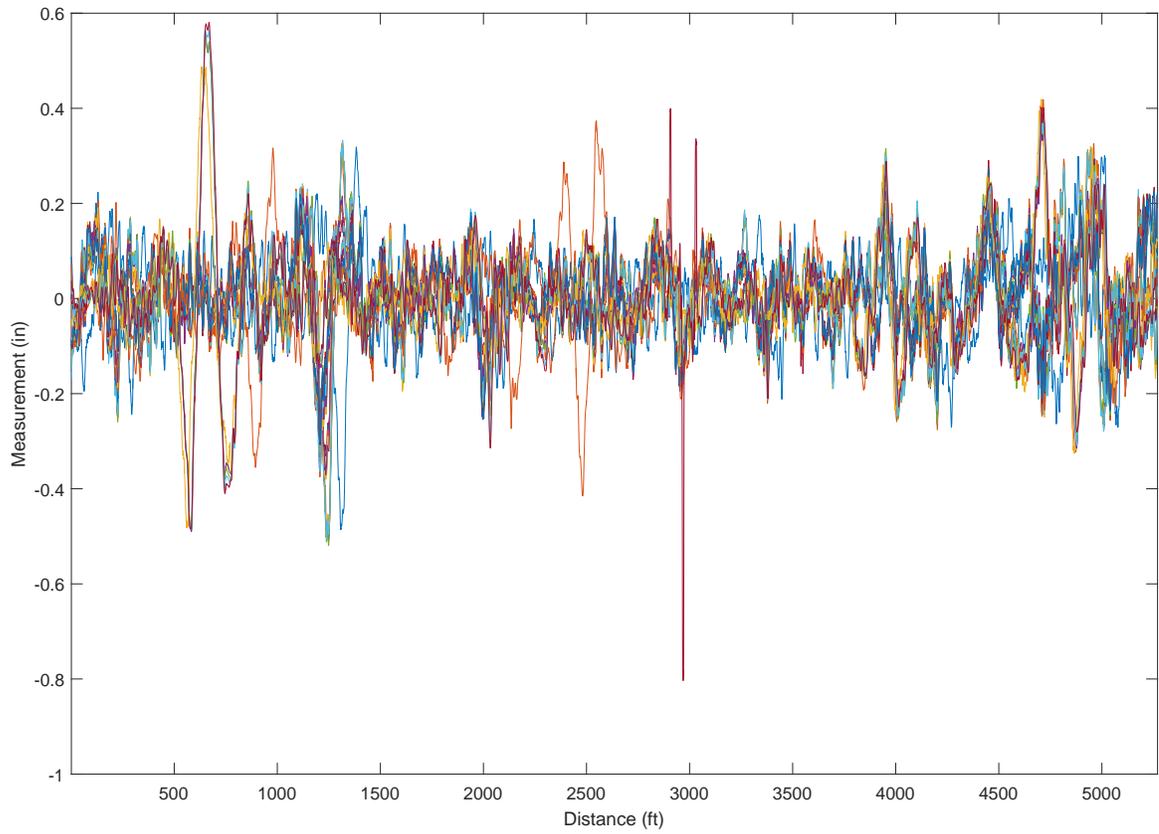


Figure A.8: Illustration of alignment right (124 ft) data at multiple inspection dates

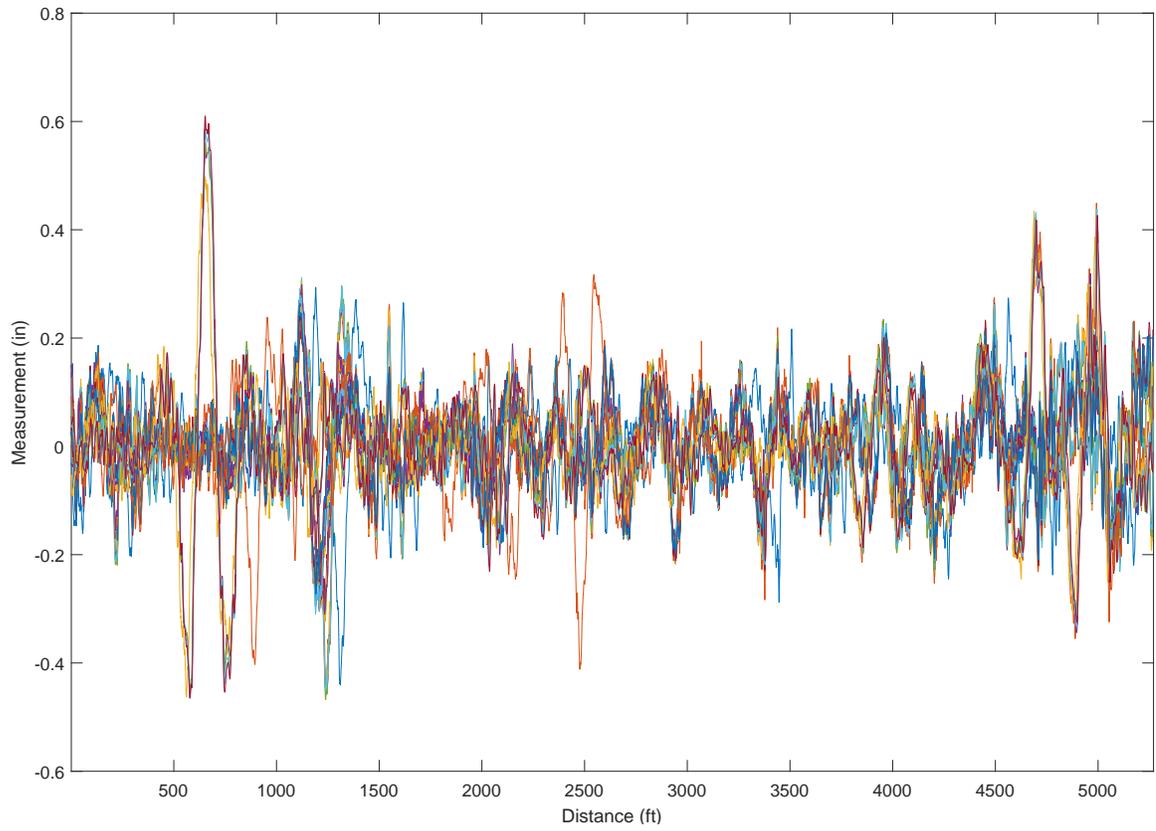


Figure A.9: Illustration of alignment left (124 ft) data at multiple inspection dates

A.2 Box Plots for Longitudinal Data

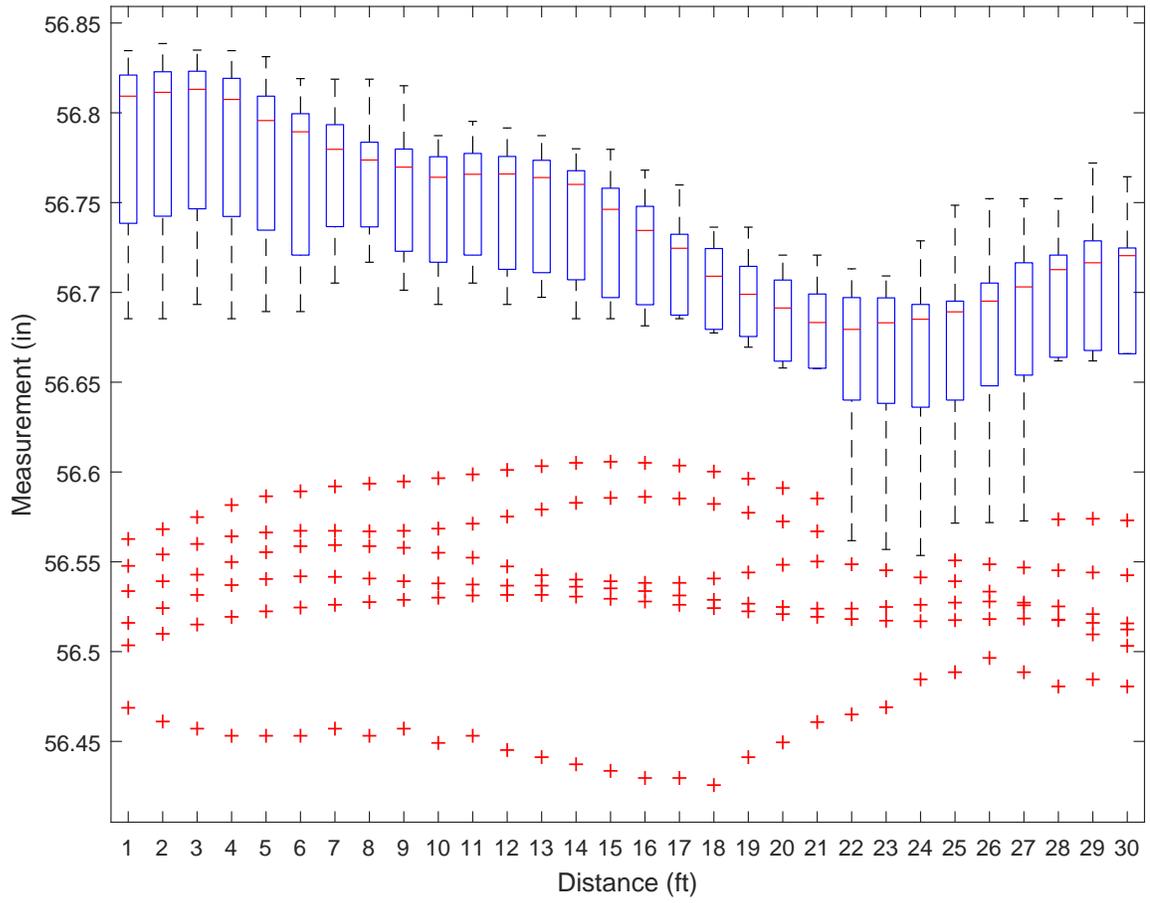


Figure A.10: Illustration of box plot for gage

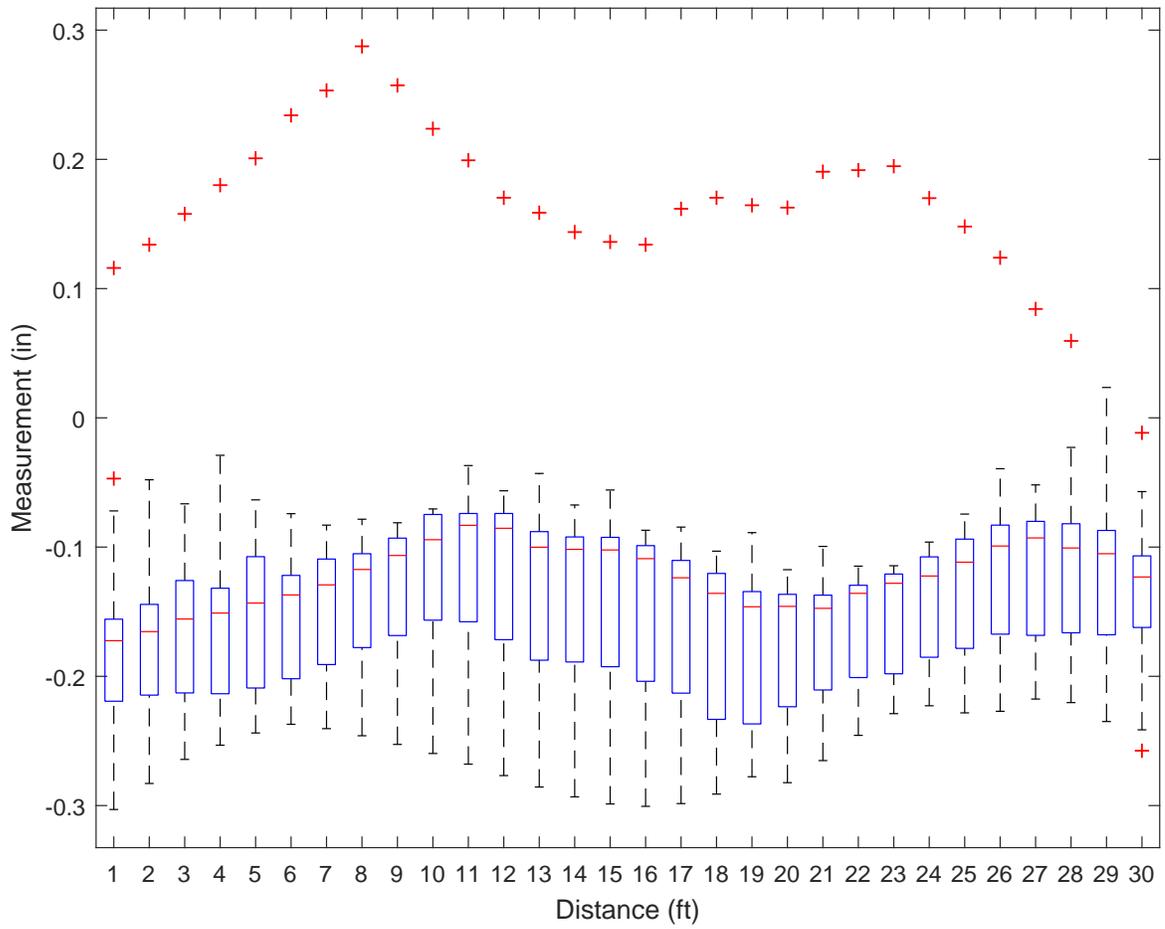


Figure A.11: Illustration of box plot for surface left (124 ft)

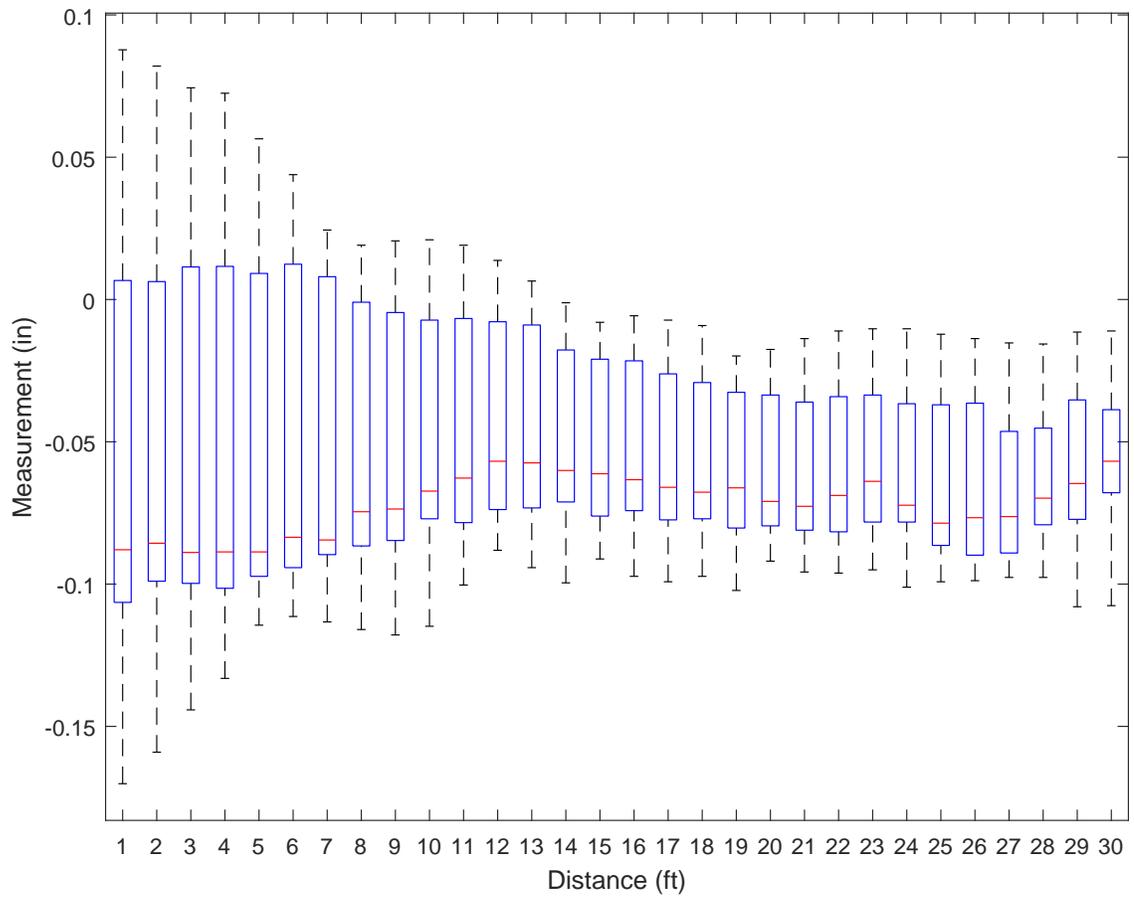


Figure A.12: Illustration of box plot for alignment right (124 ft)

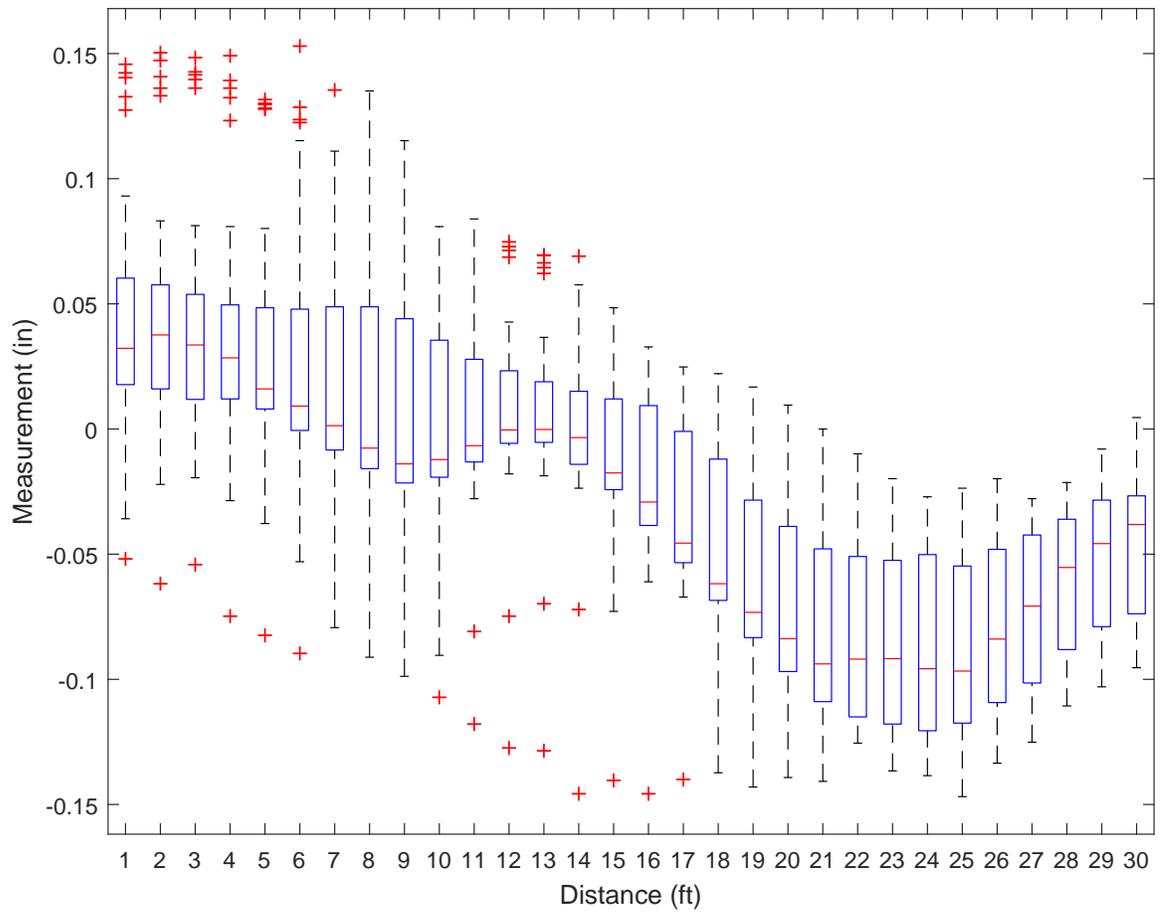


Figure A.13: Illustration of box plot for alignment left (124 ft)

Appendix B

MCMC OUTPUT PLOTS

B.1 Output Plots for 500-foot sections

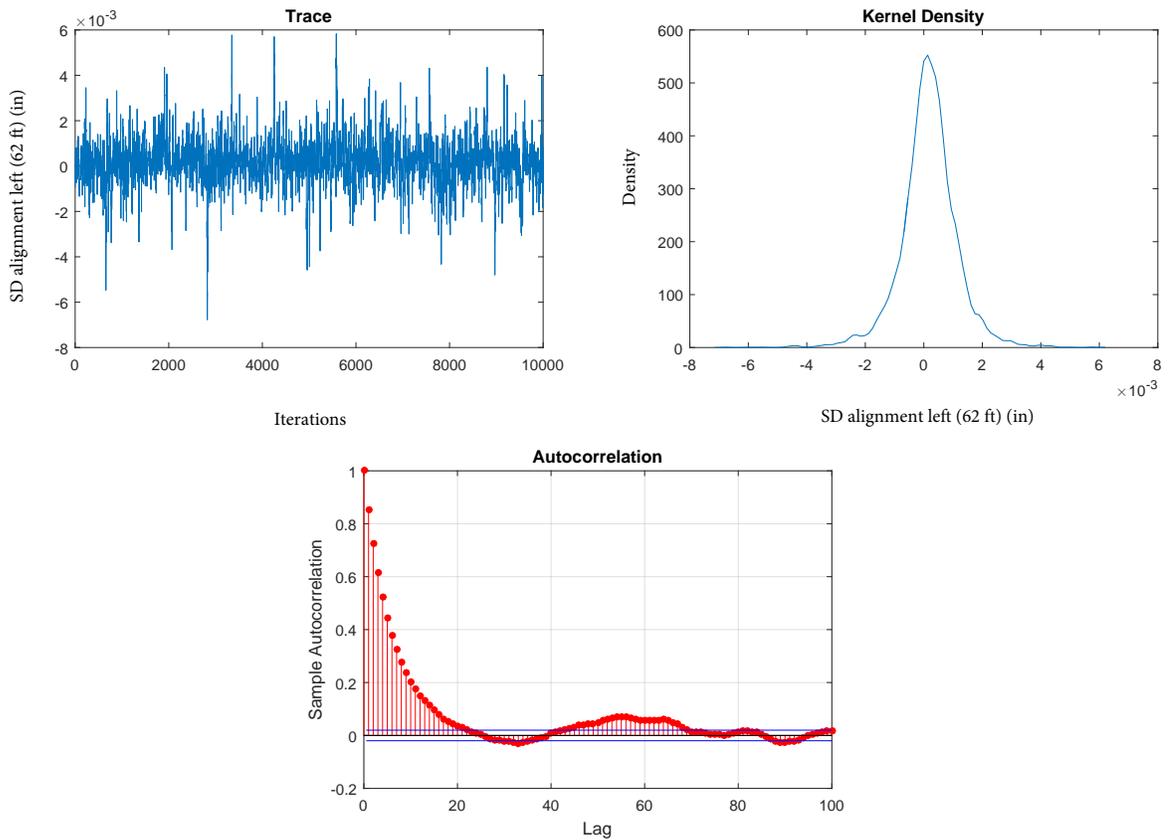


Figure B.1: Alignment left (62 ft): MCMC posterior plots for drift parameter

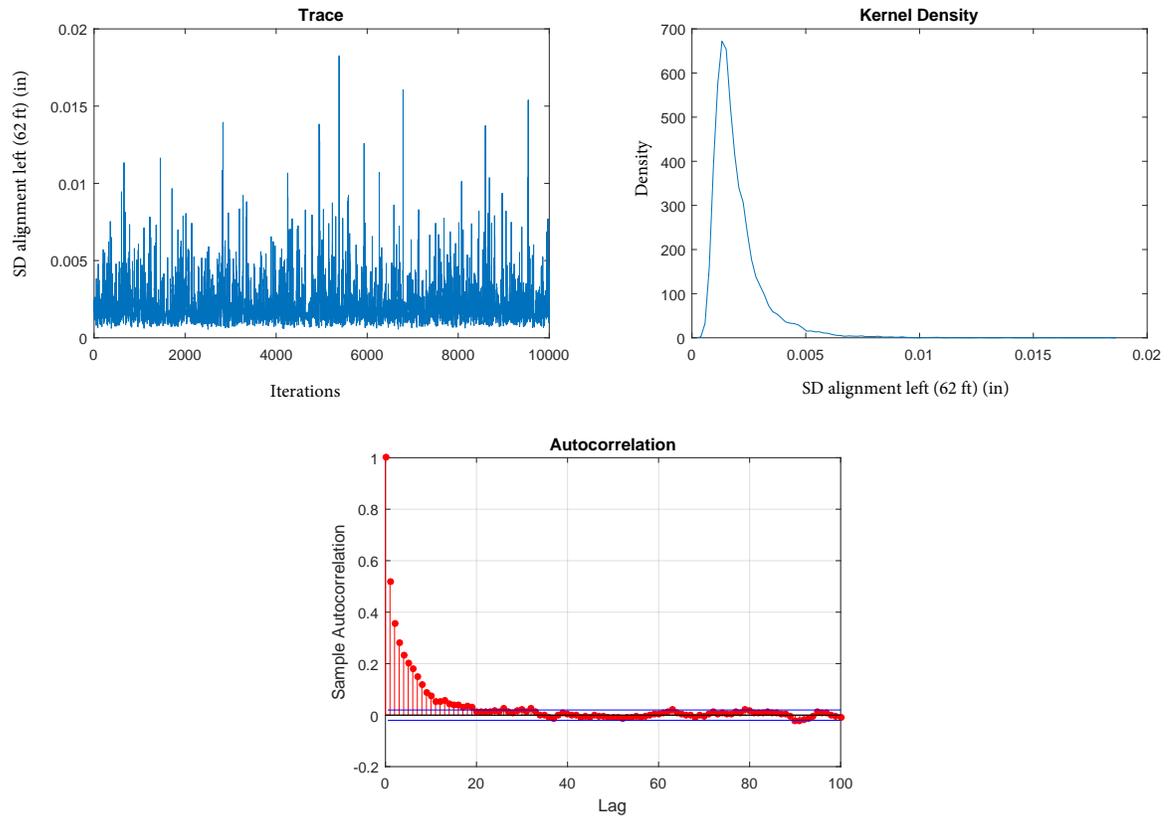


Figure B.2: Alignment left (62 ft): MCMC posterior plots for diffusion parameter

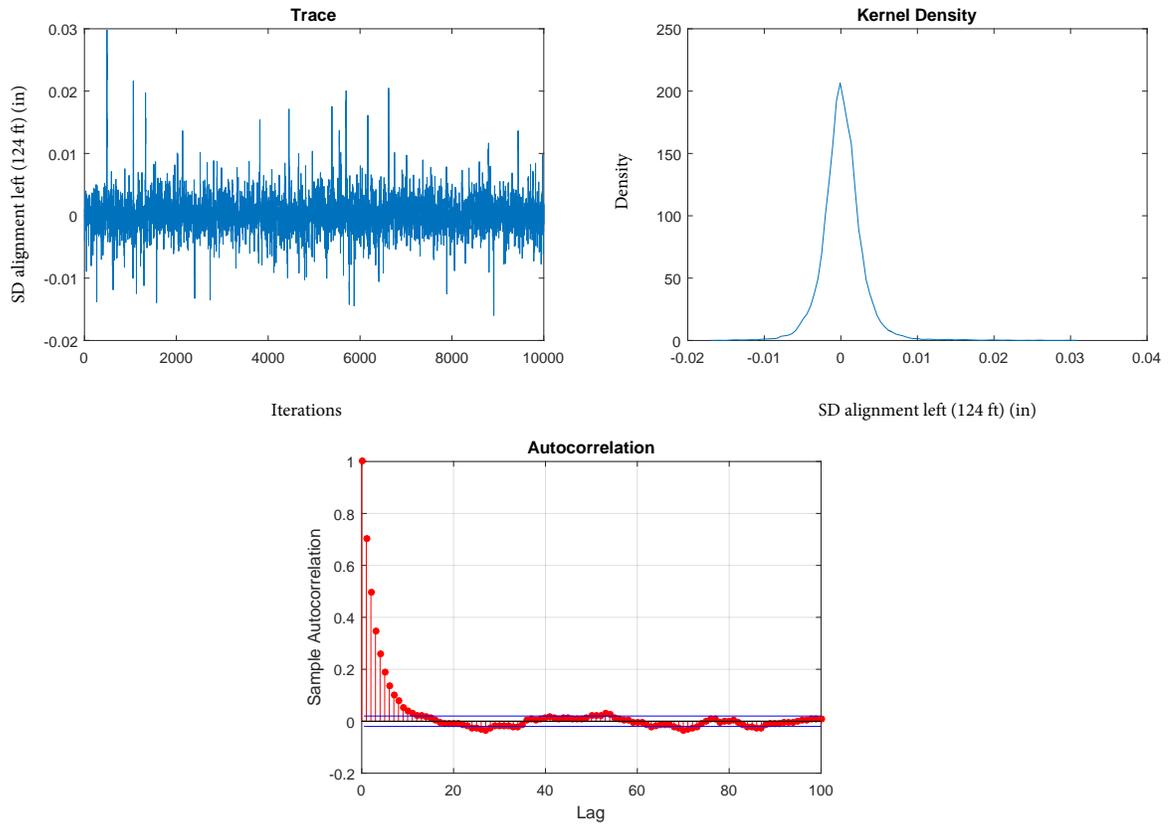


Figure B.3: Alignment left (124 ft): MCMC posterior plots for drift parameter

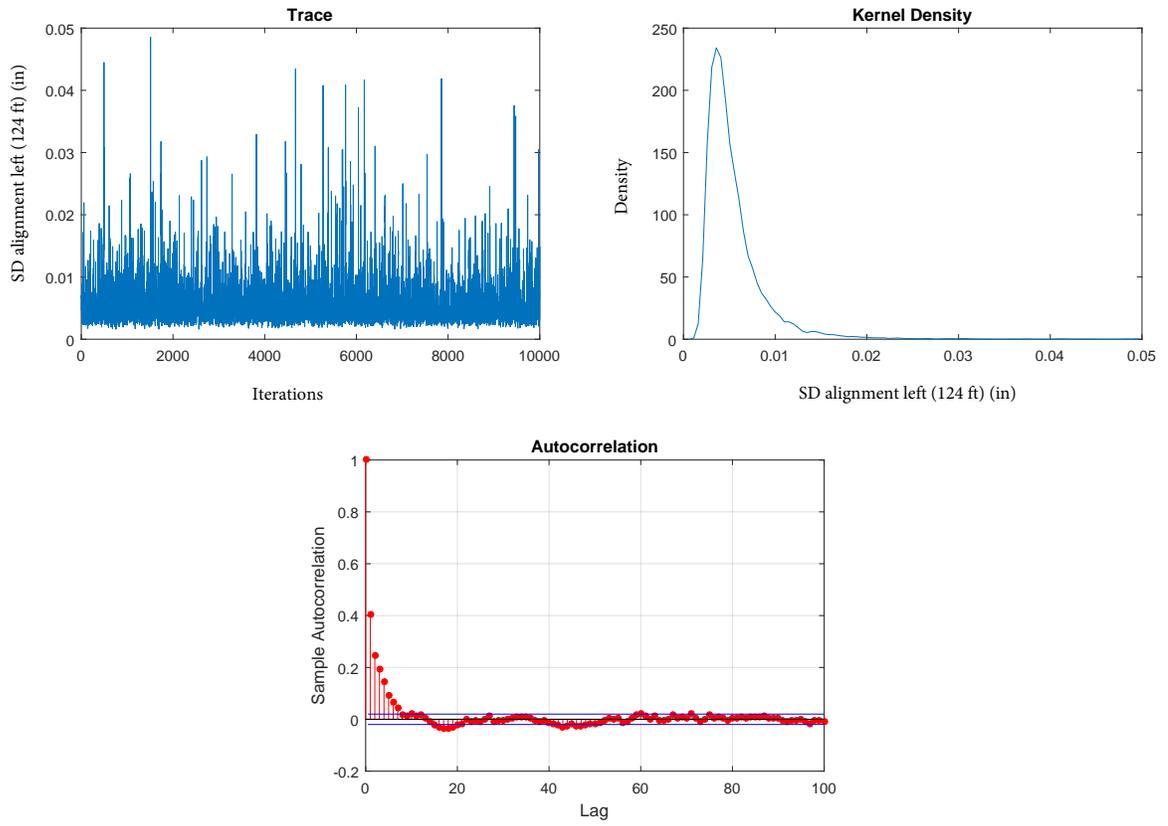


Figure B.4: Alignment left (124 ft): MCMC posterior plots for diffusion parameter

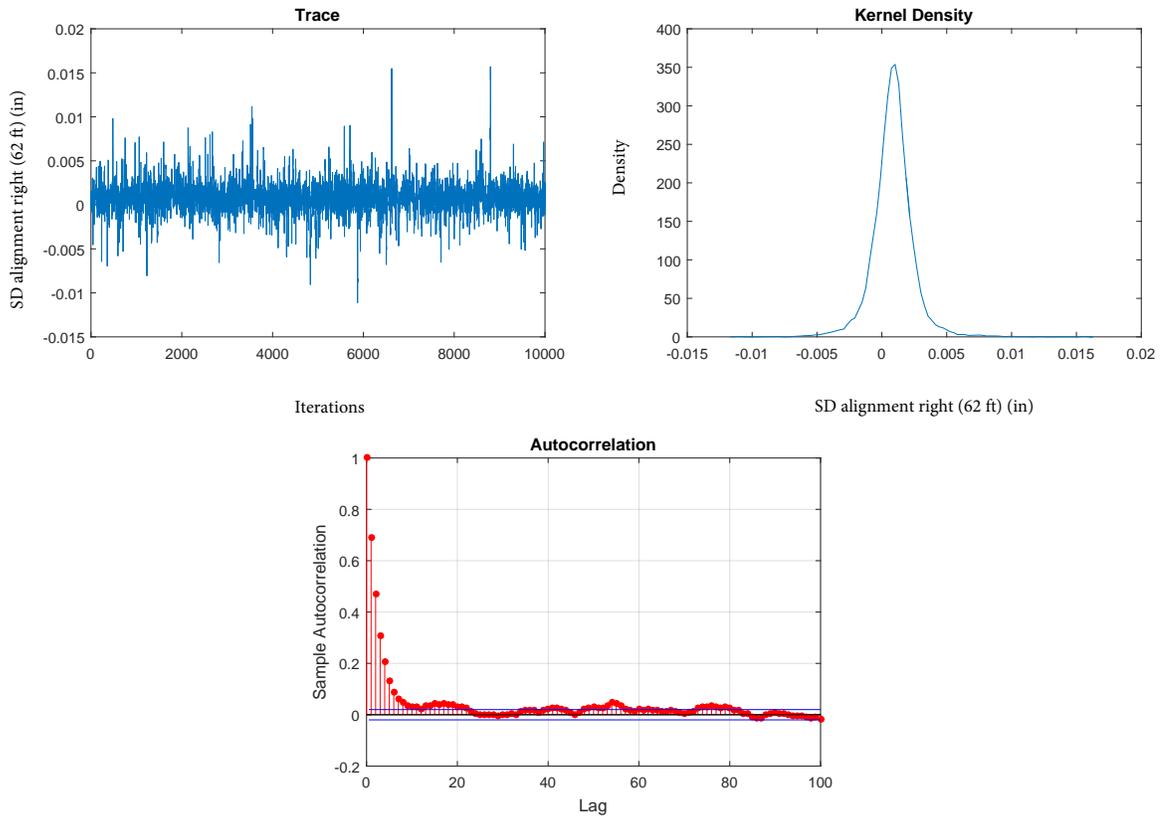


Figure B.5: Alignment right (62 ft): MCMC posterior plots for drift parameter

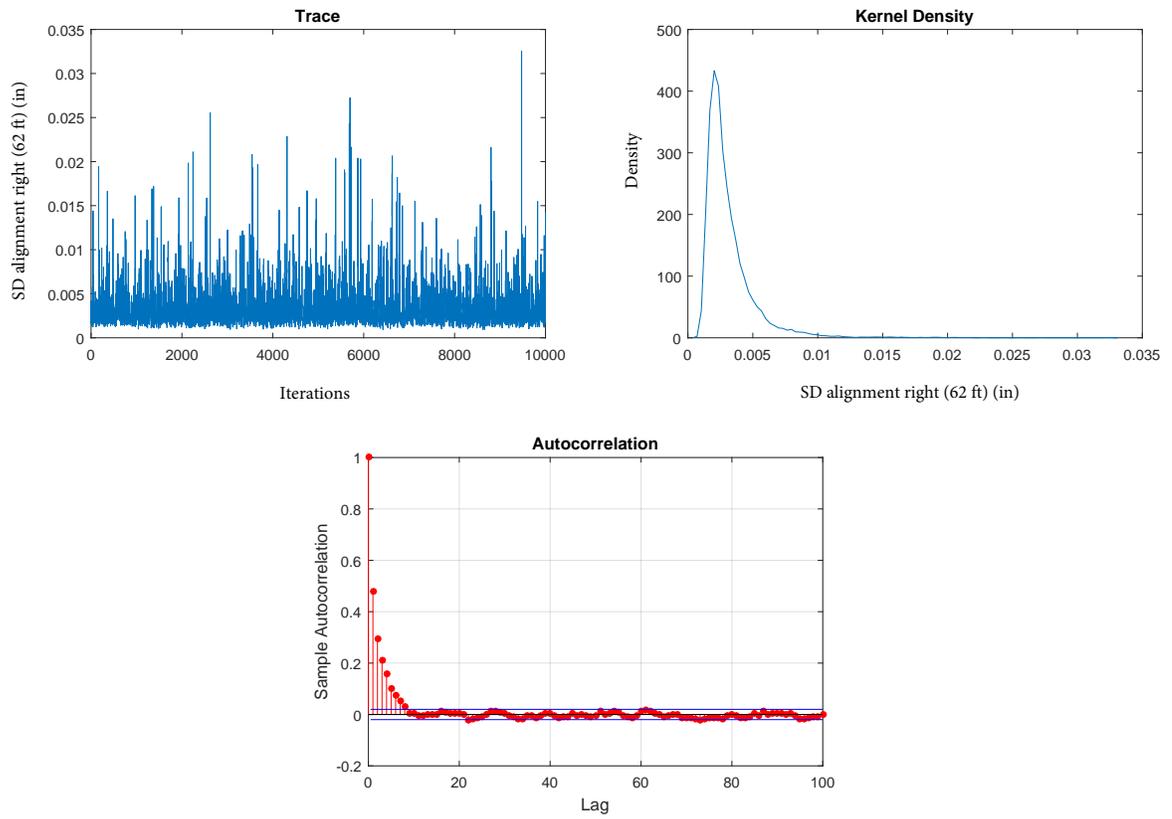


Figure B.6: Alignment right (62 ft): MCMC posterior plots for diffusion parameter

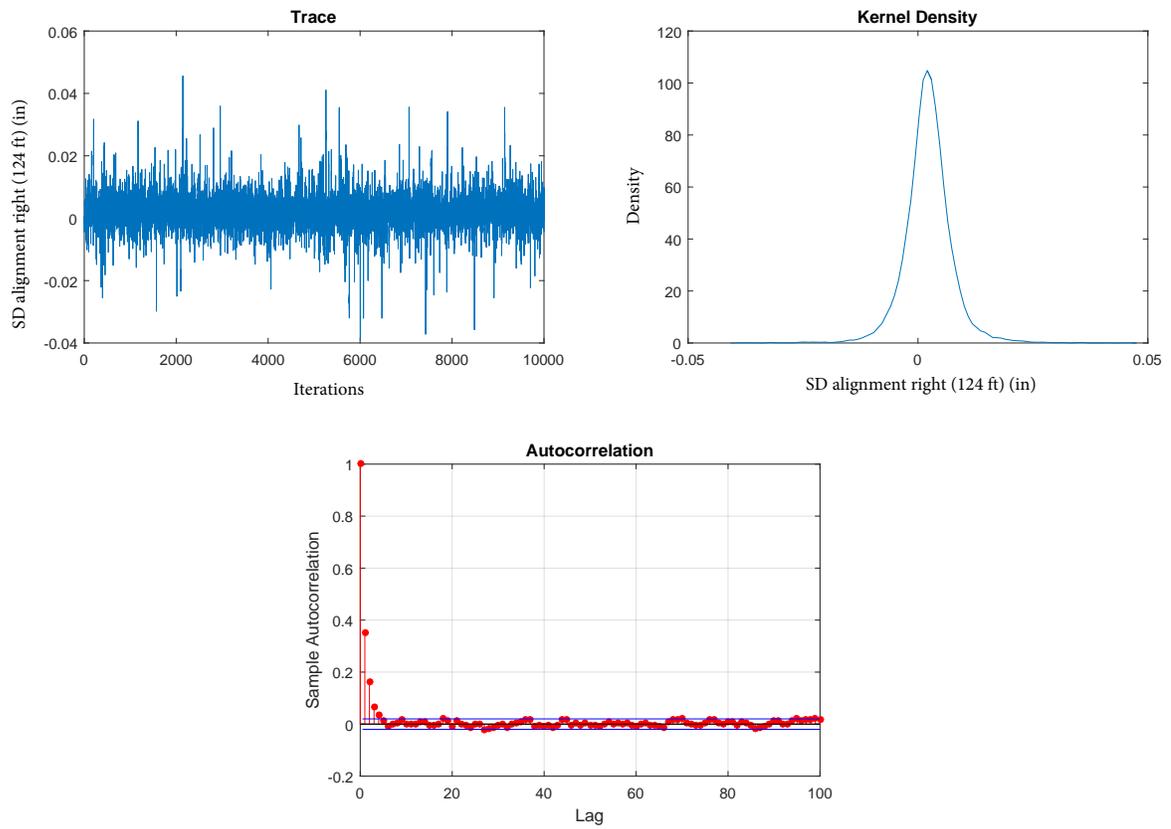


Figure B.7: Alignment right (124 ft): MCMC posterior plots for drift parameter

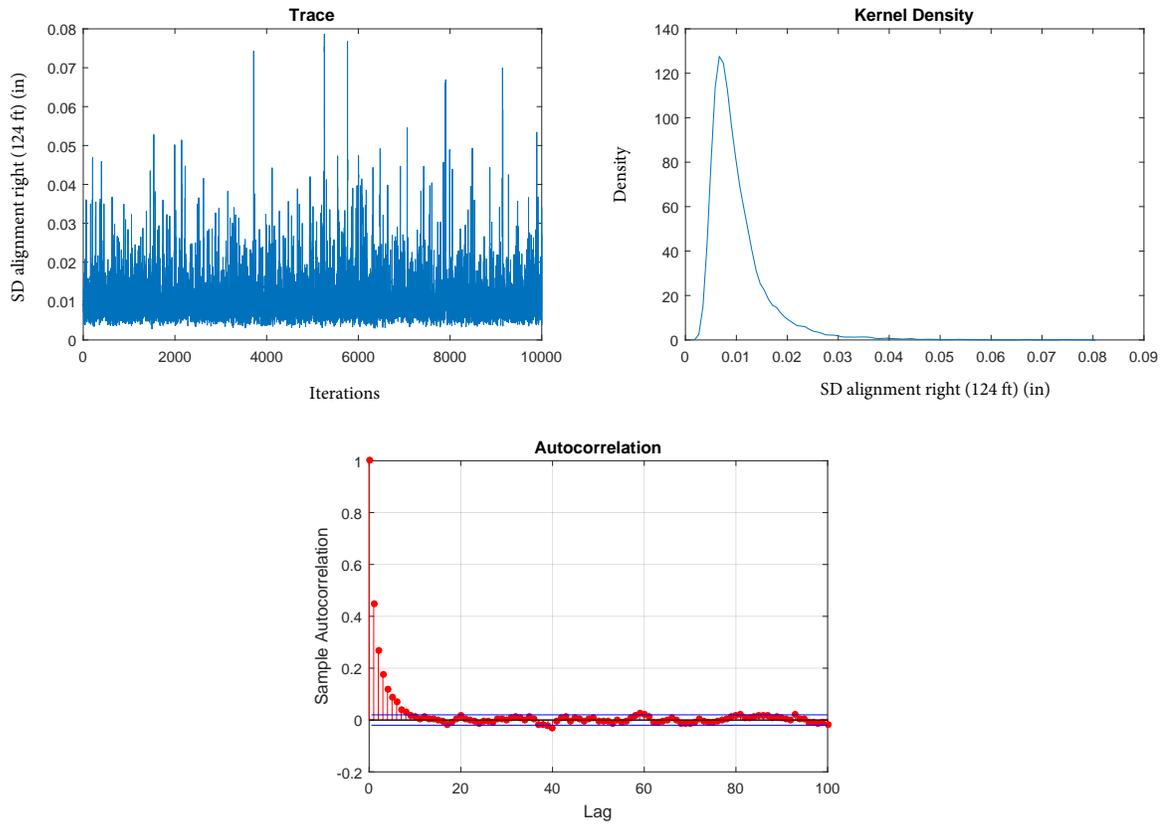


Figure B.8: Alignment right (124 ft): MCMC posterior plots for diffusion parameter

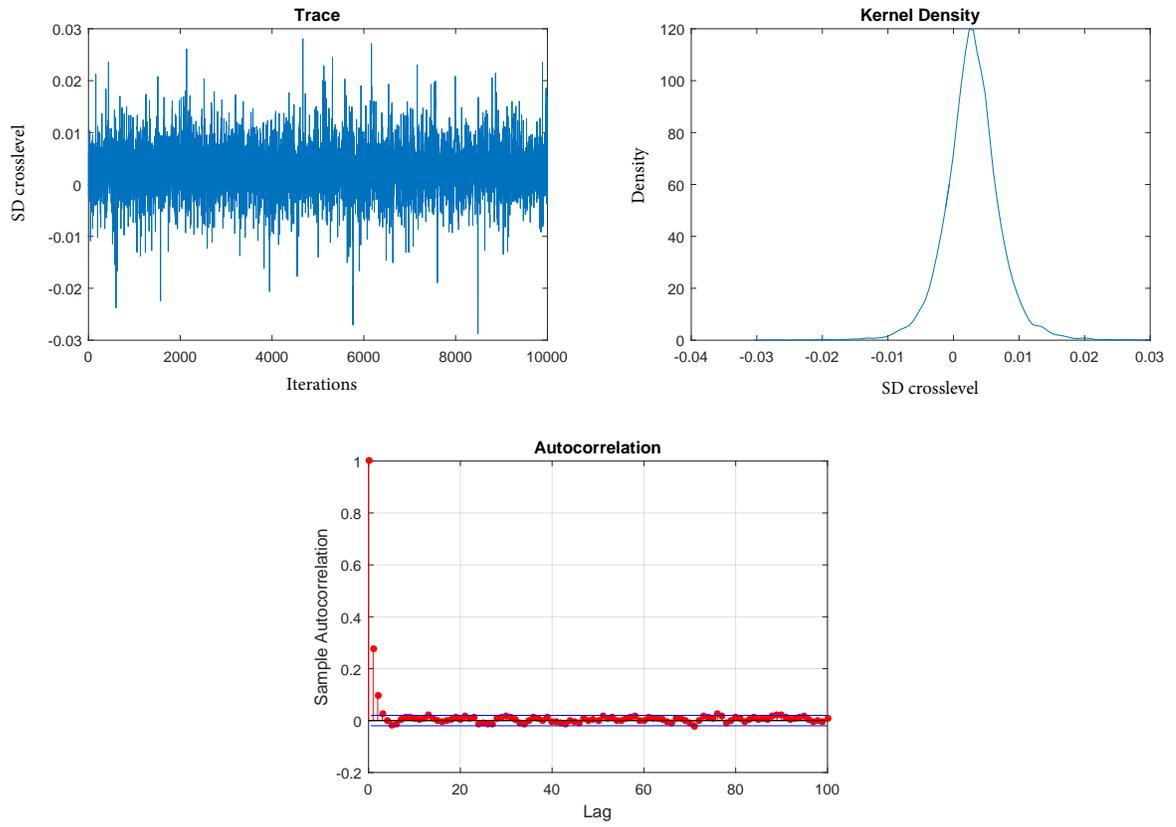


Figure B.9: Crosslevel: MCMC posterior plots for drift parameter

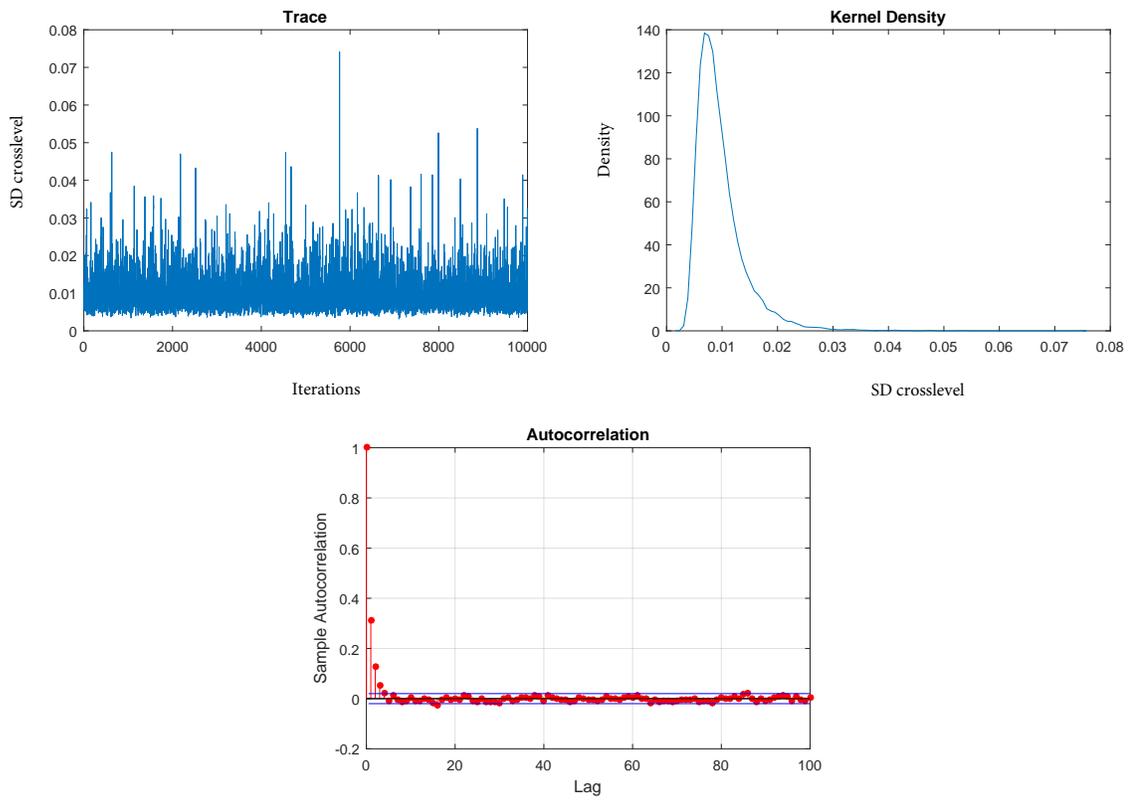


Figure B.10: Crosslevel: MCMC posterior plots for diffusion parameter

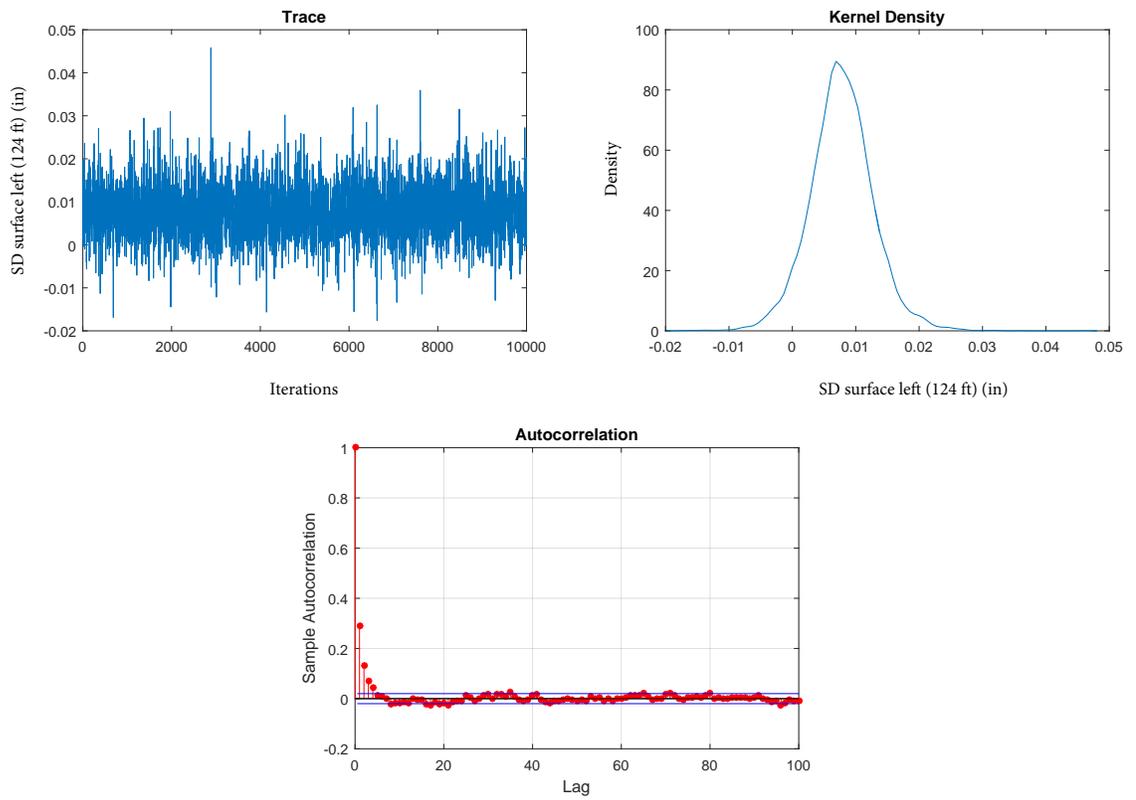


Figure B.11: Surface left (124 ft): MCMC posterior plots for drift parameter

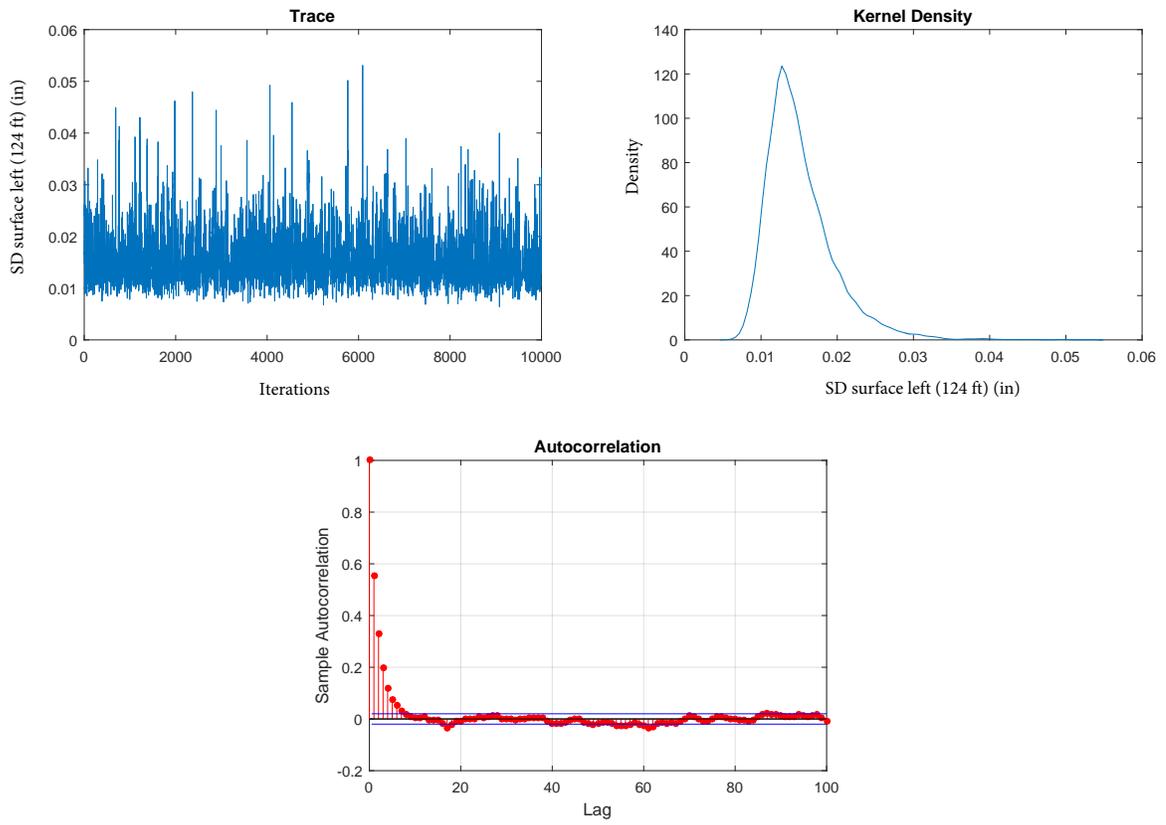


Figure B.12: Surface left (124 ft): MCMC posterior plots for diffusion parameter

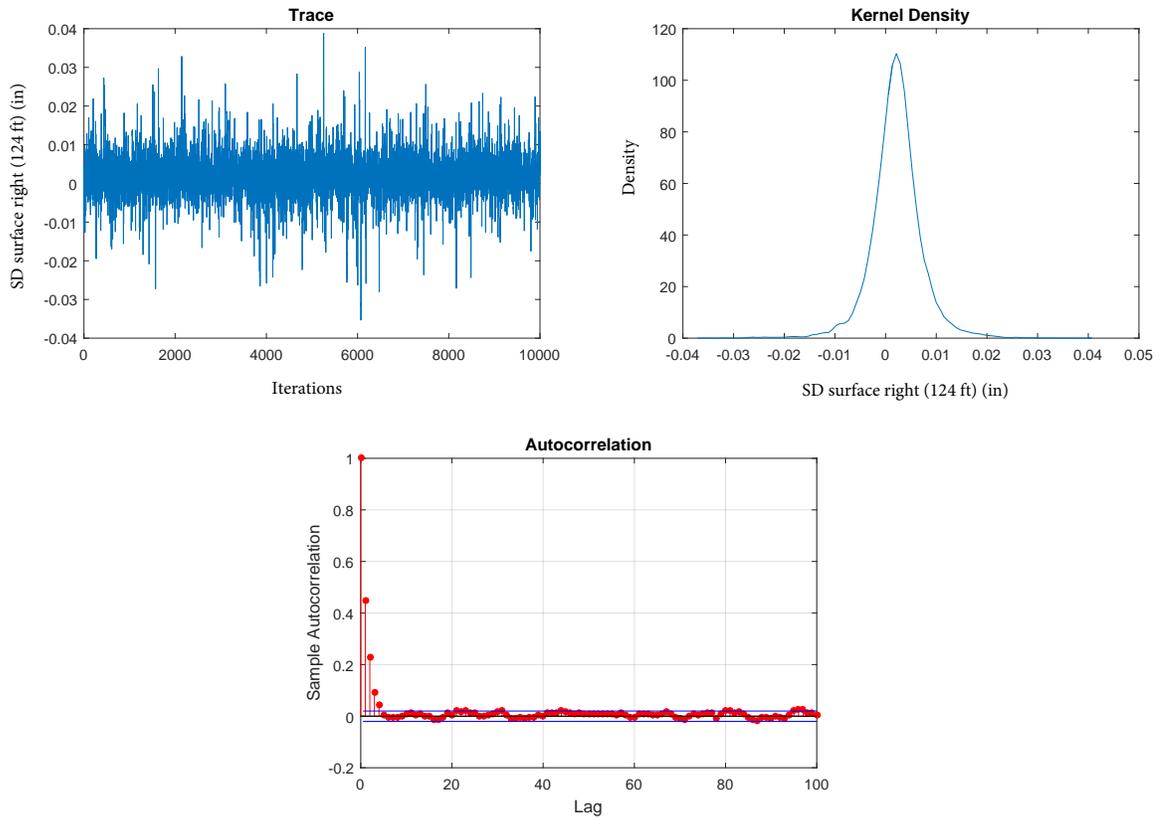


Figure B.13: Surface right (124 ft): MCMC posterior plots for drift parameter

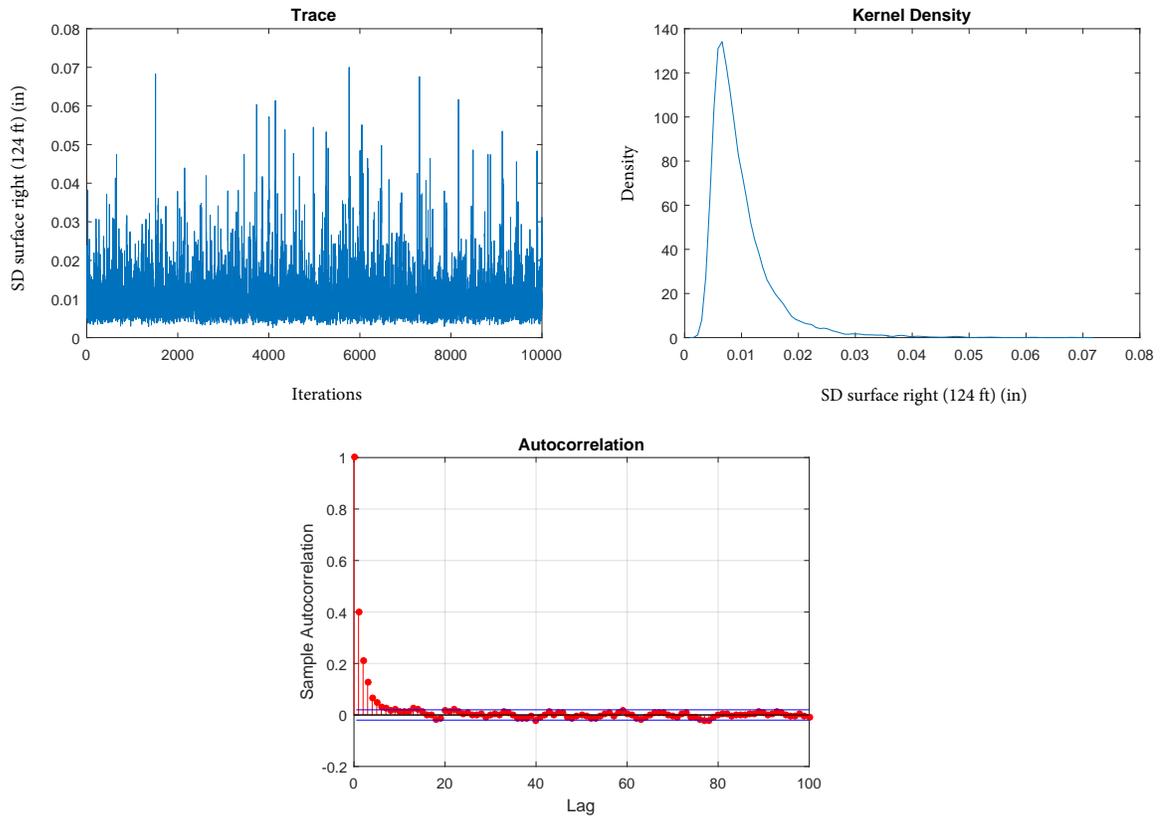


Figure B.14: Surface right (124 ft): MCMC posterior plots for diffusion parameter

Appendix C

PERMISSIONS

6/29/2017

RightsLink Printable License

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Jun 29, 2017

This Agreement between University of Delaware -- Silvia Galvan Nunez ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4138280985844
License date	Jun 29, 2017
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Wiley Books
Licensed Content Title	Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering
Licensed Content Author	Nii O. Attoh-Okine
Licensed Content Date	May 1, 2017
Licensed Content Pages	272
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	2
Original Wiley figure/table number(s)	Figure 1.1 Track structure components Table 9.1 Conjugate prior distributions
Will you be translating?	No
Title of your thesis / dissertation	Hybrid Bayesian-Wiener Process in Track Geometry Degradation Analysis
Expected completion date	Aug 2017
Expected size (number of pages)	170
Requestor Location	University of Delaware 127 The Green, 301 DuPont Hall NEWARK, DE 19716 United States Attn: University of Delaware
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing Address	University of Delaware 127 The Green, 301 DuPont Hall NEWARK, DE 19716 United States Attn: University of Delaware
Total	0.00 USD

<https://s100.copyright.com/AppDispatchServlet>

1/5

Figure C.1: Permission to use Figure 1.1 and Table 3.2