Identification and Characterization of Damaging Road Events

Craig T. Altmann

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mechanical Engineering

John B. Ferris, Chair
Montasir M. Abbas
Corina Sandu
Tana Tjhung
Brian Vick

May 18, 2020
Blacksburg, Virginia

Keywords: Pseudo Damage, Customer Usage, International Roughness Index (IRI), Event Identification, Event Characterization, Vehicle Durability

Identification and Characterization of Damaging Road Events

Craig T. Altmann

(ABSTRACT)

In the field of vehicle durability, many individuals are focusing on methods for better replicating the durability a user will experience throughout the typical design lifespan of a vehicle (*e.g.*, 100,000 miles). To estimate user durability a means of understand the types of damaging events and driving styles of uses must be understood. The difficulty with accurately estimating customer usage is, firstly, there is a large pool of possible roads for a user to drive along, for example, there are over 4 million miles of public roads in the United States, alone [1]. In addition, while measurements of these surfaces could be collected it would be impractical for two reasons, the first is the financial and extreme time burden this would take. Second, when collecting measurements of a road surface only the current state of a road surface can be measured, thus as a road deteriorates or is repaved the measurements collected would no longer be an accurate representation of the road. It should be mentioned that even, if all of the road surfaces were measured performing simulation and analysis of all of these road surfaces would be computationally intensive. Instead, it would be beneficial if select events that account for a significant portion of the damage a vehicle experiences can be identified. These damaging events could then be used in more complex vehicle simulation models and as input and validation of proving ground and laboratory durability testing. The objective of this research is to provide a means for improved estimation of vehicle durability, specifically a means for identifying, characterizing, and grouping unique separable damaging events from a road profile measurement.

In order to achieve this objective a measure that can be used to identify separate damaging events from a road profile is developed. This measure is defined as Localized Pseudo Damage (LPD), which identifies the amount of damage each individual road excitation makes to the total accumulated damage for a single load path in a vehicle system. LPD is defined as a damage density to minimize the effect of measurement spacing on the resulting metric. The developed LPD measure is causal in that the value of LPD at a location is not affected by any future locations. In addition, for a singular event (*e.g.*, impulse or step) in the absences of other excitations, the LPD value at the singular event location is equivalent to the total pseudo damage divided by the step size at the location.

Once a measure of pseudo damage density is known at multiple locations along a road profile for multiple load paths of interest, then separable damaging events can be identified. To identify separable damaging events the activity of the vehicle system must be considered

because separate damaging events can only occur when a region of inactivity is present across all load paths. Subsequently, an optimization problem is formed to determine the optimal active regions to maintain. The cost function associated with the optimization problem is defined to minimize the cost (number of locations maintained in damaging events) and maximize the benefit (the amount of pseudo damage maintained).

Lastly, a statistical test is developed to assess if separate damaging events can be considered to be from the same general class of events based on their damage characteristics. The developed assessment methods establish the similarity between two more separable damaging events based on application specific user defined inputs. In the development, two example similarity metrics are defined. The first similarity metric is in terms of distance and the second is in terms of likelihood (probability). The developed statistical analysis uses the current state-of-the-art in clustering algorithms to allow for multiple damaging events to be identified and grouped together.

Identification and Characterization of Damaging Road Events

Craig T. Altmann

(General Audience Abstract)

In the automotive field determining the level of damage a typical production vehicle experiences over its lifetime has always been a desirable criterion to identify. This criterion is commonly referred to as customer usage. By understanding the typical customer usage of a vehicle over the lifetime of a vehicle, automotive engineers are able to improve the design of vehicle components. The issue with defining customer usage is that there are millions of miles of roads that a customer can travel on and millions of customers that all have unique driving characteristics. While it is possible to collect measurements of these road surfaces to use in further vehicle simulations, it is not feasible both from a financial and time perspective. In addition, the simulation and analysis of all road surfaces would be computationally intensive. However, if select damaging events (regions of the road surface that excessively contribute to accumulated damage) are identified, then they can be used in more complex vehicle durability analyses with lower computational efforts.

In conventional damage analysis a total amount of accumulated damage is established for a known road surface. The issue with defining damage this way is that unique events which likely contributed a large amount of the accumulated damage cannot be identified. The first objective of this research is to define damage as a function of the vehicle's location along a road surface. Then, unique and separable damaging events can be identified and separated from sections of the road that do not significantly contribute to the accumulated damage. After defining this measure, an optimization problem is developed to identify damaging events based on maximizing the benefit (amount of damage accounted for in damaging events) and minimizing the cost (amount of road surface retained). Unique and separable damaging events are identified by solving this optimization problem.

While the optimization problem identifies unique, separable damaging events, it is likely that some damaging events contain similar characteristics to each other. When performing additional durability analysis, it would be beneficial to form connections between similar damaging events to allow for analysis to be performed based on groups of events. To identify damaging events with similar characteristics, a statistical analysis is developed as the last contribution of this work. By combining this analysis with current state-of-the-art clustering algorithms and user provided definitions based on applications, similar damaging events are able to be grouped together.

Dedication

This dissertation is dedicated to my family.

To my mother, Karen Altmann, your love and unwavering faith in me provided me with the confidence I needed to never stop pursuing my dreams.

To my father, Chester Altmann, your constant support and tenacious attitude taught me to always push myself and never give up.

To my fiancée, Katie, you have been there for me every day over the past five years. You have celebrated my achievements with me and supported me during the setbacks of my graduate studies. Words cannot describe my appreciation for your compassion and encouragement throughout this process.

To my siblings, Alex and Katelyn Altmann, at times we may annoy each other, but we are always there for each other and provide constant support and encouragement.

To my grandmother, Selma Merkley, who taught me to always see the positive in life and realized my potential even before I did. The life lessons she taught me, made me the compassionate and focused individual I am today.

To my grandmother, Rose Ann Altmann, your love and support has continued to encourage me to pursue my dreams.

To my grandfathers, Norbert Merkley and Chester Altmann Sr., both who passed away early in my life. While my time with them was short, I believe their spirit still lives on inside of me and their memories cary on through other family members.

To my family dog, Gunner, who I consider a second brother. You were always my side kick since grade school and would never leave my side when I came home from college. I will always remember your energetic personality and wagging tail.

To my uncle, Jim Merkley, your support and belief in me since my undergraduate studies is greatly appreciated.

Lastly, to all other extended family members who have provided me with the support and encouragement to continue pushing my limits and pursuing my dreams.

Acknowledgments

I would like to express my greatest appreciate for my advisor, Dr. John Ferris. His guidance, assistance, and belief in my research and approach to identifying and characterizing damaging events is greatly appreciated. In addition, his support and willingness to let me take a leadership role in sponsored projects in the lab was very much appreciated and I believe a great preparation for my future career. In addition to my advisor, it was a great pleasure working with former members of the Vehicle Terrain Performance Laboratory, especially Yong Suk Kang and Savio Pereira. I would also like to thank my committee members Dr. Tana Tjhung, Dr. Montasir Abbas, Dr. Corina Sandu, and Dr. Brian Vick for their support and assistance throughout my graduate studies.

I would also like to thank all of the sponsors I have worked with throughout my time in the Vehicle Terrain Performance Laboratory (VTPL). I would like to thank the individuals at Harley Davidson who have provided helpful feedback and recommendations regarding the research presented in this work. I would also like to thank the Federal Highway Administration (FHWA), specifically Andy Mergenmeier, for his support throughout the sponsored FHWA project.

Lastly, I would like to thank the faculty in the mechanical engineering department at Rose-Hulman Institute of Technology. There were numerous faculty members that were always there to help me with my studies and encouraged me to pursue a PhD degree. One faculty member that was the most significant help was Dr. Daniel Kawano. Throughout my years at Rose-Hulman he worked with me on numerous independent studies and activities with the Formula SAE team all of which prepared me for my time at Virginia Tech. I appreciate the time he set aside to talk with me, at least once a semester, throughout my graduate studies. These conversations really helped me keep sane and gave me the encouragement to keep pushing forward.

Contents

List of Figures

List of Tables

Nomenclature

D_T	Total accumulated pseudo damage
C_D / C	Damage scaling factor
β	Fatigue exponent (typically ranges between 3 and 8 for vehicle applications [2])
L_ℓ	Suspension force load amplitude identified using rainflow counting
N_ℓ / M	Number of load amplitudes
$\dot{D}_j$	Pseudo damage rate at the jth location
G	Damage rate scaling factor
$F_{susp,j}$	Suspension force at the jth location
F_j	Force through a load path at the jth location
$F^p_{sups,j}$	Peak suspension force corresponding to $F_{susp,j}$
$F_{p,j}$	Peak force in a load path corresponding to F_j
$\dot{F}_{susp,j}$	Suspension force rate at the jth location
$\dot{F}_j$	Rate of change of force through at load path at the jth location
d_i	Localized pseudo damage, defined as a pseudo damage density
Δt_j	Time step corresponding to location j
u_i	Longitudinal distance traveled at the ith location of the road profile
Δu_i	Longitudinal step spacing at the ith road excitation at the jth location
$\alpha_{i,j}$	Relative contribution factor for the ith road excitation at the jth location
ε_j	Scaling factor for the relative contribution
$\gamma_{(j-i)}$	Scaling factor for the relative contribution when a linear time invariant assumption is made
$z_{r,i}$	Road profile amplitude at the ith location
N	Total number of locations in a road profile

$\boldsymbol{A_{p,i}}$	Activity in the pth load path at the ith location
$\boldsymbol{L_{k,p}}$	Inactive window length for the kth location in the pth load path
$\boldsymbol{T_{s,p}}$	Settling time of the pseudo damage in the pth load path
$\boldsymbol{W_{k,p}}$	Set of locations that lie in the kth window for the pth load path
$\boldsymbol{\psi}$	Measure on the set of activity values for each window $W_{k,p}$
$\boldsymbol{V_{k,p}}$	An inactive window for the kth window of the pth laod path
$\boldsymbol{\gamma_p}$	Activity threshold value for the pth load path
$\boldsymbol{U_p}$	The union of inactive locations across all load paths of interest
$\boldsymbol{I}$	The intersection of inactive locations across all load paths of interest
$\boldsymbol{AR_r}$	A separate unique active region for a road profile
$\boldsymbol{E_r}$	The rth damaging event for the road profile based on the solution of an optimization problem
$\boldsymbol{E}$	The set of locations for all damaging events
$\boldsymbol{FD_{E,p}}$	Fraction of pseudo damage for the pth load path in the set of damaging events
$\boldsymbol{\eta_p}$	Minimum threshold on the fraction of pseudo damage in damaging events for the pth load path
$\boldsymbol{FL_E}$	Fraction of locations that are members of damaging events
$\boldsymbol{FM_{E,p}}$	Figure of merit of the pth load path
$\boldsymbol{\phi}$	Objective function defined by a selected measure on the figure of merits for all load paths
$\mathbf{d}_{\boldsymbol{r}}$	Matrix of pseudo damage densities for an rth damaging event, each row in the matrix corresponds to the pseudo damage density in a load path of interest and each column is a discrete sample
$\boldsymbol{N_r}$	Number of unique pseudo damage density samples for the rth damaging event
$\boldsymbol{n_p}$	Number of load paths of interest when clustering damaging events
$\bar{\mathbf{d}}_{\boldsymbol{r}}$	Average pseudo damage density vector for the rth damaging event
$\boldsymbol{\Sigma_r}$	Pseudo damage density sample covariance matrix for the rth damaging event
$\boldsymbol{\Sigma_{q,r}}$	Pooled pseudo damage density sample covariance matrix between the qth and r^{th} damaging events
$\boldsymbol{\rho_{q,r}}$	Mahalanobis distance squared between between the damaging event q and damaging event r

$\boldsymbol{\Lambda_i}$	Total population size associated with the reference distribution when the samples are used to establish the parameters of the distribution
$\boldsymbol{\mathbb{E}_C}$	Cluster of two or more damaging events events
$\boldsymbol{\rho_{q,C}}$	Set of Mahalanobis distances squared between the new event q and all events in the cluster C
$\boldsymbol{p_{q,C}}$	Set of likelihood measures between the new event q and all events in the cluster C
$\boldsymbol{s_{q,C}}$	General similarity definition between the a new event q and all events in the cluster C; the similarity metric is defined in terms of distance ($\boldsymbol{\rho_{q,C}}$) or likelihood ($\boldsymbol{p_{q,C}}$)
$\boldsymbol{L(s_{q,C})}$	Linkage function based on similarity between a new event, q, all an events in a cluster C
$\boldsymbol{S_{q,C}}$	Overall similarity between a new event q and a cluster of events $\mathbb{E}_C$
$\boldsymbol{i}$	Road location index
$\boldsymbol{j}$	Time location index
$\boldsymbol{q}$	Index for a new event considered in a cluster
$\boldsymbol{r}$	Event index within a cluster

Chapter 1

Introduction

The focus of this dissertation is identifying, characterizing, and grouping damaging road events into a vehicle. The theoretical developments in this work make no assumptions regarding vehicle modeling or road surface characteristics allowing for application to a wide range of models and applications. Throughout this work examples are provided using a Golden Quarter Car model to illustrate the process and provide results which are applicable to a range of vehicles (typical sedan passenger cars). The quarter car model allows for adequate representation of ride characteristics which are typically important in vehicle durability analysis. In this work, damage is characterized by pseudo damage, which is damage calculated using load-time history. Pseudo damage is selected for this work because no material or geometric part assumptions are necessary allowing for events to be identified by information which is commonly known during pre-production stages of vehicle design and analysis.

This work begins by developing a novel measure of damage density, referred to as Localized Pseudo Damage (LPD). The benefit of LPD over current pseudo damage analysis is that pseudo damage is linked to specific locations along a road profile. By defining LPD as a damage density the sensitivity of the measure to differences in spacing and sampling rate is minimized. During the development of LPD, the unique properties of the selected model are explained and demonstrated. These properties include: causality and ability to distinguish between damage inducing road excitations and damage reducing road excitations based on the specific context in which the excitations occur. Multiple examples including an impulse excitation, step excitation, and a real-world road surface are provided to demonstrate these properties.

Through calculation of damage density for defined load paths of interest in a vehicle model, a road profile can be separated into damaging events and non-damaging regions. Damaging events are parts of the road that contain significant damage density in at least one load path of interest, non-damaging regions are portions of the road surface that are non-damaging to all load paths of interest. Unique damaging events are separated by non-damaging regions

and are considered separate damaging events. In the methodology developed, a set of LPD values must be known for each load path of interest; however, the definition of which load paths are of interest is application specific and outside the scope of this work. To allow for flexibility in the load paths of interest, the identification methodology is developed to allow for any desired number of load paths. In addition to the definition of load paths of interest, two user provided pieces of information are needed for each load path of interest. The first necessary piece of information is the minimum total accumulated damage to account for in the identified damaging events. The second piece of information is a settling time threshold magnitude on the damage rate impulse response. Using these pieces of information non-damaging regions can be identified which, in turn, allows for identification of damaging events.

Lastly, the characteristics of damaging events are used to identify events that are similar to each other. The first challenge when defining the similarity (or dissimilarity) between damaging events is that the measure must account for the uncertainty in the estimate in damage associated with the event. The second challenge is handling the potential scalability of damaging events, or the ability repeat an event an integer number of times and achieve similar damage characteristics. In the work developed both challenges are addressed by using the set of pseudo damage density values known in each load path for a given event.

The remainder of this chapter provides motivation for the research of identification and characterizing damaging road events. Subsequently the problem this research focuses on addressing is provided in Section 1.2. In Section 1.3 the thesis statement and overall scope of the work are presented. The main contributions of this research are provided in Section 1.4. All current publications are provided in Section 1.5. Lastly, a brief outline of the dissertation is provided in Section 1.6.

1.1 Motivation for Research

When predicting vehicle durability, it is critical to accurately model the vehicle and to excite the model with damaging road events that are representative of the real-world. Without accurate representation of these damaging events the calculated vehicle damage may not be representative of the damage users' experience. The ability to accurately collect high-resolution road surface measurements continues to improve. However, simulation of these surfaces can take significant computational effort. Specifically, for durability analysis, it is likely that the majority of the damage accumulated is attributed to a few localized events. Thus, it would be useful if these events could be identified *a priori* allowing for simulation of only the most damaging road segments, thereby resulting in a savings of computation time. To identify these discrete events, the effect discrete road excitations have on the accumulated pseudo damage must be identified. One method for identifying this effect is calculating a pseudo damage density, which provides a measure of pseudo damage over a specified length of the road profile.

When defining customer usage and vehicle durability analysis, it would be ideal to know every damaging event that a customer will experience during the life of a vehicle. Two major issues stand in the way of this ideal. First, it is only feasible to collect high-resolution samples of the current state of road surfaces. There are about four million miles of road surfaces in the United States [1] and these surfaces degrade and are repaired over time. Second, vehicle simulations and analyses of vehicles driving on these surfaces can be computationally intensive. One way to address these issues is to identify damaging events in measured road surfaces. This damage characterization could then be used to develop proving ground durability test cycles and laboratory tests based only the damaging road segments.

When events are identified from the current state of a measured surface, there is always the chance that the road surface can deteriorate further and become more damaging. In addition, it not reasonable to measure every mile of public roads. Therefore, it is necessary to use a known subpopulation of damaging events to estimate customer usage. If groupings of damaging events based on similarities of damage characteristics can be established, then durability testing cycles can be improved to better represent groups of damaging events. In addition, the number of necessary repetitions of durability events (or set of durability events) can be established based on the number of damaging events in the group along with a likelihood measure of the customer experiencing each event in the group. Thus, durability test cycles will not only closely match the accumulated damage in a vehicle system across load paths, it will match the damage characteristics associated with each damaging event.

This work proposes a measure of pseudo damage density that allows for pseudo damage to be linked to individual locations along a road profile. Through development of the pseudo damage density, unique damaging events can be identified from a road profile. The damaging event identification method proposed in this work ensures that events are separate and the dynamics of the vehicle have sufficiently settled between consecutive damaging events. Lastly, once multiple damaging events are identified, a statistical test is developed to identify if two or more damaging events are considered to be from the same population of damaging events based on the damage characteristics (pseudo damage density) of each event.

1.2 Problem Statement

An important component of developing and improving vehicles produced is accurately modeling customer usage. Due to the variability between customers and regions this is a challenging problem to solve. In addition, to the variability due to customers, it is desirable to define customer usage early in the design cycle and potentially as a function of vehicle class rather than a targeted component on a specific vehicle.

In prior work, customer usage was defined based on road roughness [3]. However, over the year's vehicle modeling and simulation along with ability to accurately measure road surfaces has improved. The improvements in these areas has made it possible to determine a

load time history for a specific load path in a vehicle (or class of vehicles depending on the model used). Once this load time history is established conventional damage models can be applied to calculate a pseudo damage. A study performed by Zamora-Alvarez reveled that pseudo damage is not one-to-one with respect to road roughness (specifically the International Roughness Index) [4]. Therefore, it would be beneficial if the pseudo damage can be expressed as locations along a road profile to make identification of separate damaging events from a road profile possible.

Defining a relationship between pseudo damage and an overall measure of the road roughness is difficult and, in some cases, not one-to-one as shown by Zamora-Alvarez for the International Roughness Index. However, it is hypothesized that if a road profile is segmented into non-damaging regions (segments of the road where little to no pseudo damage is accumulated in any load path of interest) and damaging events then similarities between separate damaging events can be captured. These similarities can be used to develop a set of general (basis) damaging events that have similar damage characteristics to all events within the same group or to develop a durability testing cycle. By identifying similarities between damaging events, the similarities between proving ground durability events and public roads can be established to identify the best set of durability events to reproduce the damage a customer experiences.

1.3 Thesis Statement and Scope of Work

Thesis Statement: A measure of damage density can be developed to allow for pseudo damage to be linked to a road segment and this measure allows for damaging events to be identified and provides characteristics for grouping similar events.

The primary goal of this research is to develop a measure of damage density which defines the amount of damage each individual road excitation contributes to the total accumulated pseudo damage. The measure developed in this research, localized pseudo damage, is causal and correctly identifies the location and magnitude of damaging events. The theoretical developments in this work makes no vehicle, speed, or road modeling assumptions, resulting in a measure that is applicable to all vehicle and road models. Examples are provided using a golden quarter car to provide results which pertain to a range of vehicles rather than a single vehicle. The limitations of this work are: having a known/measured road surface and a verified vehicle model that accurately models the characteristics of interest.

1.4 Main Contributions

This research focuses on development of damage density for vehicle-terrain interactions, identification of events based on damage density, and grouping similar damaging events

using clustering algorithms. The methodology developed in this work is not limited to the example cases provided. For example, any vehicle model of interest can be used to identify a force response for calculating pseudo damage and the measure of damage rate developed in this work. The main contributions of this work are:

1. Representation of pseudo damage as a damage density
   (a) A metric, which quantifies the amount of pseudo damage contributed by a road excitation, is developed in terms of a density.
   (b) The developed metric is shown to be causal, accurate in identifying damaging locations, and capable of distinguishing between damaging inducing and damage reducing excitations, based on the context in which they occur.
   (c) A proof of concept is provided through three example longitudinal road profiles (impulse and step excitation, a dual impulse excitation, and measured road surface).
2. A novel method for identifying damaging events
   (a) A road surface is represented as a pseudo damage density for any number of vehicle load paths.
   (b) Each location of a road surface can be characterized as a portion of a damaging event or a portion of a non-damaging region based on a relationship between damage density and the load path dynamics.
   (c) Candidate non-damaging regions are defined using a windowing approach where the window length is dependent on the settling time for the load path dependent damage rate function.
   (d) Non-damaging regions are defined as the intersection of candidate non-damaging regions across all load paths of interest.
   (e) The complement of the non-damaging regions provides the complete set of damaging events.
   (f) Unique damaging events are the sets of contiguous locations within the complete set of damaging events.
   (g) Proof of concept is provided using publicly available longitudinal road profile.
3. Method for grouping similar damaging events using common clustering algorithms
   (a) Events are decomposed into a mean pseudo damage density vector and a sample covariance matrix for the pseudo damage density.
   (b) A method for defining the similarity between two or more damaging events is defined in terms of distance and likelihood (probability)

(c) The Mahalanobis distance metric is provided as an example distance similarity metric between damaging events.

(d) A Beta distribution is used to provide a reference distribution for the Mahalanobis distance to define the similarity in terms of probability.

(e) Two examples are provided using damaging events identified from measured highway surfaces to illustrate how the application should inform the selection of a clustering algorithm and the criteria for rejecting the hypothesis that two events are drawn from the same population.

1.5 Publications

Published Journal Articles:

1. Altmann, Craig T., and Ferris, John B. "Theoretical Development of Localized Pseudo Damage," SAE Journal of Passenger Cars - Mechanical Systems, 13(1):2020

Journal Articles Under Review:

1. Altmann, Craig T., and Ferris, John B. "Identification of Damaging Road Events Using Pseudo Damage Density," SAE Journal of Passenger Cars - Mechanical Systems, Submitted for Review on April 6, 2020

Published Conference Proceedings:

1. Altmann, Craig T., and Ferris, John B. "Customer Usage Based on Pseudo Damage." Proceedings of the ASME 2017 Dynamic Systems and Control Conference. Volume 3. Tysons, Virginia, USA. October 11-13, 2017.

2. Pereira, Savio J., Altmann, Craig T., and Ferris, John B. "Theoretical Development of a Modified RANSAC Algorithm for Identifying Outliers in Road Surface Data." Proceedings of the ASME 2017 Dynamic Systems and Control Conference. Volume 2. Tysons, Virginia, USA. October 11-13, 2017.

Presentations at Professional Conferences:

1. Altmann, Craig T., "Customer Usage Based on Pseudo Damage.", Presentation, ASME 2017 Dynamic Systems and Control Conference, Presentation, Tysons, VA

2. Altmann, Craig T., Ferris, John B., Mergenmeier, Andy, "Calibration, Certification, and Verification of Transverse Profile Measurements", 2019, Webinar, Transportation Pooled Fund - TPF-5(299)

3. Altmann, Craig T., and Ferris, John B., "TPF-5(299) - Calibration, Certification, and Verification of Transverse Profile Measurements", 2019, Presentation, Road Profiles Users' Group and Pavement Evaluation Conference, Roanoke, VA

4. Altmann, Craig T., "Using Localized Pseudo Damage as a Characterization Tool for Longitudinal Profiles", 2019, Presentation, Road Profiles Users' Group and Pavement Evaluation Conference, Roanoke, VA

5. Altmann, Craig T., and Ferris, John B., "Calibration, Certification, and Verification of Transverse Pavement Profile Measurements", 2019, Presentation, Road Profiles Users' Group and Pavement Evaluation Conference, Roanoke, VA

6. Altmann, Craig T., "Road Surface Measurement Spacing Based on Delaunay Triangulation", 2019, Presentation, Road Profiles Users' Group and Pavement Evaluation Conference, Roanoke, VA

1.6 Outline

This work is organized as follows. Chapter 1 includes a motivation of the research and presentation of the scope of the work, problem statement, thesis statement, and main contributions. In Chapter 2 recent literature pertaining to road measurement and characterization, current state-of-the-art regarding damage prediction for vehicle-terrain interactions, common damage methods, and current state-of-the are in clustering algorithms are provided. A novel method of identifying the contribution of discrete road excitations to the total accumulated damage of a vehicle, referred to as Localized Pseudo Damage (LPD), is developed in Chapter 3. In Chapter 4 a method for identifying damaging events in a road profile using LPD across multiple vehicle load paths is developed. In Chapter 5 a similarity test between damaging events is developed using known damage characteristics for each event and exercised with applications using clustering algorithms. Lastly, in Chapter 6, applications of this research, conclusions which can be drawn for this dissertation, and areas for future work are presented.

Chapter 2

Background

Chapter 2 provides general background information for the development in this dissertation. First, typical road roughness measures used to characterize measured road surfaces are presented. Following, current methods for predicting vehicle damage and the shortcomings of the process are presented. Next methods for modeling roads using a power spectral density (PSD) are provided. The process of performing cycle counting on stress/strain/load time history to define a set of stress/strain/load amplitudes is provided. Next, damage and life prediction models are provided along with the development of pseudo damage. A brief overview of methods for detecting localized damaging events is provided along with the shortcoming of these methods. Lastly, current state-of-the-art in clustering analysis is introduced.

2.1 Roughness Measures

Measuring and characterizing road surfaces are a critical part of performing any vehicle-road interaction analysis. Therefore, before discussing different road roughness models, a few words about measuring road surfaces are presented. Terrain measurement started with using vehicle-response systems that was commonly towed behind a vehicle [5, 6]. The issue with using a vehicle-response unit is that the data collected was vehicle dependent. In more recent times, vehicle-independent systems have been developed and are frequently used [7, 8]. These vehicle independent systems typically utilize optical/laser measurement units along with methods for canceling out the body motion to determine the relative height of the road profile. It is assumed in this work that road surfaces are properly measured and available for further analysis.

2.1.1 Roughness Indices

In order to categorize road profiles roughness indices were developed. Roughness indices are used to specify the unevenness of road profiles relative to a vehicle response. One important artifact of using roughness indices is that it is not necessary for two roads to have the same profile to achieve the same roughness index. A roughness index is measure representing the typical roughness of a road surface and this measure can be achieved through an infinite number of road profiles. Below are some of the common roughness indices that are used along with some brief details.

International Roughness Index (IRI)

IRI is considered the standard index for determining the road roughness [9, 10]. Road agencies across the word commonly use the IRI to monitor pavement health because of it can be calculated independent of the vehicle equipment being used [11]. The IRI is calculated by simulation of the suspension travel of the 'Golden Quarter-Car Model' and dividing the resultant suspension travel by the longitudinal distance traveled, giving the IRI units of mm/km or in/mi.

Ride Condition Rating (RCR)

The Ontario Ministry of Transportation (MTO) uses a system that collects the root-mean-square (RMS) vertical acceleration of a trailer axle. These vertical accelerations can then be used to determine the RCR. The RCR is a simple rating from 0 to 10 on the roughness of the road. A rating of 10 is a smooth road while a rating of 5 or 6 indicates a rough road. In modern times MTO has since switched to using the IRI because it does not require the use of a response-type road roughness measurer like the trailer axle used. In addition, it has been shown that the IRI is better at predicting the perceived roughness of road profiles over the RCR method, thus making the IRI a better roughness index.

Full-Car Roughness Index (FRI)

Capuruço developed the FRI because he claimed that the a full-car model would more accurately represents vehicle responses than a quarter-car model [11]. The FRI utilizes four quarter-car models connected together by a rigid rectangle symbolizing the front axle, rear axle, and longitudinal chassis bars. The FRI model restricts the four quarter-car models to only vertical displacements. Details of how the FRI is calculated are provided by Capruruço [12]. In different studies performed it has been found the the FRI results in less variance and higher repeatability in the resulting roughness values than the IRI, but at the cost of higher

computational efforts. The FRI should be chosen over the IRI for higher fidelity simulations if computational time allows.

Root Mean Square of Terrain Height ($\mathrm{RMS_H}$)

The IRI is insensitive to roughness caused by certain wavebands, therefore the root mean square of terrain heights was developed. The root mean square of terrain height (RMS_H) is useful because it uses an RMS approach to analyze the importance of short, medium, and long road unevenness effect on the vibration of the vehicle [13, 14]. It is important to decompose the road into these different wavelengths because the vehicle response differently to different wavelengths. RMS_H can be calculated using Eq. (2.1).

$$RMS_H = \sqrt{D_H} \tag{2.1}$$

In Eq. (2.1), D_H is a value of the combined variance of the random and harmonic components of a road profile [13].

Ride Number (RN)

The RN roughness index is a spinoff of the IRI and is suggested that it does a better job at identifying the condition of the pavement over the IRI. The RN can be calculated for a single or double wheel-track. Details of how the RN is calculated are provided by Sayers [15]. The bias and random error associated with the RN method is on par with that of the IRI calculation when optical and laser profilers are used.

Roughness Index for Driving Expenditure (RIDE)

RIDE is based off of the sprung mass acceleration to the response of a road profile and is intended to be used as a relationship between pavement roughness and user cost [11]. RIDE is calculated from the power spectral density (PSD) of the acceleration of the sprung mass and full details of this calculation are provided by Papagiannakis [16]. The RIDE index is claimed to have several advantages over other indices [11]:

- Direct relationship to riding comfort
- Related to the dynamic axle loads and reflects heavy vehicle ride and damage
- Fully compatible with ISO standard 2631
- Sensitive to road excitation frequencies near the resonant frequencies of the sprung mass

One issue with the RIDE method is that it has a strong emphasis on the sprung mass of the vehicle which could cause it to yield flaws in the actual condition of the pavement depending on the choice of tire, tire quality, and the state of the vehicle suspension [11].

2.1.2 International Roughness Index Shortcomings

Out of the roughness indices presented above IRI is the most widely accepted roughness measure. While the IRI is a reasonable measure for identifying general pavement roughness that can cause the suspension to move and indicate the need for maintenance it does not come without its shortcomings. The IRI is a good measure for identifying areas of roadways that need repair because it is a measure that is independent of the vehicle collecting the data and identifies areas that cause large suspension travel to a 'Golden Quarter Car' model. The problem with using the IRI in a process to develop target roads based on the damage caused by customer usage is that in calculating the IRI it is assumed that

1. Vehicle damage is linearly related to accumulated suspension travel.
2. An average measure of damage is sufficient to capture the complexity of road surfaces.
3. A LTI 'Golden Quarter Car' model simulated at 80 kph represents all vehicles on all roads.

These three major issues with using the IRI to develop target roads for customer usage analysis are addressed in turn.

Vehicle Damage is not Linearly Related to Accumulated Suspension Travel
The issue for vehicle engineers is that fatigue damage is not simply linearly proportional to suspension travel. According to Basquin's relation the damage experienced by a part is exponentially related to the load; for vehicle components this exponent is usually between 3 and 8 [2].

Average Measures of Damage are Insufficient
Due to this assumption there can be a multitude of roads that fit into IRI groups that do not produce similar levels of pseudo damage. For example, take three roads, one that is smooth aside from a large pothole, one that has low frequency long wavelength swales, and one that has high frequency small amplitude excitations. Each of these road profiles has similar IRI values because the average roughness is the same, but they excite different components of the vehicle. The pothole excitation will likely have majority of the loads travel through the jounce bumper and the control arms. The second road will likely have most of the loads carried through the springs. The third road will have most of the forces carried through the dampers. The different load paths will have a significant effect on the resulting pseudo damage of the vehicle and of specific components.

A LTI 'Golden Quarter Car' is Insufficient
The parameters for an LTI 'Golden Quarter Car' model were chosen such that they model a typical passenger car [17, 18]. Therefore, when analyzing the damage associated with a vehicle that does not fall into the select category a different model is needed. In recent work done by Zamora Alvarez vehicles other than the Golden Quarter car are simulated at speeds other than 80 kph in an effort to achieve a better roughness measure. Specifically Zamora Alvarez created a Discrete Roughness Index (DRI) which allows for vehicle models other than the Golden Quarter Car model to be used [4].

2.2 Predicting Vehicle Damage From Road Roughness

2.2.1 Estimating Customer Usage Through Road Roughness

Improving the modeling and simulation of vehicle-road interaction prior to production of a vehicle has always be desired. In 2002 Ferris and Larson developed a method for identifying the 95th percentile damage of a vehicle based on road roughness. In their work they defined the 95th percentile damage by the 95th percentile IRI, International Roughness Index, value. Now this method has a short coming and that is using the IRI as a measure of damage. As shown by Zamora-Alvarez IRI does not contain a one-to-one relationship with damage making it difficult to relate an IRI value to damage [4]. Since Ferris and Larson's development of determining the 95th percentile damage new methods for determining the vehicle damage based on road profiles have been developed. One method, outlined below, is using the suspension force(s) to calculate the pseudo damage of the vehicle.

2.2.2 Calculating Vehicle Damage Through Pseudo Damage

Bogsjö and Rychlik found that not only was the roughness important when predicting the vehicle fatigue but also the sharpness of the curves, occurrence rate of irregularities, degree of roughness, etc. [19]. They found that the majority of fatigue damage is caused by local areas of a higher degree of roughness, therefore they modeled the fatigue accumulation process for Gaussian loads with superimposed transients due to sudden short changes of surface variability [19]. With these assumptions the true stress that the vehicle exhibits will not match the calculated stress, but the results can be used to quantitatively compare designs. The first step in Bogsjö and Rychlik's method is finding the fatigue damage due to an irregularity. After finding the damage due to a single irregularity the expected damage by a road with several irregularities is found, where the total damage to the vehicle is defined in Eq. (2.2).

$$D_{total} = \sum_{j=1}^{N+1} D_j^{stat} + \sum_{j=1}^{N+1} D_j^{irr} \tag{2.2}$$

In Eq. (2.2) D_j^{stat} is the damage due to a stationary section at j and D_j^{irr} is the damage due to an irregularity at j. The process of calculating the pseudo damage is outlined in [20] and can be summarized as followings. In addition, a more detailed review of the pseudo damage process is provided in Section 2.5. Miner's rule can be used to estimate the total fatigue damage caused by a given sequence of loading events. This sequence can be used to calculate the damage, D, through Eq. (2.3).

$$D = \sum_i \frac{1}{N_i} \tag{2.3}$$

In Eq. (2.3), N_i is the number of load cycles to failure at load index i. In order to find N_i from some input load profile the Basquin Relation can be used, which is summarized in Eq. (2.4), where C is a material constant, U_i is the ith load amplitude, and β is the fatigue exponent. The fatigue exponent β usually takes a value in the range of 3-8 for vehicle components [2].

$$N_i = C^{-1} U_i^{-\beta} \tag{2.4}$$

In 2011 Zaabar and Chatti used a similar method of calculating the pseudo damage as Bogsjö and Rychlik used, above. Zaabar and Chatti were interested in estimating the effect of road roughness and its relationship to repair and maintenance costs. The analysis contained the following steps which were repeated for each road roughness value [20].

1. Generate a synthetic road profile
2. Use numerical modeling to estimate the vehicle response
3. Use the rainflow counting algorithm along with the Miner's rule to compute the damage to the vehicle suspension

For Zaabar and Chatti's method they generated the synthetic road profiles following the approach developed by Cebon, see Section 2.3 for details on this method of road modeling. To compliment the road model Zaabar and Chatti used a quarter-car vehicle model. A quarter-car model was used was because the vertical mode is the dominant component of vibration produced by road irregularities [20]. After generating the synthetic road profile and performing the numerical simulation of the vehicle they found that the accumulating of damage is not linear in relation to the IRI used. Figure 2.1 illustrates their findings for a vehicle traveling at 60 mph and 70 mph.

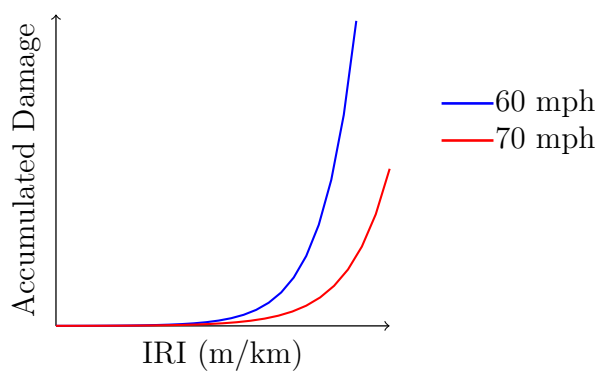


Figure 2.1: Effect of IRI value on the accumulate vehicle damage [20]

Although there were numerous assumptions and averages performed in this study, it is clear that the calculated pseudo damage is not linearly related to the road roughness. In addition the damage is dependent on the speed of the vehicle. In the IRI effect on the damage the results are approximately exponential, $y = ae^{ax}$, where the table below summarizes the values for the two different reported speeds.

Table 2.1: Exponent values from an exponential fit to the report data

Speed	a	b
60 mph	0.0005991	2.183
70 mph	0.0005725	1.823

2.3 Road Modeling Through Power Spectral Density

Modeling road profiles using a power spectral density (PSD) approach has been around since the 1950s. One of the first published uses of spectral analysis in a road application was in 1955 by Houbolt *et. al* which was later improved in 1962 by Houbolt. Over the years different methods for defining the PSD approximation have been proposed and developed. In use today is a linear approximation, a specific version of which was already used in 1955. The specific approach used in 1955 is shown in Eq. (2.5), as defined in terms of spatial frequency n and scaling constant C.

$$G_d(n) = Cn^{-2} \tag{2.5}$$

Since 1955 the exponent of -2 has been replaced with different values in order to better approximate the straight-line fit. It was not until 1995 that a standard, ISO 8608, was developed for road modeling using a PSD. This standard is covered in more detail below.

2.3.1 ISO 8608

In ISO 8608 the process for forming a single and multi-track profile is outlined along along with an agreed upon standard of what PSD to use and smoothing the PSD data. When generating a single-track profile either the displacement PSD, G_d, can be used and/or the acceleration PSD, G_a, can be used [21]. The displacement PSD is simply the PSD of the vertical road profile displacements and the acceleration PSD is the is the PSD of the rate of change of the vertical road profile per unit distance traveled. When using the displacement or acceleration PSD it should be plotted against the spatial frequency, n, or the spatial angular frequency, Ω, on a log-log scale [21]. The two methods, displacement PSD and acceleration PSD, can be related through Eq. (2.6).

$$\begin{aligned} G_a(n) &= (2\pi n)^4 G_d(n) \\ G_a(\Omega) &= \Omega^4 G_d(\Omega) \end{aligned} \tag{2.6}$$

For finding multiple-track profiles the same PSD approach for the single-track approach is used and the two profiles can be related through their coherence [21]. The equation for the coherence, y, is defined as Eq. (2.7).

$$y^2 = \frac{G_{12}(.)^2}{G_1(.)G_2(.)} \tag{2.7}$$

In Eq. (2.7), $G_{12}(.)$ is the cross spectrum between the first and second tracks, $G_1(.)$ is the PSD of the first track, and $G_2(.)$ is the PSD of the second track.

2.3.2 Summarization of the PSD Approximation Methods

Over the years individuals have found that the ISO 8608 standard for approximating the power spectral density of roads to be inadequate when modeling profiles they had collected. For this reason they developed different methods for approximating the PSD of the road profile. These approximation methods are summarized in Table 2.2.

Table 2.2: Methods for approximating the PSD of road profile roughness spectral

Method Name	PSD Approximation	
ISO 8608	$G_d(n) = Cn^{-w}$	$0 \leq n \leq \infty$
BSI	$G_d(n) = \begin{cases} Cn^{-w_1} \\ Cn^{-w_2} \end{cases}$	$0 \leq n \leq n_0$ $n_0 \leq n \leq \infty$
Two Split	$G_d(n) = \begin{cases} Cn^{-w_1} \\ Cn^{-w_2} \\ Cn^{-w_3} \end{cases}$	$0 \leq n \leq n_1$ $n_1 \leq n \leq n_2$ $n_2 \leq n \leq \infty$
Sayers	$G_d(n) = \frac{C_1}{n^4} + \frac{C_2}{n^2} + C_3$	$0 \leq n \leq \infty$
Gillespie	$G_d(n) = C\left(1 + \left(\frac{0.066}{n}\right)^2\right)/n^2$	$0 \leq n \leq \infty$
Marcondes *et al*	$G_d(n) = \begin{cases} C_1 e^{-kn^p} \\ C_2(n - n_0)^q \end{cases}$	$0 \leq n \leq n_0$ $n_0 \leq n \leq \infty$
Sussman	$G_d(n) = \frac{C}{\alpha^2 + n^2}$	$0 \leq n \leq \infty$
Macvean	$G_d(n) = \frac{C}{(\alpha^2 + n^2)^2}$	$0 \leq n \leq \infty$
Sussman	$G_d(n) = \frac{C\left(n^2 + \alpha^2 + \beta^2\right)}{(n^2 + \alpha^2 + \beta^2)^2 + 4n^2\alpha^2}$	$0 \leq n \leq \infty$
Xu *et al*	$G_d(n) = A/\,(2a)e^{\left(-n^2/(2a)^2\right)}$	$0 \leq n \leq \infty$
Kozin and Bogdanoff	$G_d(n) = A/\,a\,e^{\left(-n^2/a^2\right)}$	$0 \leq n \leq \infty$

From Table 2.2 it is easy to see that the approximations have a large variation in complexity and the amount of information that is needed in order to generate road profiles. For the BSI and the Two Split method it is trivial to see that they are closely based on the ISO 8608 standard, but instead of using a single line to approximate the PSD when plotted in a log-log format they use two or three linear lines, respectively. For the remaining approximations the resulting equations were typically formed based on the best fit to gathered road data which

means that they worked for the specific set of road profiles that the respective individual(s) had collect, but they many not be proper approximations to all road profiles.

2.4 Cycle Counting

When load cycles are cyclic and constant amplitude determining the relative load amplitude of the component is fairly simple, since the load amplitude is a constant value calculated from the difference between the maximum and minimum values of the load time history. Unfortunately, in practice most components, especially those on vehicles, seldom experience constant amplitude cyclic loading. For this reason a counting scheme must be used to make a series of constant amplitude events from a series of variable amplitude events. Typical cycle counting can be broken up into two types: one-parameter cycle counting and two-parameter cycle counting methods.

One-Parameter Cycle Counting Method
This type of cycle counting is the easiest to perform and understand, but it is inadequate for fatigue damage analysis and is typically not used. One-parameter cycle counting is inadequate because it does not capture the load cycles based on completion of a loading hysteresis loop, and thus fails to link the loading cycles to the local strain-strain behavior [22]. Some common examples of one-parameter cycle counting are level crossing, peak valley, and range counting. While one-parameter cycle counting methods is quick to perform it does not provide the accuracy desirable and is typically not useful for the type of damage analysis of interest in this work, fatigue damage analysis. Thus, this method will not be discussed in any further detail.

Two-Parameter Cycle Counting Method
This type of cycle counting method is usually used when performing fatigue damage analysis on a complex non-cyclical load-time history. Two-parameter cycle counting identifies unique cycles in the time history by using a defined algorithm for estimating hysteresis loop clusters. Common two-parameter cycle counting algorithms work in a methodical order to identify cycles throughout the load-time history. When a hysteresis loop is identified by the algorithm a resulting constant load amplitude is identified allowing for a damage estimate across the cycle to be defined. The primary example of two-parameter cycle counting is the rainflow counting method. There are many different ways that the rainflow counting method can be performed, the commonly used methods are covered in the following subsections.

2.4.1 Introduction to Rainflow Counting

The rainflow counting method was developed by Endo. The original document was written in Japanese and was translated into English by Matsuishi and Endo. Endo's definition of the method involved a complicated recursive algorithm that involved finding the closed

hysteresis loops in the stress-strain plane [23]. It is well known that rainflow counting method corresponds to cyclic strain-strain behavior where through cyclic loading closed hysteresis loops are used to calculate damage. Figure 2.2 shows the relationship between collected stress-time data and its conversion to the stress-strain domain. In Figure 2.2, the left plot window illustrates the collected stress-time data with peak stress values labeled 1-7. The center plot window illustrates how the 7 stress peaks translates to hysteresis loops in the stress-strain domain. Lastly, the right plot window illustrates the stress amplitudes of the three resulting closed hysteresis loops from the stress time history. The stress amplitude is defined as the difference between the maximum and minimum stress values for a closed hysteresis loop.

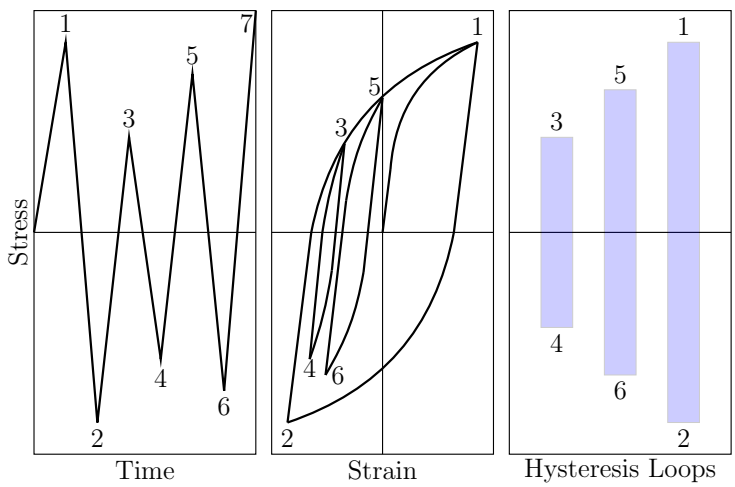


Figure 2.2: Relationship between stress-time and hysteresis loops

Most rainflow counting algorithms are based on one of two concepts. The first concept is the 'availability matrix' and the second concept is the 'vector' approach. The 'availability matrix' algorithm was developed by Wetzel [24]. In Wetzel's algorithm the input signal is divided into a finite number of bins which contains the range and mean of each reversal. Each bin corresponds to a single element in the 'availability matrix'. The 'availability matrix' is then used to determine when a rainflow cycle is formed. The 'vector' based approach uses an array to keep track of the peak and valley that did not form a closed loop [25]. As soon as a closed loop is identified the associated peak and valley is removed from the array. The first non recursive method was proposed by Rychlik.

Two commonly used methods for performing rainflow counting are the three-point cycle counting method and the four-point cycle counting method. The terms three and four refer to the number of load points that are used to determine if a cycle occurs. Both of these methods are outlined in greater detail below.

2.4.2 Three-Point Cycle Counting Method

The process for performing the three-point rainflow cycle counting method has primarily been outlined in three different locations: SAE handbook, ASTM standard E1049, and by Rice *et al.* [26, 27]. The three-point cycle counting uses three consecutive points in a load-time series to identify a cycle. Figure 2.3 shows the three types of events that can occur in a load-time history. The red shaded region in the center and right events illustrates the two ways a cycle occurs when using the three-point cycle counting method. In addition, Figure 2.3 shows that if the change between points 1 and 2, $\Delta S_1 = |S_1 - S_2|$, is larger than the change between points 2 and 3, $\Delta S_2 = |S_2 - S_3|$, than no cycle occurs. Therefore, the only way for a cycle to occur is if $\Delta S_1 \leq \Delta S_2$.

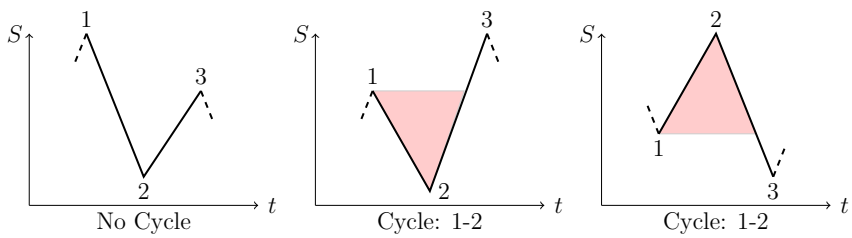


Figure 2.3: The three cycle types that can occur in the three-point cycle counting method

For the process presented in SAE and ASTM the stress-time history must be rearranged so that it only contains the peaks and valleys of the load-time history and has the highest peak or lowest valley (which ever has the largest magnitude) at the start. After rearranging the load-time history the cycle identification can be applied to identify three consecutive points that generate a cycle, like the center and right side events in Figure 2.3. Once a cycle is found the load amplitude is recorded, the two points that constitute the cycle (i.e. points 1 and 2 in Figure 2.3) are discarded, and the remaining load points are connected together. This process is repeated until all of the data is exhausted. A schematic outlining this process can be found in Appendix A.1.

Rice developed a different method for performing the three-point cycle counting. This method is referred to as the range-pair counting method. The method is the same as the SAE and ASTM counting methods outlined above but with one change, the data is not rearranged. This means that for the range-pair counting method the load-time history does not need to be altered before cycle counting is performed.

2.4.3 Four-Point Cycle Counting Method

The four-point cycle counting method is similar to the three-point cycle counting method, but uses four consecutive points instead of three. The four-point cycle counting method

is considered the most generic rainflow counting method. In fact, the three-point cycle counting method can be derived from the four-point cycle counting method. In addition to being generic the four-point method is useful because there is no need to rearrange the load-time data and has the ability to be implemented to extrapolate cycles and reconstruct load-time history. Figure 2.4 illustrates the two types of cycles that can occur in four-point cycle counting. Only through one of these two load-time history arrangements can it be ensured that a hysteresis loop closure occurs.

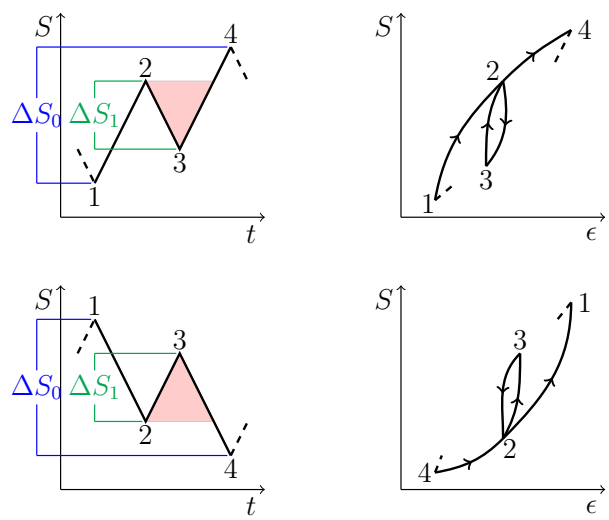


Figure 2.4: Two cycle types that can occur in the four-point cycle counting method

A cycle is counted if the inner stress range is less than or equal to the outer stress range ($\Delta S_1 \leq \Delta S_0$) and the two points that constitute the inner range are bounded by the two points that constitute the outer range. The cycle that is counted is the inner cycle. Therefore, when a cycle occurs points 2 and 3 are removed and points 1 and 4 are connected. If $\Delta S_1 > \Delta S_0$ then a cycle does not occur. If this happens then point 1 is omitted and the next point after point 4 is added. The cycle counting process is repeated until no countable cycles remain. Unlike the three-point cycle counting method the four-point cycle counting method can have data points that do not form closed cycles. If this occurs then these remaining points are stored in the residue. A standardized method for handling the residual is covered in the section below, Section 2.4.4. A schematic outlining the process for performing the four point cycle counting can be found in Appendix A.2.

2.4.4 Standardization of the Rainflow Counting Algorithm

The rainflow counting algorithm is used by many different industries to determine the damage accumulated by a wide range of parts all of which have different applications. For this reason, it was recognized by Amzallag *et. al* that some standard was needed so that rainflow

cycle counting was the same throughout out all industries. The basic outline of rainflow counting can be broken up into the following categories: treatment of the sequence (initial data), extracting the cycles, performing the rainflow counting algorithm, and dealing with the residue. This section covers a standardization developed by Amzallag *et. al* to perform rainflow counting.

Treatment of the Sequence

Before cycle extraction occurs, preliminary treatment of the sequence is necessary. The preliminary treatment consists of sampling the data, extracting the local extrema, and placing the values into defined classes. In order to store the values in a method that is targeted towards speed of the analysis the sequence is divided into k classes of constant-width intervals and all of the maximum and minimum values in each of the classes are replaced by the mean value [28]. The standard practice for generating the constant-width classes is to set the number of classes equal to 64 [29, 30]. When performing the four-point cycle counting method the maximum number of residue points depends on the number of classes, this relationship is that the maximum number of points cannot exceed $2k - 1$.

Handling the Residue

Residue of the rainflow counting algorithm is the left over load-history points that do not complete a full cycle. When dealing with the residue there are two common processes that can be implemented depending on the application. The first method is decomposing the cycles and the second is storage of the residue to allow for sequence reconstruction [28]. Both of these methods are discussed in greater detail below.

In method one the residue is decomposed in order to accurately predict the fatigue life of the part being analyzed. The standard technique used to pull the cycles from the residue is to add the residue to itself and apply the rainflow technique to the two joined residues [28]. This method is used because the resulting final residue is identical to the first residue. This property is outlined in Eq. (2.8), where [residue] is the complete set of stresses that are left over after the first time performing the rainflow counting algorithm. The reason this method is used is because of its simplicity. The same procedure is applied repeatedly until the whole sequence is decomposed into cycles.

$$[\text{residue}] + [\text{residue}] = [\text{residue}] + \text{cycles} \tag{2.8}$$

In method two, the residue is stored, as is, to be used to reconstruct a sequence using the cycles identified during the rainflow counting that resulted in the residue. In general, the process used to reconstruct the load time history is the reverse of the procedure used to extract the cycles [28]. One rule that must be followed when inserting cycles into a reconstructed sequence is that the peaks of the cycle must be entirely included within the peaks where the cycle is inserted [28]. The one issue with the reconstruction method is that the reconstructed sequence is not unique because there are multiple places a cycle may be inserted. One important limitation of the reconstruction method is that it cannot be used to truncate high-amplitude cycles [28]. The reason for this is because small cycles contained

in the high cycles transitions cannot be properly connected with the other portions of the sequence.

2.4.5 Keeping the Load Sequence

One issue with the cycle counting methods covered thus far is that the order in which the load sequences occur are not kept. This can pose an issue for components that have load sequence effects where changing the peak/valley locations in the load-time history can have a severe effect on the life of the component. In order to keep the load sequence order, Anthes developed a new method that keeps the sequence order. The order is maintained by counting every rising half cycle and treating it as a damaging event. If the half cycle is not closed by a hysteresis loop it is classified as a virtual hysteresis loop [31]. There are three different cases for a virtual hysteresis loop:

1. The virtual histories will be closed by a falling half cycle in the future. For this case there is no further operations that need to be performed because the damage associated has already been accounted for.
2. The raising half cycle that is treated as a virtual hysteresis loop will continue on later. If this occurs then a new virtual hysteresis loop is created that is larger. The damage is then calculated by finding the damage associated with the larger virtual hysteresis loop and subtracting out the damage of the smaller virtual hysteresis loop.
3. The virtual hysteresis loop will not close until failure occurs. Even if this does occur the damage calculated for the virtual hysteresis loop is still valid.

The first two forms of virtual hysteresis loops are shown in Figure 2.5. The the top of Figure 2.5 illustrates a falling half cycle is illustrated by the complete hysteresis loop closure when the stress-time history falls below the stress amplitude at the first load reversal (point a) after reaching the peak stress of the hysteresis loop at point b. The bottom of Figure 2.5 illustrates a raising half cycle because a small hysteresis loop is formed by the stress reversals at points b and c and a new larger stress reversal occurs at point d which may lead to a larger stress hysteresis loop, and a complete cycle, between points a and d later in time. Otherwise, the load reversals at points a and d form a half cycle.

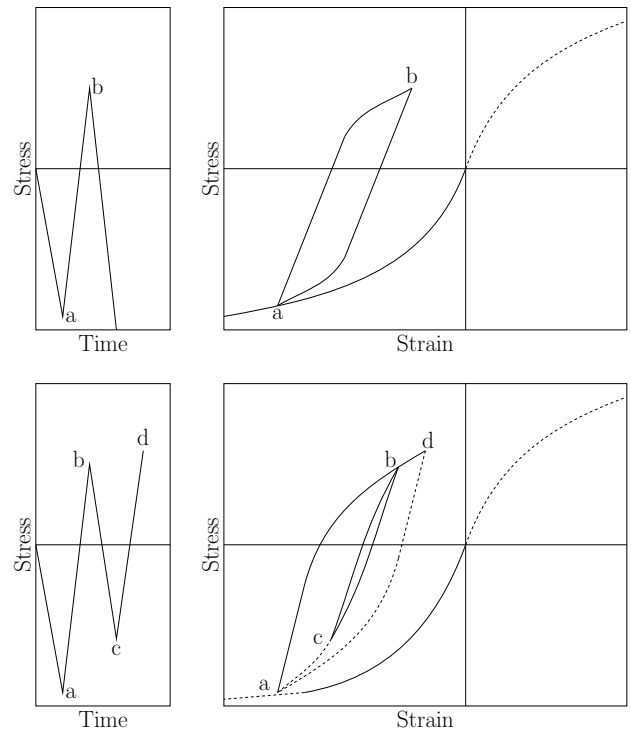


Figure 2.5: Examples of case 1 and 2 virtual hysteresis loops [31]

The algorithm for performing the rainflow counting while keeping the loading sequence developed by Anthes is presented in a schematic format similar to the prior rainflow counting methods in Appendix A.3. For this method the algorithm contains two residual stacks, the stress and strain values, that are initially empty but are to be filled with the half cycles that are not closed. In addition there are two counters used for the organization of the two stacks. The first counter, IZ, points to where the last starting point of a half cycle is located. This means that IZ represents the entire size of the stack. The second counter, IR, points to where the first starting point of where a half cycle is capable of forming a closed hysteresis loop [31].

The stress values are found using a cyclic stress-strain curve relationship, FC, and an equation for the half cycle, FH. The cyclic stress-strain curve relationship is commonly found using the Ramberg-Osgood formulation [32], shown in Eq. (2.9).

$$\epsilon = \frac{\sigma}{E} = \left(\frac{\sigma}{K'}\right)^{\frac{1}{n'}} \tag{2.9}$$

where E is the modulus of elasticity, K' is the stress hardening coefficient, and n' is the strain hardening exponent. Reforming Eq. (2.9) it can be written in terms of half cycles as shown in Eq. (2.10).

$$\Delta\epsilon = \frac{\Delta\sigma}{E} = 2\left(\frac{\Delta\sigma}{2K'}\right)^{\frac{1}{n'}} \tag{2.10}$$

For Anthes' method to work the maximum and minimum stress-stain values of the real or virtual hysteresis loops must be known. Additionally, the damage model must take into account the maximum strain value of the history. If the maximum strain value for the history is not known the first stress reversal point would not be considered, thus causing discrepancies. Three different damage events are recorded by Anthes modified method.

1. Closed hysteresis loops
2. Rising half cycles that form a closed hysteresis loop
3. Rising half cycles that do not form a closed hysteresis loop due to the last load reversal considered

In Anthes' method, the complete stress-strain path is used, except for the part following the rising cyclic stress-strain curve for first load reversal in time. Due the use of the complete path, the integer number of damaging events, n_d, can be written in terms of the number of reversals, n_r based on the first peak in the stress-strain path being positive or negative as shown in Eq. (2.11).

$$\begin{aligned} \textbf{Positive First Peak:}\ n_d &= \mathrm{INT}\left(\frac{n_r - 1}{2}\right) \\ \textbf{Negative First Peak:}\ n_d &= \mathrm{INT}\left(\frac{n_r}{2}\right) \end{aligned} \tag{2.11}$$

The reason Anthes modified method is better than the classical rainflow counting method is because of the way it handles load sequence effects. It is well known that load sequence effects can have a huge effect on damage accumulation. These load sequence effects can be influenced by the preceding load history along with the actual crack opening level [31]. Typically models that taken into account load sequence are based on the plastic deformation around the crack tip, but Anthes modified rainflow counting method keeps the load sequence without using crack propagation analysis.

Anthes method will show improvements over the classical rainflow counting method for stress-strain input data where fatigue damage significantly relies on load sequence effects. For other input data using Anthes model might not produce results that are more accurate than the classical rainflow method. Significance of load sequence effects for input data depends on three factors: load-time sequence, crack length, and the structure. Some of these factors (e.g., load sequence effects and crack growth) are addressed in Section B whereas other factors (e.g., component structure) are application specific and are therefore not discussed in detail in this work.

2.5 Damage Prediction and Pseudo Damage

The development in this work is focused on estimation of vehicle durability when limited information is known about the components in the vehicle. In this section, current state

of the art in pseudo damage, usage of load-time history to estimate accumulated damage, is presented. To calculate the damage a method for modeling the accumulation of damage is needed. Typically Miner's linear damage rule is used to model accumulation, however other non-linear methods are introduced. Following, a method for modeling the fatigue limit for constant amplitude load cycles is presented. Next, the combination of Miner's rule and Basquin's relation is presented to determine the total pseudo damage in a load path. Lastly, an introduction to more advanced damage and life prediction theories that focus on material properties is provided for completeness.

2.5.1 Damage Accumulation Models

It is believed that one of the first damage theories was developed in 1933 where H.J. French reported investigation of the over stress effect on the endurance limit [33]. In 1937 Langer proposed that the damage process can be separated into stages: crack initiation and crack propagation [34]. Just shortly after Langer's proposal, in 1938, Kommers suggested that the change in endurance limit could be used as a damage measure [35]. Prior to these three individuals Palmgren had proposed the Linear Damage Rule (LDR) in 1924. But, it wasn't till 1945 that Miner first expressed this concept in the mathematical format shown in Eq. (2.12). In Eq. (2.12) the total damage , D_T, is expressed as a summation of the fraction of the number of applied cycles, n_ℓ, and the total number of applied cycles under the ℓth constant amplitude loading level, N_ℓ.

$$D_T = \sum_{\ell} \frac{n_\ell}{N_\ell} \tag{2.12}$$

The LDR is commonly used because of its simplicity of calculating the damage based on the cycle ratio with the two basic assumptions that a constant amount of work is absorbed per cycle and there is a characteristic amount of work absorbed at failure [36]. Through the relationship between work and energy it is possible to see that there is an accumulation of energy and a linear summation of damage can be used.

The main issues with the LDR is its load-level dependence, inability to account for load sequences, and its lack to account for load-interaction [36]. Since the original development of the LDR individuals have attempted to fix these issues. In 1949, Machlin proposed a method based off of the cumulative damage theory which utilized metallurgy [37]. Throughout the 1950s Coffin *et al.* used a strain based LDR approach in order to correlate their experimental results [38, 39]. Lastly, in 1970 Miller showed how LDR could be used in an application for strain-controlled fatigue damage analysis [40]. In 1948, Richart and Newmark introduced the damage curve (D-r) diagram and speculated that at different stress levels the D-r curves would be different [41]. Marco and Starkey were the first to proposed a nonlinear load-dependent damage model in 1954 [36]. The damage was represented by a power a power

relationship as shown in Eq. (2.13) where x_ℓ is a quantity related to the ℓth loading level [42].

$$D_T = \sum_\ell \left(\frac{n_\ell}{N_\ell}\right)^{x_\ell} \tag{2.13}$$

Figure 2.6 shows this power relationship. In the figure, four different values for the power were used to understand the effect of the power rule. The first value for the power is $x_\ell = 1$, which is equivalent to the Miner Rule (or LDR). The following three lines in Figure 2.6 illustrate the damage accumulation affect at higher exponent values ($x_\ell > 1$). In the figure four points are highlighted and a dashed line is drawn between the points. The dashed line between points 1 and 2 and again between 3 and 4 indicates that when x_ℓ increases, then a higher cycle ratio is needed to accumulate the same amount of damage. The dashed line between points 2 and 3 indicates that at a constant cycle ratio when x_ℓ increases a lower damage accumulation is expected. Thus, Miner's rule tends to be the safest method for estimating the damage curve, especially when specific material information is not known.

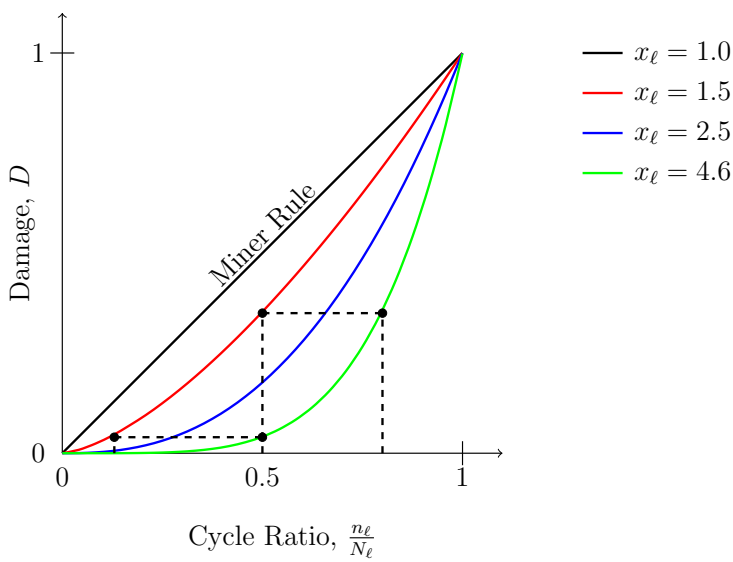


Figure 2.6: Representation of the damage curve for the Marco-Starkey theory

2.5.2 Basquin Relation and Total Pseudo Damage

In the Section 2.5.1 all presented damage accumulation models are defined as a function of the total number of load amplitudes, N_ℓ, that can be achieved prior to failure at a constant load amplitude cycle. A commonly used method for estimating N_ℓ that typically models both low and stress fatigue well is the Basquin relationship presented in Eq. (2.14) [43]. In Basquin's model C is a material coefficient and β is the fatigue exponent.

$$N_i = C^{-1} \left|L_i\right|^{-\beta} \tag{2.14}$$

Since the focus of this work is on pseudo damage, the damage analysis is load path dependent and therefore used as a relative measure, therefore C is typically 1. The value of the fatigue exponent is dependent on the application. For vehicle components the fatigue exponent is typically between 3 and 8 [2]. When the Basquin damage model is combined with the Palmgrin-Miner Linear Damage Rule, the total damage for a load-time history can be expressed as a summation of individual load amplitude as shown in Eq. (2.15).

$$D_T = C \sum_{\ell} |L_\ell|^\beta \tag{2.15}$$

2.5.3 Typical Pseudo Damage Calculation Process

In general the following three step process is used to determine the total pseudo damage accumulated:

1. Determine a load time history for a load path of interest through vehicle simulation or through instrumentation of a vehicle.
2. Select a cycle counting method and identify a set of load amplitudes to characterize the load-time history.
3. Use the combination of Miner's rule and Basquin's method to estimate the total pseudo damage accumulated.

This process is presented graphically in Figure 2.7 as a flow chart. In the flowchart it is clear that the force-time history is used as an input into any selected cycle counting method. The output of the cycle counting method is load amplitudes which are inputs into a damage model. By modeling the damage using Basquin's damage relationship and Miner's rule a total pseudo damage for a given load-time history is established.

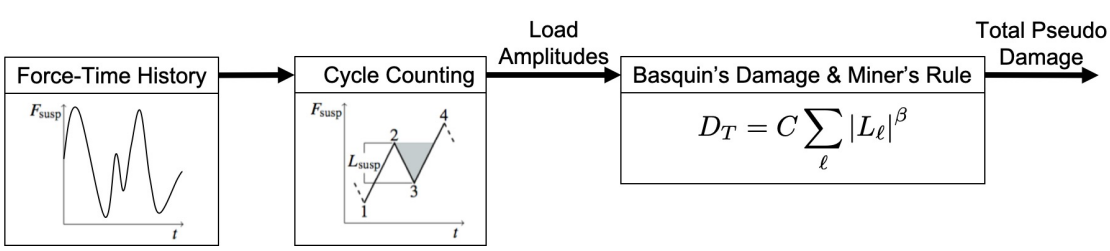


Figure 2.7: Flow diagram of the pseudo damage calculation process.

2.5.4 Stress/Strain Damage and Life Prediction

There have been many additional damage and life prediction theories developed based on experimental data and testing. While damage prediction is an old problem and has been

studied by many individuals, it has yet to be fully resolved. In the sections prior, an overview of the current state-of-the-art in estimation of damage using load-time history has been presented. However, to provide a complete review of current damage and life prediction models stress and strain analysis were investigated. Since these developments focus on material specific information, they are not included in the main portion of this dissertation and have been moved to Appendix B for reference.

2.6 Detection of Localized Road Event

Using a PSD to represent a road profile was first developed in 1973 by Dodds and Robson [44]. One of the issues with using a PSD to model a road profile is that the data must be homogeneous, or better stated wide sense stationary (WSS) when dealing with the time signal. If these criteria are not meet for the collected data the resulting PSD will produce an average that is not truly representative of the road profile and thus may not meaning. For this reason it became apparent that there was a need for identifying local transient events which do not meet this criteria. Once the transient events are identifying and removed from the road profile the remaining data meets the criteria to be able to calculate the PSD. The following sections cover methods proposed to identify these localized transient events.

2.6.1 Spatial Acceleration

In 1999, Bruschella *et al.* showed that PSD of the road elevations does not expose portions of the road surface that are non-Gaussian, non-stationary, or transient [45]. In this work Bruschella *et al.* showed that transient events can be easily found using the spatial acceleration instead [45]. In addition to being able to more easily identify the events using spatial acceleration, identifying transient events based on spatial acceleration is more representative of the severity of vehicle vibration [45].

Rouillard *et al.* implemented the use of spatial acceleration to characterize road surfaces. When implementing the method of spatial acceleration Rouillard *et al.* used two criteria to identify transients portions of the road profile. The first criteria was that transient events have short moving mean square drop-off distance [46]. The second criteria was that transient events have a sufficiently large local crest factor for the spatial acceleration data [46]. Using these criteria and developed methodology regions which contained non-Gausian characteristics were identified. These regions were typically between 0.25m and 0.5m long [46].

2.6.2 Median Filtering

Kropáč and Múčka decided that median filter method would be a practical approach for identifying localized events based on its ability to identify and smooth large large obstacles [47]. A median filter was chosen because median filters exhibit two important properties that make them a useful for identifying faults in road profiles. The first property is that median filters remove the significant irregularities in the road height without having an effect on the randomness of the surrounding parts of the road profile [47]. Secondly, if a discrete step to a new steady level is present (e.g. transition between road surfaces) a median filter will not affect the random nature of the surrounding road profile like a moving average filter will [47].

In addition to using a median filter, Kropáč and Múčka characterized obstacles using six indicators: obstacle length, maximum obstacle height, ratio of the obstacle length to its max height, obstacle profile area, obstacle volume, and distance between two successive obstacles [47]. Through this characterization and median filter approach Kropáč and Múčka were able to identify large obstacles with reasonable accuracy when they were the six indicators were in a desired range.

2.6.3 Local Areas of High Error

In more recent years alternative approaches like using Markov chains [48] and Hilbert spaces [49] have shown promising results. Construction of a Hilbert space has the inherent ability to identify localized events. This ability is because an error function is defined to decide how many basis vectors are needed to model the road profile. Therefore, local portions of the road profile that have high error components after enough basis vectors are added can be concluded to be localized events.

2.7 Clustering Analysis

When multiple events are known, clustering analysis is used to group events into a set of discrete non-empty groups such that events which are more similar to each other are placed in one group/cluster and events that are dissimilar are placed in other groups/clusters. The objective of clustering analysis to to highlight useful information or trends in a set of objects with similar features which may have otherwise gone unnoticed [50]. Over the last decade clustering analysis has experienced great success across a variety of application.

Over the last decade clustering analysis has experienced great success across a variety of applications. Clustering analysis has four main goals: (1) Development of a typology or classification scheme, (2) investigation of useful conceptual schemes for grouping entities, (3) hypothesis generation through data exploration, and (4) hypothesis testing [51]. Clustering analysis aims to categorize a data set with some known features into clusters where

each cluster contains high similarities while low similarities exist between clusters. Two main forms of clustering algorithms are partitional and hierarchical clustering. Partitional clustering is the division of a set of data into non-overlapping clusters such that each data object is in exactly one cluster and each cluster contains at least one object [52, 53]. Hierarchical clustering is the development of a hierarchy between multiple data objects [54]. Both of these clustering analysis methods are discussed in greater detail below.

2.7.1 Partitional Clustering

Partitional clustering is the division of a set of data into non-overlapping clusters such that each data object is in exactly one cluster and each cluster contains at least one object [52, 53]. Given a data set with N objects, the objective of partitional clustering is to devise k clusters, where $N \geq k$ [55]. Typically the objective of partitional clustering is to minimize an objective function. In many common partitional clustering algorithms the objective is to minimize the distance between an object and the center of a cluster [55]. For a data set with individual objects referenced as O_i for the ith object and N total objects in the dataset for partitional clustering the objective function, ζ, can be defined in general terms in Eq. (2.16) where $dist$ is any selected distance (or similarity) function and C_j is the center of the jth cluster [56].

$$\zeta = \sum_{i=1}^{N} dist\left(d_i, C_j\right) \tag{2.16}$$

Two commonly used partitional clustering algorithms are K-means and K-medoids. While these two algorithms vary in the method of solving and implementation, the same generalized representation of the results can be used to highlight these clustering methods. Figure 2.8 provides a generalized representation of partitional clustering. In Figure 2.8, an open circle is used to indicate the cluster center and each black dot indicates an object from the date set. The black lines in Figure 2.8 are used to help illustrate the three clusters that are formed from clustering the data set.

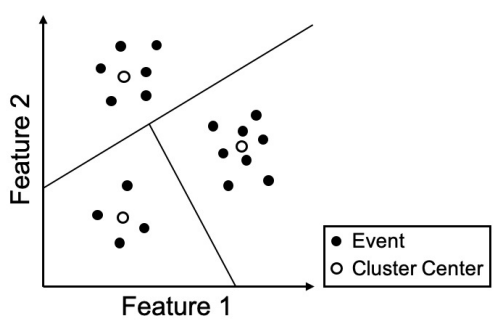


Figure 2.8: Generalized representation of partitional clustering.

In the K-means algorithm, each cluster is represented by the mean value of the points in the cluster. In Figure 2.8 the open circle would be the corresponding mean value for each cluster. Instead, for the K-medoids algorithm each cluster is represented by one of the points located near the center of the cluster, commonly referred to as a medoid [57]. In Figure 2.8 each open circle would be representative of this medoid. In short, the K-medoids algorithm differs from the K-means in that the centers of clusters are defined by the user rather than letting the algorithm define the centers.

2.7.2 Hierarchical Clustering

Hierarchical clustering is a form of clustering analysis where the objective is to form a hierarchy of clusters. The process of performing hierarchical clustering can be considered in three steps: (1) Evaluation of a distance between each pair of points, (2) Grouping of the points into a binary hierarchical tree, (3) separating/cutting the hierarchical tree into clusters [54]. This process is developed in general terminology as follows. For a data set two objects shall be identified (m and n). For each object a set of features that characterize the object shall already be established. Using this features a distance between objects can be calculated. Using the distance between objects, the next step is to begin form a hierarchical tree between objects using the relative distance between objects to identify at what level two or more events should branch together.

An example hierarchical clustering results is provided in the form of a dendrogram in Figure 2.9. In this example eight objects are present in the dataset. The x-axis indicates each of the objects and the y-axis indicates the level of dissimilarity between the objects. Thus, horizontal lines near the bottom of the y-axis indicate low levels of dissimilarity, while horizontal lines near the top of the y-axis indicate high levels of dissimilarity. Thus, as you follow a branch of the tree from the bottom to the top the events which are most similar and should likely be clustered together are identified by the horizontal links near the bottom while events should likely should not be clustered together have horizontal links near the top. For example, Objects 3 and 4 are more similar to each other than they are to Object 2 or Object 1. In addition, Objects 3 and 4 and not very similar to Object 5, 6, 7, or 8 since those four events completely separate branch that is not linked until the highest dissimilarity level.

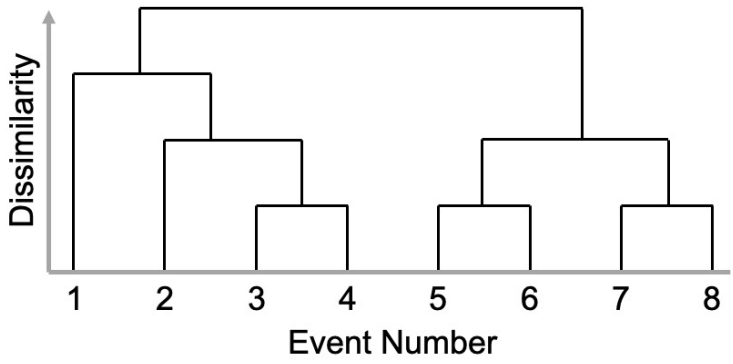


Figure 2.9: Generalized representation of hierarchical clustering.

2.7.3 Clustering Validation

For any clustering algorithm, it can be difficult to decide the number clusters to break the data up into when there is no *a priori* knowledge regarding the characteristics of the dataset. To decide the number of clusters to form either an external or internal cluster validation can be performed. If a set of training data is known where events are predefined to be similar, then a distribution of similarity values can be defined and a confidence on acceptance/rejection of events being similar can be established. This threshold definition can be considered external because an external label/known information is used to establish the threshold. However, when no training data is known, an internal validation method should be used to decide which events are similar to each other. Specifically an internal validation method evaluates the partitioning of a dataset into similar clusters to determine if the partitioning is sufficient [58]. When performing internal validation of a partitioned dataset two criteria that should be evaluated are: separation of the clusters and the compactness of each cluster [59]. Several internal validation methods exist [60] and have been successfully used to determine the optimal partition of a dataset. One of these measures was the S_Dbw metic, which is a linear summation of the intra-cluster variance and the inter-cluster density [59].

Chapter 3

Theoretical Development of Localized Pseudo Damage

3.1 Introduction

When predicting vehicle durability, it is critical to accurately model both the vehicle and to excite the model with real-world damaging road events. Without accurate representation of these damaging events the calculated vehicle damage may not be representative of the damage a user experiences. The ability to accurately collect high-resolution road surface measurements continues to improve. However, simulation of these surfaces can take significant computational effort. Specifically, for durability analysis, it is likely that the majority of the damage accumulated is attributed to a few localized events. Thus, it would be useful if these events could be identified a priori allowing for simulation of only the most damaging road segments, thereby resulting in a savings of computation time.

Not only would a method for identifying damaging road segments be useful for simulation, it would be useful for physical testing. Targeted testing can be conducted on the most damaging segments to reduce testing time. Additionally, if damaging events can be found from the results of a road surface measurement then proving ground testing can be improved. Durability cycles can be constructed based on real-world damaging events and proving ground events can be updated or improved to better represent the damaging events on road surfaces experienced by drivers.

The objective of this work is to develop a measure of the amount of pseudo damage each individual road excitation contributes to the total accumulated pseudo damage. The first

challenge in identifying the damage contribution of individual road excitations is that vehicles are dynamic systems; hence at any given instant the response is dependent, at varying levels, to all prior excitations. This challenge is amplified for damage analysis because damage models are nonlinear and require cycle counting methods that depend on previous peak responses. The measure developed in this work is referred to as Localized Pseudo Damage (LPD) and captures the causal relationship between the vehicle dynamics and the road surface excitations. The theoretical development in this work makes no vehicle, speed, or road modeling assumptions, resulting in a measure that is applicable to all vehicle and road models.

This chapter is organized as follows. First, background on pseudo damage, calculation of pseudo damage rate, and current state-of-the-art is provided. Next, a constraint is established on any LPD model: that for an impulse or step excitation, the model must attribute all of the local damage to the singular event itself. The LPD model proposed in this work is then developed in general terms of vehicle suspension force, allowing for varying vehicle speed and longitudinal spacing of road profile points. Three examples demonstrate the properties and usefulness of the LPD. First, that the proposed LPD satisfies the constraints for impulse and step responses. Second, that although the damage rate is non-decreasing, the LPD may be negative, indicating damage-reducing road excitations. Third, the simulation results of a measured road surface in Tennessee demonstrates the capabilities of the LPD measure on a real-world road surface and usefulness in identifying the most damaging road events. Lastly, three rejected LPD models are briefly discussed, followed by a discussion of future applications, concluding remarks, and references.

3.2 Background

In applications where limited vehicle information is known, pseudo damage can be used as a measure of relative vehicle damage from a known road surface [2, 19]. Pseudo damage is different from conventional damage and fatigue in that the force-time history is used in place of the stress/strain-time history. For vehicle applications, in which the pseudo damage associated with road interactions is desired, the suspension force is used as the input to the pseudo damage model.

Total pseudo damage can be modeled using the Palmgrin-Miner linear damage accumulation method and the Basquin relation is used to estimate the number of cycles to failure [20]. The results of the coupled Palmgrin-Miner and Basquin damage models is presented in Eq. (3.1). To calculate the total pseudo damage, D_T, a set of load amplitudes, L_ℓ, must be known along with a fatigue exponent, β, and a damage scaling factor, C_D.

$$D_T = C_D \sum_{\ell=1}^{N_\ell} |L_\ell|^\beta \tag{3.1}$$

When the pseudo damage model in Eq. (3.1) is applied to vehicle pseudo damage, the load amplitude is most commonly based on an irregular suspension force time history which requires a cycle counting method. The most general method for identifying cycles is the four-point cycle counting method [22]. Most recently, Altmann and Ferris [62] used the four-point cycle counting method to calculate the total pseudo damage associated with synthetic road profiles. The primary issue with the four-point cycle counting method, or other simplified forms, is that the time associated with a cycle is not considered. Disregarding time leads to load sequence affects and greater uncertainty and error in the pseudo damage estimate.

The Top-level-Up Cycle (TUC) counting method developed by Rychlik to analyze the statistical properties of damage estimates allows for damage to be assigned to specific points within a cycle rather than over a complete cycle [63, 64]. The TUC counting method is different from conventional cycle counting methods (e.g. three-point and four-point cycle counting) in that it is not a sequential process. Instead, the TUC counting method maintains the order of the load reversals during the process of defining hysteresis loops allowing for a relationship between damage and time. In 2009, Bogsjo and Rychlik showed that pseudo damage rate (rate of change of pseudo damage over time) can be calculated using the Palmgrin-Miner and Basquin damage model shown in Eq. (3.1) when the load amplitude is defined using the TUC counting method.

Consider an example of a suspension force time history shown in Figure 3.1 where the suspension force at time index j is $F_{susp,j}$. Four examples of $F_{susp,j}$ are shown in the figure as open circles. For any particular $F_{susp,j}$, consider the prior suspension force oscillations. At some prior time, the suspension force will have crossed the particular value $F_{susp,j}$ (except for the special case when j is near the beginning of the time history). The time of the most recent prior crossing of $F_{susp,j}$ is denoted as time index j^+ and shown as a cross in the four examples in Figure 3.1. The dotted lines in Figure 3.1 reference the most recent prior crossing with a particular value $F_{susp,j}$. In the special case when there is not prior crossing of $F_{susp,j}$, the time index $j^+ = 0$. Finally, consider the suspension forces that occur in the time interval (j^+, j). The local peak force, $F^p_{susp,j}$, associated with any particular $F_{susp,j}$, is defined as either: the minimum suspension force in the time interval (j^+, j) if $F_{susp,j}$ is increasing, or the maximum suspension force in the time interval (j^+, j) if $F_{susp,j}$ is decreasing. The local peak force, $F^p_{susp,j}$, for each of the four examples is shown as a solid dot in Figure 3.1.

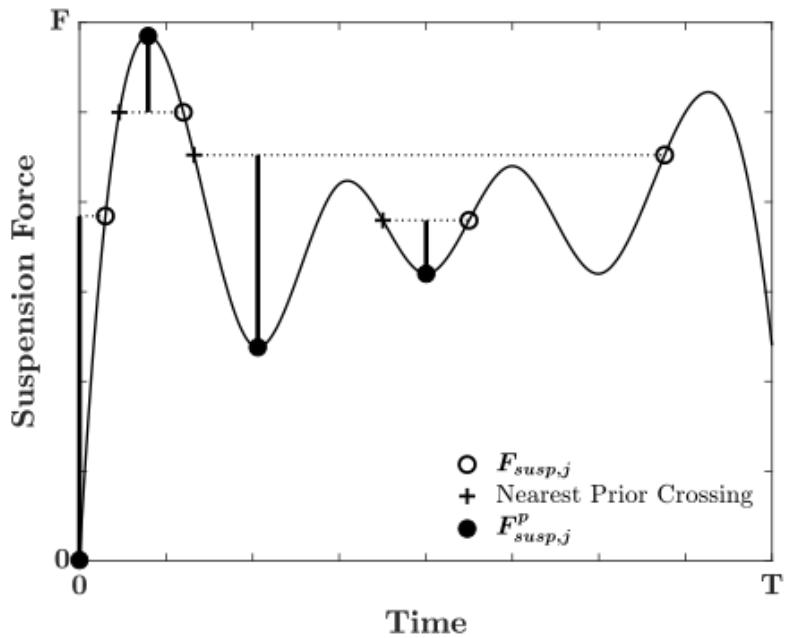


Figure 3.1: Identification of peak suspension forces.

By identifying prior local minima/maxima values as defined in the TUC counting method the time history of irregular data can be maintained during cycle counting. The resulting definition for the pseudo damage rate based on the TUC counting method is provided in Eq. (3.2).

$$\dot{D}_j = G|F_{susp,j} - F^p_{susp,j}|^{(\beta-1)}|\dot{F}_{susp,j}| \tag{3.2}$$

For identification of damaging events from a road profile, there are two commonly used methods: identification based on instrumentation of a vehicle and estimation of pseudo damage based on vehicle responses due to the road surface (typically suspension travel as defined by roughness measures). For identification based on an instrumented vehicle, sensors (e.g. accelerometers, potentiometers, or force transducers) are used to collect data used to determine the total pseudo damage. Once data has been collected for a set of roads, the critical portions of the data can be used to develop testing cycles by matching the critical portions to existing data from proving ground data or using the critical portions as input signals to laboratory and computer testing. Matching or extraction of data (typically forces or accelerations) can be performed in the time domain [65, 66] or the frequency domain [67, 68], depending on the characteristics of interest. Recently, Ferreira et al. compared multiple time and frequency domain fatigue models and provided examples where both models provided comparable results [69]. However, they noted that the time and fatigue domain results did not always agree [69].

For estimation based on vehicle response, when a road surface is modeled as a stationary process with superimposed random irregularities, then a roughness criterion can be used to

identify two categories of irregularities (short and long wavelength) [2]. These locally rough sections can be considered damaging events. Bogsjo and Rychlik developed a model for estimating the expected pseudo damage contribution for road irregularities identified using the roughness criterion [19]. More recently, in a brief investigation performed by Zamora-Alverez it was determined that the relationship between pseudo damage and road roughness, specifically the International Roughness Index (IRI), are not one-to-one [4]. Indicating that estimation of pseudo damage based on a roughness measure is not sufficient for accurately capturing damage.

3.3 Development of Localized Pseudo Damage

The objective of developing a measure of localized pseudo damage (LPD) is to identify the contribution each road profile excitation makes to the total pseudo damage for a given vehicle model and speed. It is assumed that the total accumulated pseudo damage, D_T, is known and that the damage rate at each time step, $\dot{D}_j$, is calculated using Eq. (3.2) where the suspension force is a function of the vehicle, speed, and road profile. The discretized pseudo damage rate equation, for irregular time steps, Δt_j, is presented in Eq. (3.3).

$$D_T = \sum_{j=0}^{N-1} \Delta t_j \dot{D}_j \tag{3.3}$$

The LPD, d_i, is defined as a damage density. Consider an irregularly spaced grid of road surface points in which spacing at the ith location is Δu_i. The damage produced in this interval is then $D_i = \Delta u_i d_i$ and the discretized pseudo damage density equation, for irregular longitudinal steps, Δu_i, is presented in Eq. (3.4).

$$D_T = \sum_{i=0}^{N-1} \Delta u_i \dot{d}_i \tag{3.4}$$

Defining the LPD as a damage *density* helps mitigate the sensitivity of the measure to differences in longitudinal spacing (assuming the spacing is small enough). For example, consider two cases of sampling the same road: the spacing in the second case is half the spacing of the first. The number of points in the second case doubles, but the spacing is halved so that the resulting damage density in the two cases is similar.

Before continuing to the particular model of the LPD developed in the work, consider a constraint to be satisfied by any LPD model with respect to impulse and step responses. Specifically, consider an impulse or step excitation at the first road location, where $i = 0$, resulting in the subsequent accumulation of pseudo damage. Since no additional changes in the road profile exist after $i = 0$ to affect the vehicle body motion, the total pseudo damage must be attributed to the location of the impulse or step excitation. That is, the initial

pseudo damage must be equal to the total pseudo damage, which is written in terms of damage density in Eq. (3.5).

$$d_0 = \frac{D_T}{\Delta u_0} \tag{3.5}$$

To begin developing the LPD model proposed in this work, first consider that vehicle systems respond, at varying levels, to prior road excitations. There is a relationship between the position of the excitation, denoted by subscript i, and the subsequent response, denoted by subscript j. The LPD for each excitation is modeled in this work as a linear combination of subsequent responses. Specifically, the relative contribution of the ith road excitation to the pseudo damage density at the jth time is defined as, $\alpha_{i,j}$. The LPD is modeled in terms of the relative contributions as presented in Eq. (3.6).

$$d_i = \sum_{j=0}^{N-1} \alpha_{i,j} \tag{3.6}$$

Since the LPD is modeling the pseudo damage of a causal vehicle system, the LPD model must also be causal. Thus, the ith road excitation cannot contribute to the pseudo damage that occurred at a prior time and, conversely, the damage rate at a given instance is not affected by future events. Thus, the responses in Eq. (3.6) must satisfy the equality presented in Eq. (3.7).

$$\alpha_{i,j} = 0 \quad \forall j < i \tag{3.7}$$

One of the main contributions of this chapter is the means by which the relative contribution of the ith road excitation to the response at the jth time, $\alpha_{i,j}$, is calculated. Recall that the pseudo damage rate at the jth time is written as $\dot{D}_j$. Now, consider the effect of omitting the ith road excitation on the vehicle response, specifically the pseudo damage rate. In this work, the pseudo damage rate at the jth location that would occur if the ith road excitation is omitted is defined as $\dot{D}_j^{(-i)}$. That is, to calculate $\dot{D}_j^{(-i)}$, the ith road excitation is omitted and the damage rate time series is recalculated using a double time step of $2\Delta t$ at one point in the calculation: when stepping from point $i-1$ to point $i+1$. It should be clear that the time step must be small in order to achieve negligible effects on both the temporal and spectral content of the response. Next the difference between the original damage rate and the damage rate response from the omission of the ith road excitation is multiplied by a scaling factor, ε_j, for the jth response. The chosen model for the relative contribution, $\alpha_{i,j}$, is presented in Eq. (3.8).

$$\alpha_{i,j} = \varepsilon_j \left(\dot{D}_j - \dot{D}_j^{(-i)} \right) \tag{3.8}$$

To establish the scaling factor, ε_j, an additional constraint is placed on $\alpha_{i,j}$: the change in the damage at the jth location, $\Delta t_j \dot{D}_j$, must be equal to the sum of the relative contributions of all the previous excitations, $\Delta u_i \alpha_{i,j}$. Note that since the responses in Eq. (3.6) must satisfy the equality presented in Eq. (3.7), the scaling factor, ε_j, can be determined as shown in Eq.

(3.9).

$$\sum_{i=0}^{j} \Delta u_i \alpha_{i,j} = \sum_{i=0}^{N-1} \Delta u_i \alpha_{i,j} = \Delta t_j \dot{D}_j \tag{3.9}$$

By substituting Eq. (3.8) into Eq. (3.9), the value of ε_j is established for each jth response as given by Eq. (3.10).

$$\varepsilon_j = \frac{\Delta t_j \dot{D}_j}{\sum_{i=0}^{N-1} \Delta u_i \left(\dot{D}_j - \dot{D}_j^{(-i)} \right)} \tag{3.10}$$

In practice, for a given road profile and set of vehicle speeds, the LPD is calculated as follows

1. An initial set of pseudo damage rates is calculated for all times, indexed by j, based on Eq. (3.2).
2. Each ith road excitation is individually omitted and the resulting set of pseudo damage rates for all times is calculated based on the resulting suspension force-time history.
3. The scaling factor, ε_j, is calculated using Eq. (3.10).
4. The relative contribution, $\alpha_{i,j}$, is calculated using Eq. (3.8).
5. The set of LPD values, d_i, is calculated using Eq. (3.6).

3.4 Examples

Three examples are provided to demonstrate the properties of this LPD model: (1) causality, (2) differentiation between damage inducing and damage reducing excitations, and (3) the applicability of the model to real-world data sets. First it is shown that the constraint given in Eq. (3.5) is satisfied when the road profile is an impulse or step excitation. By satisfying this condition, LPD is causal and capable of accurately capturing discrete events. The second example demonstrates that, although the pseudo damage rate and total pseudo damage must be non-negative, the LPD may be negative. That is, for a given vehicle model and speed, subsequent events may reduce the total pseudo damage caused by previous events. Lastly, the use of the LPD on a measured Tennessee highway is presented to show the applicability of the LPD on real-world data sets. The implications associated with using the LPD measure for a measured road surface are considered in the subsequent discussion section.

3.4.1 Impulse and Step

For an impulse excitation occurring at $i = 0$, the pseudo damage rate resulting from the omission of each ith road excitation is represented by the piecewise function in Eq. (3.11) for

all response times (indexed by j).

$$\dot{D}_j^{(-i)} = \begin{cases} 0 & i = 0 \\ \dot{D}_j & i \neq 0 \end{cases} \tag{3.11}$$

That is, removing the impulse at $i = 0$ removes the damage completely and removing any point other than the impulse has no effect on the damage rate at all. Figure 2 provides the suspension force response (subplot 1) and the resulting damage rate (subplot 2) for a Golden Quarter car when the impulse at $i = 0$ is omitted and when any other excitation is omitted to support Eq. (3.11). The dashed lines indicate when $i = 0$ is omitted and the solid lines indicate when any other excitation is omitted. By omitting the impulse excitation at $i = 0$, no disturbance is present to excite the causal vehicle model resulting in zero suspension force and damage rate over all time. By omitting any other excitation, $i \neq 0$, no change in the excitation is presented resulting in a suspension force time history equivalent to the impulse response and a corresponding damage rate containing non-zero values.

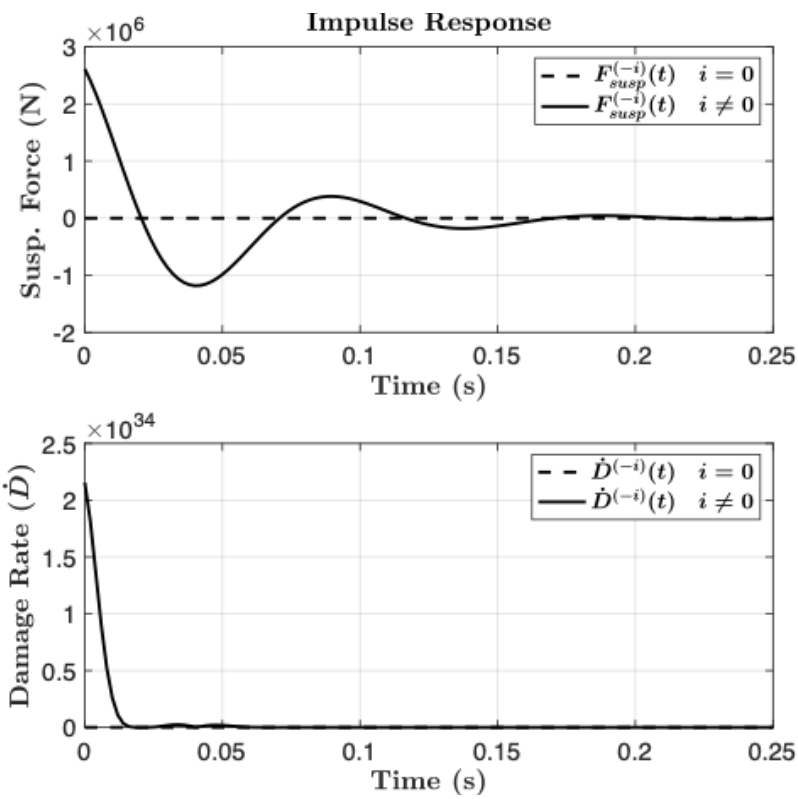


Figure 3.2: Golden quarter car vehicle load for an impulse response and the corresponding damage rate to illustrate Eq. (3.11).

The scaling factor, ε_j, defined in Eq. (3.12) results from substituting the piecewise function

defined in Eq. (3.11) into Eq. (3.10).

$$\varepsilon_j = \frac{\Delta t_j \dot{D}_j}{\Delta u_0 \left(\dot{D}_j - 0\right) + \sum_{i \neq 0} \Delta u_i \left(\dot{D}_j - \dot{D}_j\right)} = \frac{\Delta t_j}{\Delta u_0} \quad \forall\, j \tag{3.12}$$

Using the scaling factor presented in Eq. (3.12), the relative contribution for the impulse excitation can be defined over all time in the piecewise function presented in Eq. (3.13).

$$\alpha_{i,j} = \begin{cases} \frac{\Delta t_j \dot{D}_j}{\Delta u_0} & i = 0 \\ 0 & i \neq 0 \end{cases} \tag{3.13}$$

Using the summation definition for damage rate given in Eq. (3.3), along with the LPD definition in Eq. (3.6), the resulting LPD is equivalent to the total pseudo damage over the initial distance traveled at the impulse excitation and zero elsewhere, see Eq. (3.14).

$$d_i = \begin{cases} \frac{D_T}{\Delta u_0} & i = 0 \\ 0 & i \neq 0 \end{cases} \tag{3.14}$$

Thus, it is shown that the constraint given in Eq. (3.5) is satisfied when the road profile is an impulse.

Now, consider a step excitation at $i = 0$. If the excitation at $i = 0$ is omitted, then the pseudo damage rate response will be delayed by a single time step. When the time step is small, if any excitation is omitted subsequent to $i = 0$, the effect to the pseudo damage rate response will be negligible because all excitations in this region have equal amplitude and no change to the vehicle dynamics would occur. The pseudo damage rate resulting from the omission of each ith road excitation is summarized by the piecewise function in Eq. (3.15) for all time.

$$\dot{D}_j^{(-i)} = \begin{cases} \dot{D}_{j-1} & i = 0 \\ \dot{D}_j & i \neq 0 \end{cases} \tag{3.15}$$

Figure 3 provides the suspension force step response (subplot 1) and the resulting damage rate (subplot 2) for a Golden Quarter car when the step at $i = 0$ is omitted and when any other excitation is omitted to support Eq. (3.15). The solid lines indicate the omission of any excitation, $i \neq 0$, and the dashed lines indicate the omission of the first excitation, $i = 0$, in the step input. No discernible change in the suspension force response is present when excitations which are not $i = 0$ are omitted resulting in a damage rate which is equivalent to the original step input damage rate. However, when $i = 0$ is omitted a delay of one time step is present in the damage rate response; this is highlighted by the shift between the solid and dashed in Figure 3.

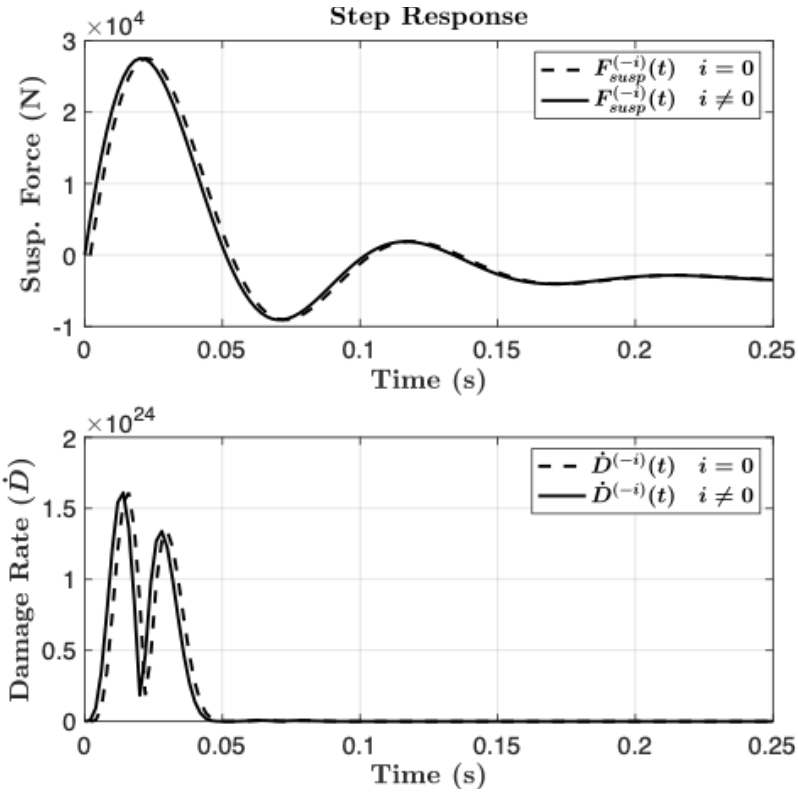


Figure 3.3: Golden quarter car vehicle load for a step response and the corresponding damage rate to illustrate Eq. (3.15).

Using the piecewise function defined in Eq. (3.15), Eq. (3.10) is evaluated for a step response as the time step approaches zero, resulting in the scaling factor presented in Eq. (3.16). In comparison to the scaling factor for the impulse response, Eq. (3.12), the scaling factor for the step response is time dependent.

$$\lim_{\Delta t_j \to 0} \varepsilon_j = \frac{\Delta t_j \dot{D}_j}{\Delta u_0 \left(\dot{D}_j - \dot{D}_{j-1}\right) + \sum_{i \neq 0} \Delta u_i \left(\dot{D}_j - \dot{D}_j\right)} = \frac{\Delta t_j \dot{D}_j}{\Delta u_0 \left(\dot{D}_j - \dot{D}_{j-1}\right)} \tag{3.16}$$

Using the scaling factor presented in Eq. (3.16) the relative contribution can be defined over all time in the piecewise function presented in Eq. (3.17).

$$\lim_{\Delta t_j \to 0} \alpha_{i,j} = \begin{cases} \left(\frac{\Delta t_j \dot{D}_j}{\Delta u_0 \left(\dot{D}_j - \dot{D}_{j-1}\right)}\right) \left(\dot{D}_j - \dot{D}_{j-1}\right) & i = 0 \\ 0 & i \neq 0 \end{cases} \tag{3.17}$$

Using the summation definition for damage rate given in Eq. (3.3), along with the LPD definition in Eq. (3.6), the resulting LPD is equivalent to the total pseudo damage over the

initial distance traveled at the start of the step excitation and zero elsewhere, see Eq. (3.18).

$$\lim_{\Delta t_j \to 0} d_i = \begin{cases} \frac{D_T}{\Delta u_0} & \forall\, j\,,\, i = 0 \\ 0 & \forall\, j\,,\, i \neq 0 \end{cases} \tag{3.18}$$

Thus, it is shown that the constraint given in Eq. (3.5) is satisfied when the road profile is a step excitation or an impulse excitation. No assumption was made regarding the vehicle model being used (a simple model was used for demonstrative purposes). In this way the LPD model is shown to be causal and able to accurately capture the damage characteristics of a discrete event.

3.4.2 Dual Impulses

Pseudo damage is a monotonically non-decreasing function and the pseudo damage rate is always non-negative. The constraint imposed by Eq. (3.4) requires that over a complete road profile, the integration of the LPD over the distance will result in the non-negative total accumulated pseudo damage. However, LPD is defined herein as a combination of subsequent relative contributions, $\alpha_{i,j}$. In contrast to LPD, these relative contributions, as defined in Eq. (3.8), may take on negative values. That is, omitting an event may increase the subsequent damage rate. A road excitation with a negative LPD indicates the excitation aided in reducing the total pseudo damage accumulated; that is to say if the event were not present then the total accumulated pseudo damage would have been greater.

An example of this reduction of accumulated pseudo damage can be produced when two impulse excitations are present. Consider two impulse excitations separated by the distance that would be covered by the vehicle during one-half of the period associated with the wheel hop frequency ($\Delta t \approx 0.05$ sec). In addition, the magnitude of the second impulse is approximately half the magnitude of the first excitation. The timing and magnitude of the second impulse is such that subsequent suspension oscillations and corresponding suspension forces are greatly reduced, resulting in less total pseudo damage than would have occurred had the second impulse not been present. Since omitting the second impulse allows the vehicle to continue oscillating, the pseudo damage continues to accumulate. In this atypical case, a negative localized damage density is properly attributed to the second impulse. Figure 4 illustrates the dual impulse example in which the LPD was calculated using a Linear Time Invariant (LTI) Golden quarter car model [9] traveling at 80 kph with damage rates calculated using a fatigue exponent of $\beta = 5$. The suspension force history for the dual impulse example is provided in the second subplot of Figure 3.4, to illustrate that the second impulse significantly reduced the corresponding suspension force. Notice that in the third subplot of Figure 3.4 the first impulse produces a large positive LPD value indicating a large component of the accumulated damage is due to the first impulse. To better illustrate the LPD value of the second impulse, a more detailed plot of the second impulse is provided in the third subplot of Figure 3.4. This example illustrates that a second

impulse has a smaller amplitude negative LPD value, indicating the second impulse assisted in reducing the total amount of accumulated damage which could have occurred.

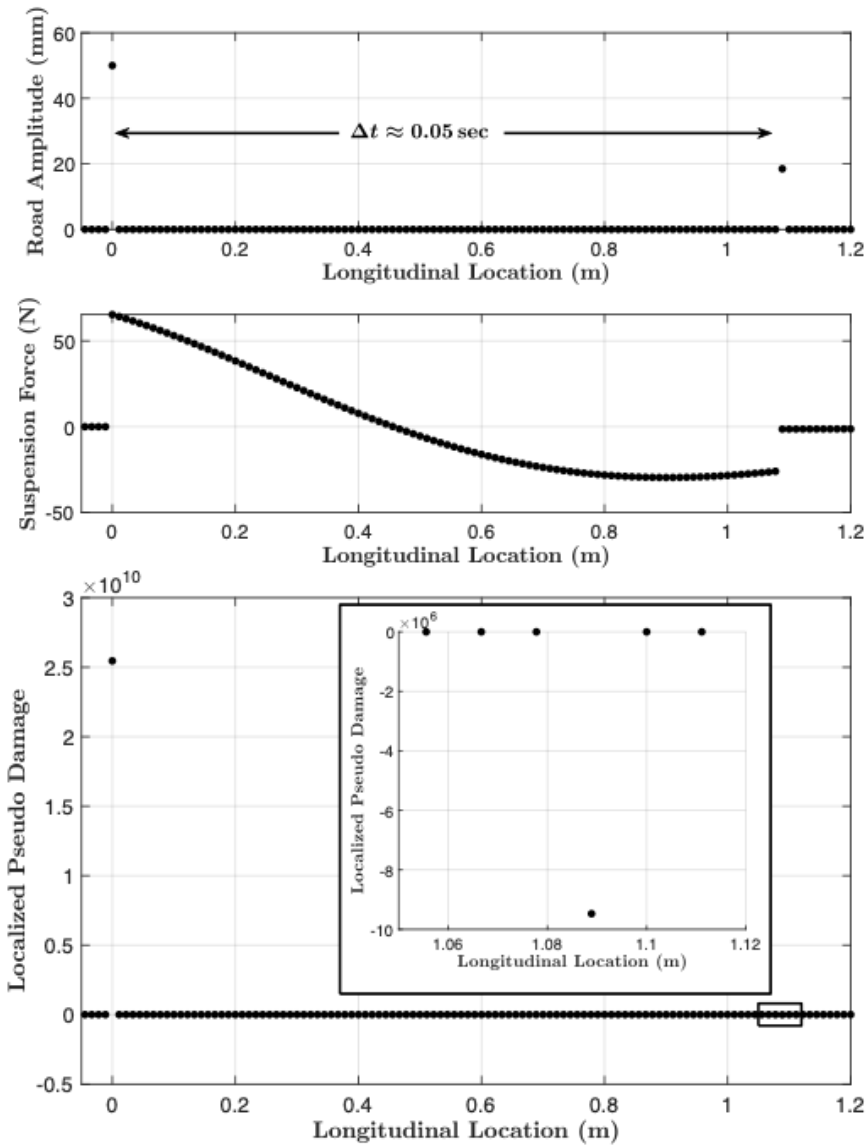


Figure 3.4: Dual impulse excitation illustrating negative LPD.

Dual impulses is a unique example, in that if the two impulses were considered separately, they would be considered damage inducing excitations. However, due to the sequence and location of the excitations, the second impulse *reduces* the amount of damage the vehicle would have experienced from the first impulse had the second impulse not been present. This example highlights an important characteristic of the LPD model: the capability to

distinguish between damage inducing and damage reducing excitations based on the sequence in which the excitation occur. Although both excitations are similar, but with different amplitudes, their damage characteristics are not similar. This is shown by the positive versus negative LPD values in the third subplot of Figure 3.4.

3.4.3 Measured Road Surface

Lastly, a measured road surface is presented to illustrate the capabilities of the LPD model developed herein on a real-world road profile. The technical specifications of the measurement procedures used when measuring the road surface follow those developed by Kern and Wagner [7, 8]. Figure 3.5 shows a 100 m long road profile collected on a Tennessee highway along with the resulting LPD results achieved using a Golden quarter car traveling at a speed of 80 kph with damage rate calculated using a fatigue exponent of $\beta = 5$. The LPD results in Figure 3.5 present several instances where the LPD: (1) significantly deviates from zero, (2) is negative, and (3) is near zero. The instances where there are significant deviations from zero indicate road segments that strongly contribute to the total accumulated pseudo damage. Negative LPD values indicate segments of the road, in the context of the entire road for a given vehicle and speed, reduce the amount of accumulated damage. Lastly, the segments with near zero LPD values indicate segments of the road that do not significantly contribute to the total accumulated pseudo damage.

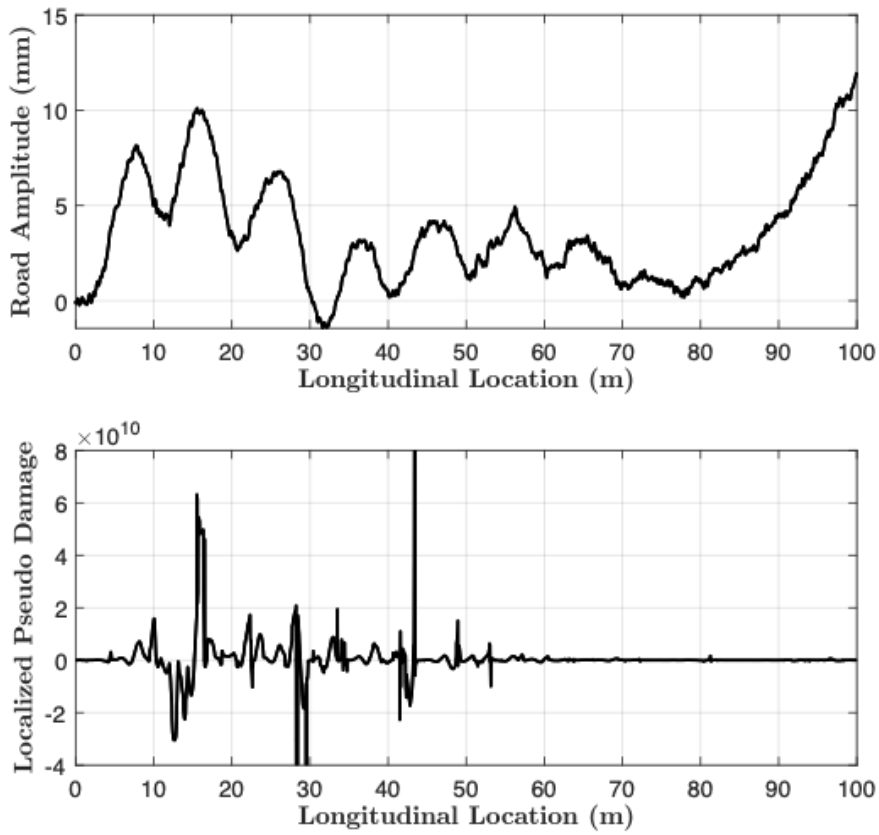


Figure 3.5: Localized pseudo damage for a measured terrain profile.

For the measured road profile in Figure 3.5 there are two significant peaks in the LPD plot (at the 15 m and 43 m longitudinal locations). Evaluation of two subplots in Figure 3.5 reveals that these peaks match up with segments of the road where significant changes in road amplitude occur resulting in increased suspension loads and consequently increased pseudo damage. Both peaks occur over a short duration and contain high damage density (high LPD magnitudes). This characteristic is similar to the LPD results of the impulse and step responses, where a high level of damage density exists only at $i = 0$, the location of the road amplitude change. In addition to the two peak values, the LPD in Figure 3.5 illustrates events which contain slightly lower damage density amplitude but over a longer duration. This is illustrated by the non-negative LPD values between the 8 m and 50 m locations. Identification of both of these types of events is necessary for determining all types of damaging events in a measured road profile.

An important characteristic of the LPD model is that the resulting measures are based on damage contribution, not road roughness. Consider the longitudinal segment of road between 70 and 100 m in Figure 3.5. This segment of road contains a level of roughness which can be visually seen in the first subplot. However, this segment of road does not contain any

excitations with strong damaging contributions therefore the LPD values are near zero, as illustrated in the second subplot. Since this segment of road contains near zero LPD values it can be removed without having a significant effect on the accumulated pseudo damage in the vehicle.

In the dual impulse example, see Figure 3.4, it was shown that a negative LPD value occurs when the vehicle response (e.g. suspension force) is reduced by a secondary excitation. In the measured road surface, presented in Figure 3.5, similar characteristics to the dual impulse can be observed. Figure 6, allows closer inspection of a region where positive and negative LPD values exist between longitudinal location 9.5 m and 10.8 m. Over this region, a set of low amplitude bumps is present. The start of the first bump of interest is highlighted by the left vertical line in a time delay callout and the start of the second bump of interest is highlighted by the right vertical line. The time delay callout is $\Delta t \approx 0.05$ which is equivalent to the time delay in the dual impulse example. The longitudinal distance associated with the time delay is defined using the 80 kph vehicle speed. Based on the LPD plot in the third subplot of Figure 3.6, over the length of the first bump positive LPD values are present, whereas, for the second bump negative LPD values are present. Comparable to the dual impulse, the location of the second bump is such that it results in a reduction of the amount of accumulated pseudo damage which would have occurred had the bump not been at that location. The second subplot in Figure 3.6 illustrates the suspension load to provide support for the reduction in accumulated damage. In the second subplot it can be seen that the second bump resulted in a decrease in suspension force.

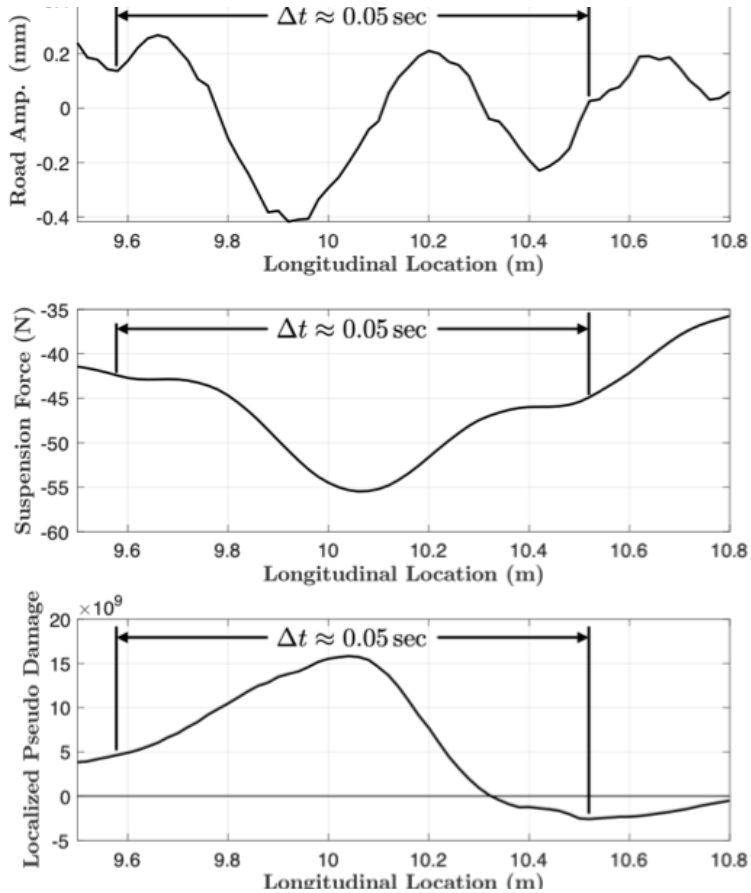


Figure 3.6: Localized pseudo damage and suspension force between longitudinal positions 9.5m and 10.8 m for the measured terrain profile.

3.5 Discrete Sampling Constraint

In practice, road profiles are a set of discrete measurements. By omitting a single measurement, the temporal and spectral content of the vehicle response could be affected, resulting in improper localized pseudo damage values. Since pseudo damage is dependent on capturing peak transient suspension force values, sampling rates that satisfy the Nyquist criteria are necessary, but not sufficient. Consider an idealized road event modeled as a half sine wave with duration T_d equivalent to half the period of the highest ride frequency of interest for the vehicle model (e.g. wheel hop, at about 70ms). The maximum discretization error when discretely sampling the transient response at a period of T_s will occur when the samples are equally spaced on opposite sides of the transient peak. A normalized transient response is provided in Figure 3.7 with discrete sampling points to illustrate the worst case scenario.

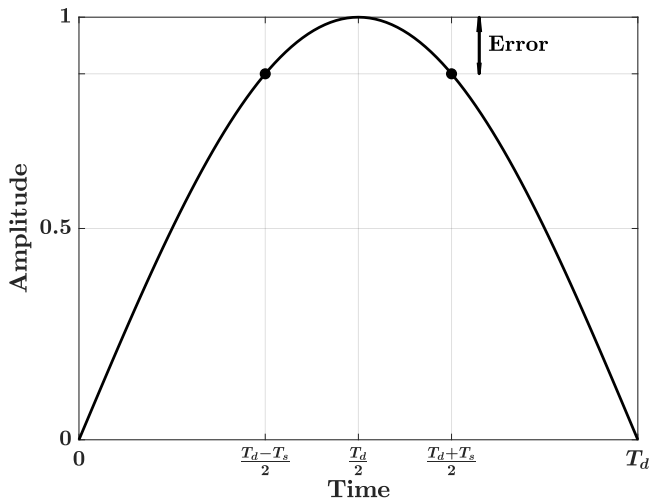


Figure 3.7: Worst case scenario when sampling a normalized transient event.

The maximum error for the normalized transient response can be written in terms of the sampling period and the transient duration, see Eq. (3.19).

$$E = \begin{cases} 1 - \cos\left(\frac{\pi}{2}\frac{T_s}{T_d}\right) & T_s \leq T_d \\ 1 & T_s > T_d \end{cases} \tag{3.19}$$

Using Eq. (3.19), a bounds on the allowable sampling rate can be determined based on the user defined maximum error allowed in the peak magnitude. Once a sampling rate is determined, then the maximum allowable distance between consecutive road profile excitations can be defined based on vehicle speed or, conversely, the maximum allowable vehicle speed can be defined based on a known road profile sampling interval.

It was shown in Section 3.4.1 that for an impulse response the LPD is equivalent to the total pseudo damage at the location of the impulse when the damage rate is a continuous time function. This characteristic can be used to identify a critical sampling time for a discrete vehicle system. The critical sampling time can be identified by calculating the LPD for discrete impulse response and investigating the ratio of the LPD at the impulse location, d_0, to the total pseudo damage, D_T. The discrete time step where the ratio deviates too far from one is considered the critical sampling time for the vehicle system and shall be used to define the required road profile spacing or the maximum allowable simulation speed based on the profile spacing.

To influence the range of discrete sampling times to investigate for a selected vehicle model, the following analysis of an idealized transient force shall be performed using the period of the highest ride frequency of interest for the vehicle model. LPD is determined by omitting individual road indices and recalculating the damage rate. Damage rate is calculated based

on relative suspension force and rate of change of suspension force, which is dependent on the vehicle response to a road surface. To maintain generality of the road surface, consider an idealized transient force modeled as a half sine wave with duration T_d equivalent to half the period of the highest ride frequency of interest for the vehicle model (e.g. wheel hop, at about 70 ms). The, idealized transient force is provided in Eq. (3.20).

$$F_{susp} = \sin\left(\frac{\pi}{T_d}t\right) \tag{3.20}$$

Since the idealized transient force is a half sine wave, the damage rate can be written as shown in Eq. (3.21)

$$\dot{D}(t) = \alpha\beta\left|\sin\left(\frac{\pi}{T_d}t\right)\right|^{\beta-1}\left|\frac{\pi}{T_d}\cos\left(\frac{\pi}{T_d}t\right)\right| \tag{3.21}$$

In conventional transient sampling error evaluations, like those shown earlier in this section, a peak value must be identified, then two transient samples can be equally spaced about the peak. However, the location of peak damage rate (i.e. the peak values in Eq. (3.21)) is dependent on the value of β, however because of the idealized transient force the damage rate peaks will be equally spaced about half the duration, $T_d/2$. Therefore, peak error for a given sampling rate will occur when the samples are equally spaced about $T_d/2$.

The error of interest for the LPD calculations is not just the error of the peak damage rate values, it is the error on the total pseudo damage, the summation of damage densities. Due to this summation criteria the error present in the discrete sampling of the damage rate to the total pseudo damage can be used. Based on the selected value of β and the transient duration, T_d, the integral of Eq. (3.21) can be calculated and used exact value of the accumulated pseudo damage. The error, E, resulting from discretely sampling the damage rate can then be calculated using Eq. (3.22).

$$E = 1 - \frac{T_s \sum_j \dot{D}_j}{\int_0^{T_d} \dot{D}(t)dt} \tag{3.22}$$

Thus, by setting a bounds on error allowed in the total damage estimate an initial estimate of the critical sampling time can be established and used as a guide for discretizing a vehicle-road response.

As an example, a Golden Quarter car model was selected. The wheel hop frequency of the model was selected as the highest frequency of interest resulting in a transient force duration of approximately 48 ms. The maximum error of various sampling rates were calculated using Eq. (3.22) and the results are provided in Figure 3.8. Based on the error reported in Figure 3.8 transient error begins to exceed 10% when the ratio of the sampling rate to the transient duration is approximately 0.2. Therefore, the ratio of 0.2 was selected as an upper bounds on the sampling rate for a Golden Quarter car when the highest frequency of interest is the wheel hop frequency.

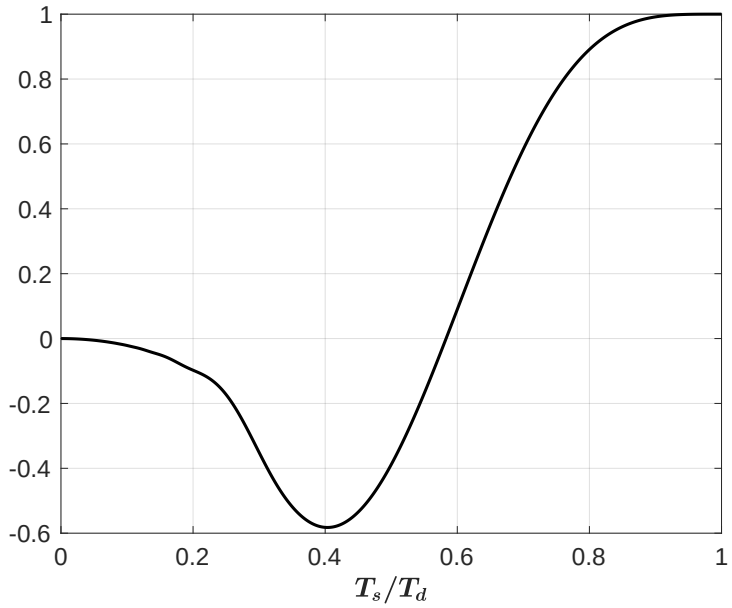


Figure 3.8: Pseudo damage error corresponding to Eq. (3.22) for a Golden Quarter Car model.

The error results presented in Figure 3.8 are based on analysis of the suspension force through a Golden quarter car. If a different vehicle model or load path is used when calculating LPD values, then the error calculation would need to be repeated/validated. By performing this discrete error analysis a required minimum sampling rate for a road profile based on a desired vehicle speed or a maximum speed that can be simulated for a road profile can be established.

3.6 Discussion

During the development of this work, three additional models for the relative contribution, $\alpha_{i,j}$, were considered. It was assumed in the first model that the relative contribution was linear and time invariant for any road profile. Based on this assumption, the relative contribution can be written in terms of a scaling parameter $\gamma_{(j-i)}$, similar in concept to the impulse response, as shown in Eq. (3.23).

$$\alpha_{i,j} = \gamma_{(j-i)} \dot{D}_j \tag{3.23}$$

This model was not selected because the resulting LPD for an impulse or step response was not equivalent to the total damage at the impulse or step location. The second model considered was a first order sensitivity of the damage rate at the jth location to the ith road

height, $z_{r,i}$. Based on this assumption, the relative contribution can be written as the first order sensitivity presented in Eq. (3.24).

$$\alpha_{i,j} = \left.\frac{\delta \dot{D}_j}{\delta z_{r,i}}\right|_{z_{r,i}} \tag{3.24}$$

This model for $\alpha_{i,j}$ was not selected because it did not satisfy the required causality condition. The third model investigated utilized the relative contribution model presented in Eq. (3.8). However, the scaling factor, ε_j, was considered a constant over all time and was prescribed as presented in Eq. (21).

$$\varepsilon_j = \frac{\Delta t_j}{\Delta u_i} \tag{3.25}$$

This relative contribution model resulted in the LPD for a given road excitation being equivalent to the difference between the total accumulated pseudo damage for the original road and the total accumulated pseudo damage when the ith road excitation was omitted over the longitudinal step length. This model was not chosen because it could not identify damage due to step excitations since total accumulated pseudo damage is unaffected for a step excitation regardless of which excitation is omitted.

The contribution of this work is a theoretical framework by which the effect a road excitation has on the accumulated pseudo damage can be identified. The work is applicable to all vehicle models excited by a single road profile. It is envisioned that the LPD developed in this work will be useful for identifying and characterizing events in road profiles which significantly contribute to the accumulated pseudo damage of the vehicle. For identification of events, segments of a road surface with large and sustained LPD values can be identified as a single damaging event (or region of damage). Similarly, road segments with LPD values which are nearly zero for sustained periods of time can be identified as having negligible contributions to the total pseudo damage. That is, a given road profile could be decomposed into multiple damaging events (or regions), separated by non-damaging stretches of road. Furthermore, these damaging events could be characterized by the shape of their LPD. In this way, vehicle testing could be based on grouping road events with similar LPD characteristics. For example, the LPD characteristics of real road events could be mapped to proving ground events to provide customer correlation to proving ground testing.

3.7 Conclusions

Identifying the contribution of each road profile excitation makes to the total accumulated pseudo damage is complex due to the vehicle dynamics coupled with the sequence dependency of excitations. This complexity was overcome in this work by modeling the Localized Pseudo Damage (LPD) as a linear combination of relative contributions an individual road exaction makes at subsequent locations. By modeling the relative contributions as a function

of the difference between the original damage rate response and the damage rate response resulting from the omission of the ith road excitation (with a time dependent scaling factor), the resulting LPD is causal and correctly identifies location and magnitude of damaging events. Through implementation of the LPD, damaging events can be identified from long segments of high-resolution road surface measurements prior to performing additional durability analysis. This additional analysis can be in the form of simulation of the identified high-resolution damaging events or further characterization to improve vehicle testing and/or customer correlation to proving ground testing. Both of these analyses are crucial for evaluating the durability of a vehicle prior to production.

Chapter 4

Identification of Damaging Road Events Using Pseudo Damage Density

4.1 Introduction

Knowing every damaging event experienced by a customer during the life of the vehicle would be ideal when modeling vehicle durability; however, the required vehicle simulations and analyses make this computationally impractical. One way to significantly improve efficiency is to identify events that would be damaging to a class of vehicles. Simpler, more general, vehicle models could be used to identify damaging events along load paths of interest for particular applications (e.g., chassis durability). These damaging events could then be used with more complex vehicle simulation models during the vehicle design and development process and used to validate the modeling with proving ground and laboratory tests based on these damaging events.

The objective of this work is to develop a method for identifying damaging events from measured road surfaces that allows for any vehicle model, load paths, speed, and road surface to be considered. Damaging Events are defined as regions of a road that cause significant pseudo damage in one or more load paths of interest and are separated by Inactive Regions (regions in which the dynamics in all load paths of interest have settled). The remainder of the developments in this work proceed from the identification of Inactive Regions. Specifically, an optimization problem is formulated through which the optimum set of damaging events is identified. This optimum set is constrained to account for some minimum required fraction of pseudo damage in all load paths and maximizes the efficiency of the representation.

This remainder of this chpater is organized as follows. First background on various road surface event detection methods is presented, followed by pseudo damage estimation techniques for vehicle applications. Next, a novel method for identifying and quantifying damaging events is developed through a multi-step process. First, Inactive Regions (IRs) are defined

and a process for identifying IRs is developed through analysis of the settling time of the dynamics and the magnitudes of the pseudo damage density. Second, an optimal search is developed based on the fraction of pseudo damage accounted for in each load path and the corresponding fraction of locations contained by the damaging events. An example is provided using a Federal Highway Administration (FHWA) Long-Term Pavement Performance (LTPP) road profile and a simple quarter car model in which two load paths of interest are investigated. Lastly, a discussion of future applications, concluding remarks, and references are provided.

4.2 Background

Mobile Road Mapping Systems (MRMS) can collect high-resolution road surface measurements that can be used to assess pavement condition and simulate vehicle dynamics. These systems typically contain localization and mapping sensors including scanning lasers, inertial navigation systems, and cameras [7, 8, 70]. Pavement distress analysis is performed on the measured road surfaces to identify segments that contain events (e.g. cracks, rutting, and potholes). The process by which events are identified in a road surface can be generalized into two categories: geometric and vehicle response. Event identification and vehicle damage estimation methods are reviewed in the remainder of this section.

4.2.1 Geometric Identification of Events

Geometric identification is commonly performed for analysis of road surface distress using image processing methods. Recent advances in geometric identification have allowed for a transition from manual analyses to semi-automated or fully-automated detection of events. Geometric identification generally proceeds by first identifying the edges of events and then the region encapsulated by the edges. Once events are identified, they can be classified based on their characteristics and compared to other road surface properties.

When two-dimensional images are acquired, a four step process is typically performed: 1) image processing, 2) segmentation of the image, 3) shape extraction, and 4) event recognition [71]. Numerous methods exist for processing and segmenting the image in an effort to enhance the edges of events [71, 72, 73]. When an event shape is extracted, the region of the road surface is typically modeled by an ellipse [74], even though an event's edges may be irregular. Several issues exist when identifying events in 2D images in this way; for example, poor lighting can affect the accuracy [75]. While 2D imagery typically uses gray scaling to emulate depth, 3D data measurement systems include a depth measurement resulting in a 3D point cloud. Zhang et. al proposed a method of identifying events (potholes) by performing a least squares fit of quadratic surfaces on the point cloud data and comparing the difference between the point cloud data and the fitted surfaces [76]. More recently, Mikhailiuk and

Dahnoun improved the modeling of the surface for event identification and used Random Sampling and Consensus (RANSAC) [77] to improve the detection of events [78]. While these methods successfully detect geometric events, no measure of vehicle damage is established. Vehicles traveling on these roads would still need to be simulated to determine if these geometric events are significantly damaging and to what extent that damage is accumulated.

4.2.2 Vehicle Response Event Identification

Vehicle Response Event Identification methods are typically defined using roughness (suspension travel) or vibration (acceleration) methods. These identification methods can be implemented on data collected from instrumented vehicles or on results of vehicle simulations based on measured road surface heights. The most widely used roughness measure is the International Roughness Index (IRI). The IRI is a measure of the average suspension travel of a golden Quarter Car traveling at a speed of 80 kph over a longitudinal road profile [10]. Since the IRI is an average measure, localized roughness levels are not identified. Zamora Alvarez *et. al* developed a Discrete Roughness Index (DRI) that provides a discrete measure of roughness for each point in a road profile and a method for identifying events that contain high levels of roughness [79]. Bogsjo proposed that roads can be considered a stationary signal with random high-roughness events with either a short or long wavelength [2]. Under this assumption, it was shown that a roughness criterion can be used to identify locally rough segments of the road profile. Bogsjo and Rychlik developed a model for estimating the expected damage contribution for road irregularities identified using the roughness criterion [19]. However, an investigation performed by Zamora Alvarez found that a one-to-one relationship between road roughness and pseudo damage does not exist resulting in inconsistencies when relating road roughness to pseudo damage [4]. An assumption in the model developed by Bogsjo and Rychlik is that the road profile must be a stationary signal, however many studies have shown that while many smooth roads are stationary, that is not typically the case for rough roads from which a majority of damage is accumulated [45, 80, 81].

Vibration-based methods for event identification use irregularities in vehicle accelerations that result from driving over a road surface; most recently, accelerometers in smartphones have been used to detect abnormal vibrations [82]. In an effort to automate the identification process, numerous instances of machine learning algorithms are deployed on coupled accelerometer and Global Position System (GPS) data [83, 84, 85]. The first drawback of using accelerometers to detect damaging events is that accelerometers measure motion and unless a detailed vehicle model is used, the forces acting on the body are not known. This issue is compounded because the acceleration signals are measured from the vehicle body (the sprung mass). However, the vehicle suspension and tires (the unsprung mass) are designed to minimize the objectionable movement of the vehicle body; they are large mechanical filters that remove significant motion from the body and must therefore go undetected by the accelerometers. In addition to these significant issues, false detections can also occur due to

movements inside the vehicle (e.g., passengers and luggage) that do not emanate from the road and due to small irregularities in rotating components (e.g., irregular tire balancing).

4.2.3 Pseudo Damage and Localized Pseudo Damage

Components are damaged when subjected to varying external forces. These forces produce a varying strain field that depends on the material properties and geometry of the component being subjected to the external loading. When the geometry and material properties are well known, the damage can be predicted based on the variation in the strain field. However, stresses and strains cannot be calculated when the material or geometry of the component are unknown. In these instances, pseudo damage can be used to estimate the relative damage produced by the external forces through a particular load path. Pseudo damage is calculated from forces without regard to component geometry, so that only relative comparisons within a particular load path can be made and pseudo damage comparisons cannot be made between different load paths.

The total pseudo damage, D_T, is typically calculated using the Palmgrin-Miner damage accumulation method and the Basquin relation to provide an estimate for the number of cycles to failure [20]. The results of coupling these damage models is provided in Eq. (4.1) where a set of load amplitudes, L_ℓ, are known along with the fatigue damage coefficient, β, and the pseudo damage scaling factor, C. Each separate load cycle is indexed by ℓ, but note that the cyles may be overlapping. Since the developments in this work are focused on the analysis of relative pseudo damage values within a load path, the value of C is inconsequential to the development of the work.

$$D_T = C \sum_{\ell=1}^{N} |L_\ell|^\beta \tag{4.1}$$

Since only a single load amplitude is defined per load cycle, the rate at which pseudo damage is accumulated within a load cycle is not considered in Eq. (4.1). However, the accumulation of pseudo damage occurring throughout the overlapping load cycles is required in this work to properly assign the pseudo damage contributed by each location.

Consider the time rate of change of the pseudo damage at some location indexed by j, $\dot{D}_j$, where the approximate integral for total pseudo damage is defined in Eq. (4.2) with respect to discrete time steps, Δt_j.

$$D_T = \sum_{j=0}^{N-1} \dot{D}_j \Delta t_j \tag{4.2}$$

Clearly, individual load amplitudes are insufficient to calculate the pseudo damage rate. Rychlik addressed this issue by developing the Top-level-Up Cycle (TUC) counting method that assigns damage to specific points within a cycle rather than over a complete cycle [63, 64]. Bogsjo and Rychlik showed that the damage rate can be calculated using the

Palmgrin-Miner and Basquin damage model at each time indexed by j as the difference between the load at j, F_j, and a corresponding peak load, $F_{p,j}$, and the time rate of change of the load at j, $\dot{F}_j$, as shown in Eq. (4.3).

$$\dot{D}_j = G\left|F_j - F_{p,j}\right|^{(\beta-1)}\left|\dot{F}_j\right| \tag{4.3}$$

Pseudo damage rate and the TUC counting method allow for representation of pseudo damage in the time domain and provide an accounting method for the pseudo damage accumulated throughout a load history. However, these methods do not explicitly provide a relationship between pseudo damage and road profile excitation. One means to address the effect of road excitations on the pseudo damage through a load path is the pseudo damage density. Pseudo damage density is the rate at which pseudo damage is accumulated with respect to the longitudinal distance traveled. Note that by capturing this effect using density, the sensitivity of the measure to differences in longitudinal spacing of measurement points is minimized. Consider a location indexed by i corresponding to a longitudinal distance traveled, u_i, with pseudo damage density, d_i. The approximate integral for total pseudo damage is defined in Eq. (4.4) with respect to discrete longitudinal steps, Δu_i.

$$D_T = \sum_{i=0}^{N-1} d_i \Delta u_i \tag{4.4}$$

Methods to account for pseudo damage density must possess several properties. First, since a vehicle is a causal dynamic system, the pseudo damage density must be causal. That is, the calculation of the pseudo damage density must not be affected by any future excitations. Second, pseudo damage density for singular events, such as an impulse or step in the absence of other excitations, must only be attributed to that singular event and no portion of the damage attributed to any other location. One method for calculating pseudo damage density that possesses these properties was developed in Chapter 3, and referred to as Localized Pseudo Damage (LPD). The LPD at each location is calculated by comparing two sets of damage rates: one from the original road profile and the second from the road profile with the individual excitation at that location removed. The LPD is a function of the difference in the responses to these two road profiles.

4.3 Identification of Damaging Events

The main contribution of this work is the use of pseudo damage density to identify Damaging Events (DE). Damaging Events are defined as regions of a road that cause significant pseudo damage in one or more load paths of interest and are separated by Inactive Regions. In identifying damaging events, the role of individual load paths and the effect of the response time for each load path must be considered. Each of these issues is discussed in turn before developing the concept of Inactive Regions (IR): a region in which the dynamics in all

load paths of a vehicle system have settled. Finally, an optimization problem is formulated through which the optimum set of damaging events is identified.

Pseudo damage must be accounted for in each load path. However, pseudo damage is calculated from forces without regard to component geometry and therefore pseudo damage values are only meaningful for comparison within a particular load path. It should be clear that the load paths of interest depend on the particular application and are assumed to be known. The response time of each load path depends on the dynamics of the load paths themselves and how they are coupled through the vehicle dynamics. Significant coupling takes place within the vehicle dynamics and so although the pseudo damage density can be attributed to specific locations for each load path, the attribution is dependent on both previous and subsequent responses. Previous responses define the dynamic state at each location and subsequent responses are used to calculate the damage density. The result is the ability to attribute the pseudo damage density to each location within a damaging event, but this attribution only has meaning when the damaging event is taken as a whole. That is, damaging events must be separated by regions of relative inactivity in order to be meaningful.

The first step in identifying damaging events is partitioning the road into the sets of Active Regions (AR) and Inactive Regions (IR) as illustrated in Figure I1. A solid line is used to indicate the partition of the space into two sets: IR and AR. Damaging events (DE) are a subset of active regions, which is illustrated by a dashed line.

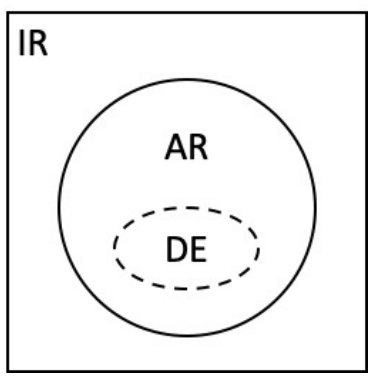


Figure 4.1: Set diagram for identification of damaging events.

Active Regions include one or more load paths in which the dynamics are unsettled, indicating potential accumulation of significant pseudo damage. Damaging Events are Active Regions that cause significant pseudo damage. The coupled dynamics requires that each damaging event must be separated by an Inactive Region to ensure the dynamics across all load paths have settled. In this way, Damaging Events are unique and separable and must be complete Active Regions. Note that there may be non-damaging segments within a Damaging Event for some load paths and there may be some load paths that do not accumulate

any significant damage within the event. The next steps are to develop a process to identify Inactive Regions (IRs) and formulate the optimization problem in which the search space for the set of Damaging Events are the Active Regions.

4.3.1 Identifying Inactive Regions

It is assumed that there are inactive regions in the dataset in which the vehicle dynamics have settled so that damaging events are separable (and the problem is non-trivial). The issue is defining the settling of the dynamics in terms of magnitude and duration. In this work the response of interest is the pseudo damage and, for a given load path p, the pseudo damage at location i is $d_{p,i}\Delta u_i$, where $d_{p,i}$ is the pseudo damage density and Δu_i is the longitudinal spacing. The magnitude of the pseudo damage at each discrete location is therefore $|d_{p,i}|\Delta u_i$, and is defined as the activity, $A_{p,i}$, for the ith location in the pth load path, shown in Eq. (4.5).

$$A_{p,i} = |d_{p,i}|\Delta u_i \ , \ \forall \, p \, , \, i \tag{4.5}$$

At each location the damage density and activity are a function of all future damage rate contributions from excitations up to the location of interest. To explain this idea consider two locations, k and m, separated by a contiguous set of inactive locations and k is located prior to m. At both locations an active damage density value is known, $d_{p,k}$ and $d_{p,m}$, respectively. The damage density at k is a function of the future damage rate responses from all excitations up to k and the damage density at m is a function of the future damage rate responses from all excitations up to m. Since location m shares contributions with location k we need to ensure the separation between m and k is sufficiently long that the damage rate responses associated with k have settled prior to m. To ensure this is the case the set of contiguous inactive locations between k and m must exist and this set must have a total duration greater than equal to the settling distance of the damage rate impulse response. By satisfying this duration condition, damage rate responses associated with k are considered separate from the responses associated with m because the damage rate response is sufficiently settled between the locations.

Next the duration over which the activity has settled must be determined. Activity is a function of pseudo damage density, which is a function of damage rate, which is a response of the vehicle system (e.g., load path dynamics, vehicle speed, direction of travel, etc.). For any load path of interest, a damage rate impulse response can be determined and a corresponding settling time, $T_{s,p}$, can be determined from the damage rate impulse response function.

For a given location, k, along load path, p, consider a window of length, $L_{k,p}$, where the length is the product of the vehicle velocity and the settling time for the pseudo damage associated with that load path $T_{s,p}$. That is, the length of the window is the minimum longitudinal distance necessary for the response (pseudo damage) from a single road excitation to sufficiently decay for that load path and vehicle speed. The locations that lie within the

k^{th} window, for the pth load path, define the set $W_{k,p}$ defined in Eq. (4.6).

$$W_{k,p} = \{i : u_i \in [u_k \, , \, u_k + L_{k,p}]\} \tag{4.6}$$

Note that individual locations that have a low activity value do not necessarily indicate that the dynamics have settled; this is ensured by a window of locations in which low-activity persists throughout. For each window, $W_{k,p}$, a measure, *psi*, on the set of activity values for the window is defined in Eq. (4.7).

$$\psi \left(A_{p,i} \, , \, \forall i \in W_{k,p} \right) \tag{4.7}$$

The measure of activity is application specific and outside the scope of this work, however common measures include the maximum, average, or RMS value. The choice of an appropriate threshold on that measure, γ_p, is also application specific, but a threshold on that measure of activity, whatever that measure may be, can be specified. The window $W_{k,p}$ is inactive if the measure of activity on this window is less than the threshold for the load path, γ_p, and an inactive window is defined as $V_{k,p}$ in Eq. (4.8).

$$V_{k,p} = \{W_{k,p} : \psi \left(A_{p,i} \, , \, \forall \, i \in W_{k,p} \right) < \gamma_p\} \tag{4.8}$$

By construction, all inactive windows are of sufficient length, $L_{k,p}$, to ensure that the dynamics of the load path are settled. However, it should be clear that there are often regions that are inactive whose length is greater than $L_{k,p}$. That is, there are often overlapping windows of inactivity that form larger regions of inactivity and, therefore, inactive regions are the union of inactive windows. More precisely, the set of locations within inactive regions is defined as the union of the inactive windows for each load path, as defined in Eq. (4.9).

$$U_p = \left\{ i : i \, \in \bigcup_{\forall k} V_{k,p} \right\} \tag{4.9}$$

The response of each load path depends on the significant coupling of the vehicle dynamics; it follows that a region can only be considered inactive if all load paths are inactive. The set of inactive locations, I, is the intersection of inactive load path locations, U_p, across all load paths as defined in Eq. (4.10).

$$I = \left\{ i : i \, \in \bigcap_{\forall p} U_p \right\} \tag{4.10}$$

Inactive regions (IR) are the sets of contiguous locations in set I and Active Regions (AR) are the sets of contiguous locations not in the set I. In this way, active regions are separated by inactive regions. Since damaging events are subsets of active regions, all damaging events are separated by inactive regions, as required. An Active Region is classified as either a Damaging Event or not. The rth Damaging Event is written as E_r as defined in Eq. (4.11).

$$E_r = \{i | i \in AR_r : AR_r \text{is a damaging event}\} \tag{4.11}$$

The set of location is damaging events, E, is defined in Eq. (4.12) as the union of all damaging events.

$$E = \cup E_r \quad \forall\, r \tag{4.12}$$

Damaging events are defined by Eq. (4.11) and Eq. (4.12) without addressing the means by which a particular AR is determined to be a DE or not. This determination is made as a result of solving the optimization problem that is presented in the following section.

4.3.2 Defining the Optimization Problem

In this work, the optimal set of damaging events is a subset of the active regions that account for at least a minimum fraction of pseudo damage for all load paths and is the optimal balance between pseudo damage accounted for and the length of the road required to achieve that pseudo damage. First consider a set of damaging locations, $i \in E$, that contributes different levels of pseudo damage to each load path. For load path p, the fraction of pseudo damage accounted for by the damaging locations, $FD_{E,p}$, is defined in Eq. (4.13).

$$FD_{E,p} = \frac{\sum_{i \in E} d_{p,i} \Delta u_i}{\sum_{\forall i} d_{p,i} \Delta u_i} \tag{4.13}$$

An imposed constraint on the optimization problem is that for each load path a minimum fraction of pseudo damage must be accounted for. The choice of an appropriate threshold on that measure, η_p, is application specific, but can be specified. The resulting set of constraints are provided in Eq. (4.14) in terms of the fraction of pseudo damage per load path.

$$FD_{E,p} > \eta_p \ , \ \forall\, p \tag{4.14}$$

Next, a figure of merit is defined for each load path and the objective function is defined as a measure on these figures of merit. First consider the fraction of locations accounted for in the set of damaging locations, $i \in E$. This fraction of locations, FL_E, is the ratio of the cardinality of E to the total number of locations in the complete road profile, N, as defined in Eq. (4.15).

$$FL_E = \frac{|E|}{N} \tag{4.15}$$

Each figure of merit, $FM_{E,p}$, is defined as the ration of two fractions: the fraction of pseudo damage retained in the load path, $FD_{E,p}$, as presented in Eq. (4.13), and the fraction of locations retained in the set $i \in E$, FL_E, as defined in Eq. (4.15). The figure of merit for each load path, $FM_{E,p}$ is given in Eq. (4.16).

$$FM_{E,p} = \frac{FD_{E,p}}{FL_E} \tag{4.16}$$

It should be clear that a unity figure of merit can always be achieved for all load paths by including all the locations as damaging events ($FL_E = 1 \rightarrow FD_{E,p} = 1 \rightarrow FM_{E,p} = 1$).

Lastly, the objective function is defined as a measure on the set of figures of merit. This measure, ϕ, is application specific and assumed to be known. Two potential measures for the objective function are the RMS or maximin across all load paths. The definition of this measure is provided in Eq. (4.17).

$$\phi\left(FM_{E,p}\ ,\ \forall\, p\right) \tag{4.17}$$

By maximizing the objective function subject to the constraints imposed on the minimum fraction of pseudo damage for each load path, the optimal set of damaging events is identified. The process of implementing this optimization problem is illustrated through an example road profile in the following section.

4.4 Example

To illustrate the process of identifying damaging events, a 152.4m (500 ft) measured segment of public highway used in the Federal Highway Administration (FHWA) Long Term Pavement Performance (LTPP) project is used [86]. The specific road profile used in this example is the right wheel path for the northbound, outside lane of I-15 near Temecula, California and is presented in the first subplot of Figure E1. The measurements of the highway were collected using equipment approved by the FHWA and road profile elevations are reported in 25 mm increments. As a proof of concept for the event identification method developed herein, a simple vehicle model with two load paths is used. The selected model is a Golden Quarter Car model [9] and the two load paths of interest are the loads through the spring and the damper. The resulting Pseudo Damage (PD) density for the two load paths is provided in Fig. E1. The second subplot of Fig. E1 contains the resulting PD density for the spring load path and the third subplot contains the PD density for the damper load path. Recall, that PD for different load paths cannot be directly compared, only relative comparisons along a load path can be made. In this example, the damage rate was scaled such that the range of PD density values is between -1 and 1 for both load paths. In both load paths, instances of positive, negative, and near zero PD densities are present. Positive PD densities indicate road excitations that increase the total amount of PD in the load path. Negative PD densities indicate road excitations that, within the context in which they occur, reduce the amount of PD that could have been experienced through a load path had that excitation been removed. Near-zero PD densities indicate locations that have minimal effect on the accumulated pseudo damage.

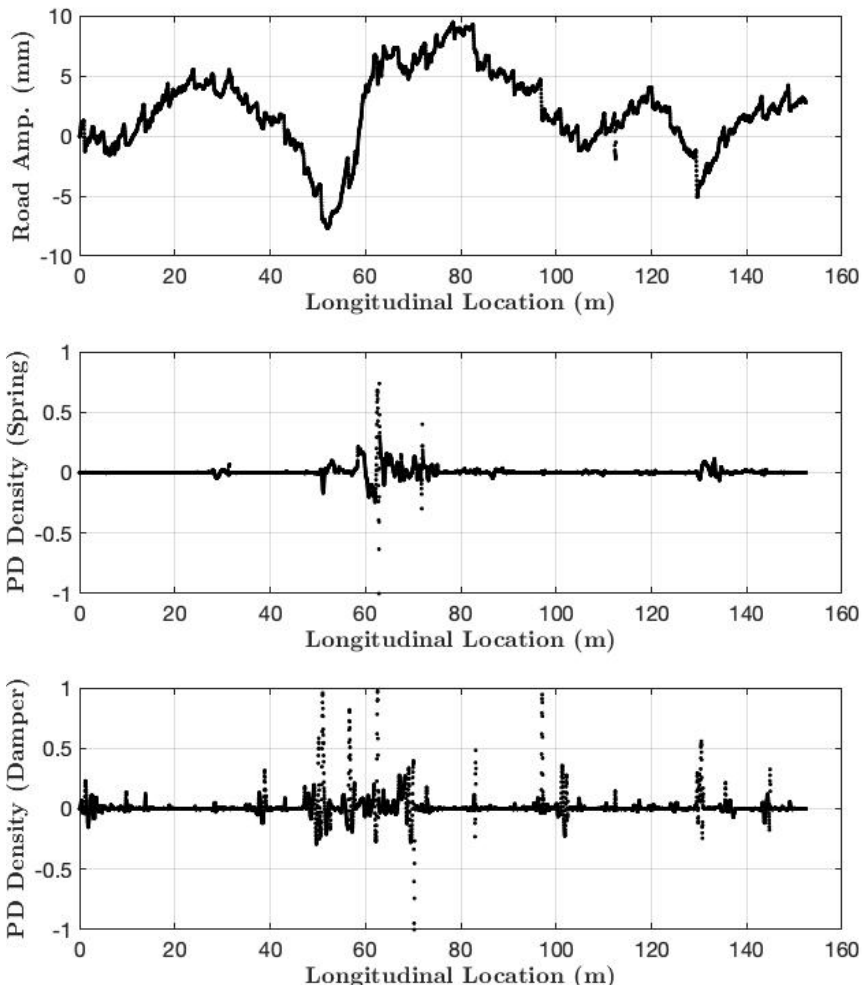


Figure 4.2: Road profile of interest and the resulting PD density for the spring and damper load paths.

To begin the process of identifying damaging events the settling distance associated with each load path must be determined to define the minimum length for inactive regions. To define the settling distance first the settling time for all load paths must be defined. For any load path of interest, the damage rate impulse response, $H_{\dot{D},p}$, can be calculated and the corresponding settling time, $T_{s,p}$, can be calculated from Eq. (4.18). The settling time is based on the time an impulse response is bounded from the remainder of time by a maximum threshold value, ϵ_p.

$$T_{s,p} = \max_{t \in \mathbb{R}} \left(H_{\dot{D},p} > \epsilon_p \right) \tag{4.18}$$

The maximum threshold, ϵ_p, is defined based on a fraction, ν_p, of the maximum damage rate in the impulse response as provided in Eq. (4.19). The fraction, ν_p, is application specific,

therefore selection of this value is outside the scope of this work.

$$\epsilon_p = \nu_p \max\left(H_{\dot{D},p}(t)\right) \tag{4.19}$$

The damage rate impulse response for the spring and damper load paths for a Golden quarter car model are provided in Figure 4.3. The dashed line in Figure 4.3 indicates when the amplitude of the damage rate impulse response fell below a selected threshold that corresponds to $nu_p = 0.02$ as defined by Eq. (4.19). This threshold resulted in a settling time of $T_{S,1} = 0.08$ s for the spring load path and $T_{S,2} = 0.05$ s for the damper load path.

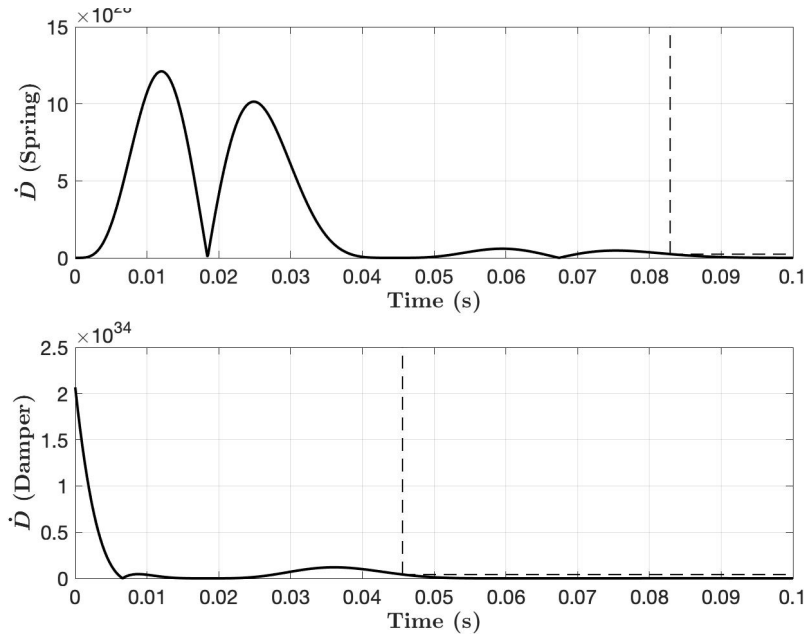


Figure 4.3: Road profile of interest and the resulting PD density for the spring and damper load paths.

A constant vehicle speed of 80 kph results in a settling distance of $L_{k,1} = L_1 = 1.8$ m for the spring load path and $L_{k,2} = L_2 = 1.0$ m for the damper load path, which are used to define the minimum duration of inactive regions for each load path. To define the measure of activity in inactive regions, ψ was defined to be the maximum function. The bounds on the maximum value for activity, γ_p, was selected to be 1% for the spring load path and 5% for the damper load path. The threshold on PD density for each load path was selected based on the effect the threshold has on the fraction of PD in active locations. For the spring load path, a threshold of 1% ensures 90% of the spring PD is accounted for in active locations. Similarly, for the damper load path a 5% threshold ensures that 90% of the damper PD is accounted for in active locations. Thus, any window, $W_{k,p}$, that satisfies the duration criteria along with the activity measure ψ is considered an inactive region, $V_{k,p}$. The union of all inactive

regions for a load path defines the complete set of inactive regions, U_p. The intersection of the inactive regions for the spring load path with the damper load path defines the inactive locations, I, in the road profile. Each set of contiguous inactive locations is illustrated by gray shaded areas in Figure 4.4. In each of these regions, the dynamics of the vehicle system are considered settled allowing for separation of active regions.

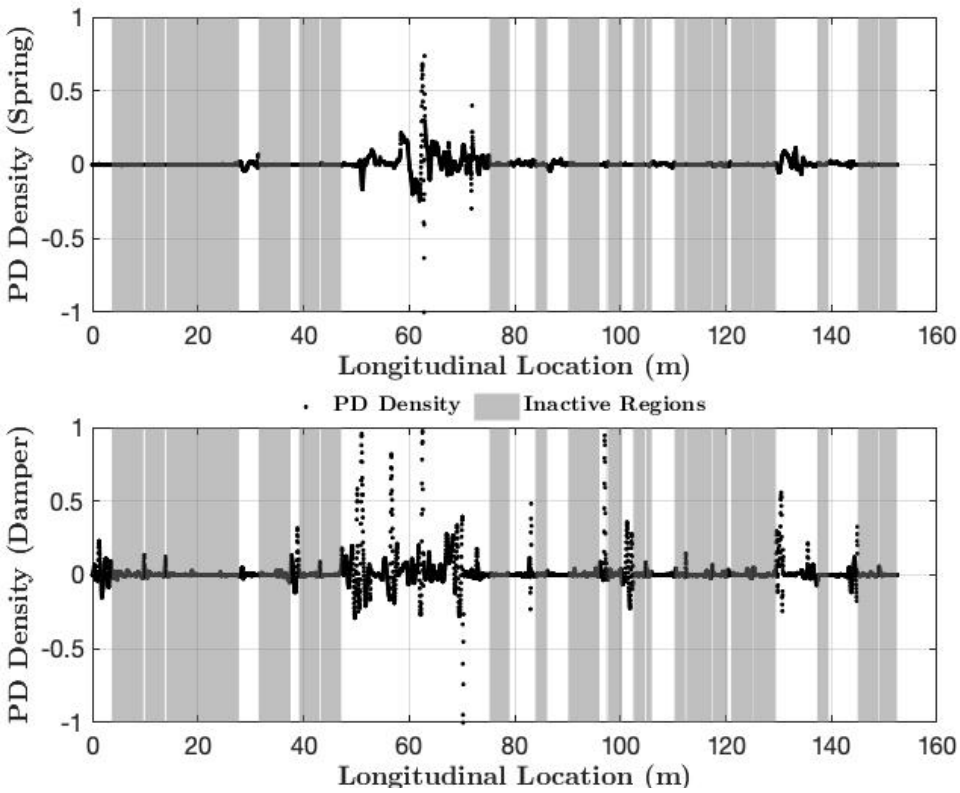


Figure 4.4: Inactive regions identified using a maximum threshold on the PD density.

For this road profile, 24 active regions were identified and are represented by the unshaded regions, separated by inactive regions, in Figure 4.4. These Active Regions define the search space for the optimization. The objective function is a measure on the set of figures of merit, ϕ; the measure chosen for this work is the RMS of the figures of merit for the spring and damper load paths. In addition, the imposed constraint, ν_p, on the minimum fraction of PD in both the spring and damper load paths was selected to be 70%. Solving the optimization problem resulted in identifying five separate damaging events, which are highlighted by the unshaded regions in Figure 4.5. The gray shaded regions are considered negligibly damaging that include all Inactive Regions (see Figure 4.4) and a subset of the Active Regions that are not significantly damaging.

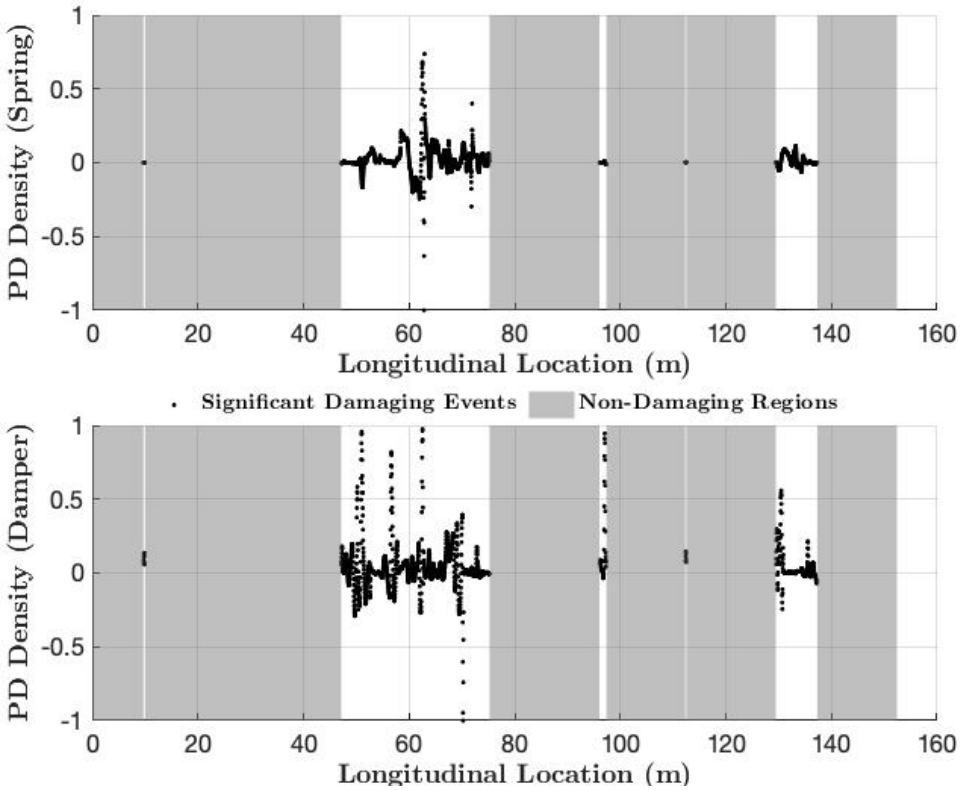


Figure 4.5: Unique significantly damaging events separated by non-damaging regions.

The total damage characteristics and fraction of road length for each damaging event is summarized in Table 4.1, where the order of the events corresponds to relative location of the event from the start of the road profile. The fraction of the PD accounted for in the spring load path, $FD_{D,1} = 82.4\%$ and in the damper load path $FD_{D,2} = 73.7\%$. Note that both values exceed the $\nu_p = 70\%$ threshold. In addition, the fraction of locations, $FL_D = 24.6\%$. This optimal solution resulted in an objective function value of 3.168. Event 1 and Event 4 are of particular interest. Both of these events could be omitted and still satisfy the $\nu_p = 70\%$ threshold. However, if Event 1 and 4 are omitted the objective function value drops by merely 0.16%, to 3.163. This highlights the importance of choosing thresholds and measures that truly capture the desired objectives.

Table 4.1: Damaging Event Summary

Damaging Event Number	Fraction of Locations, FL_D (%)	Fraction of Spring PD, $FD_{D,1}$ (%)	Fraction of Damper PD, $FD_{D,2}$ (%)
1	0.16	0	1.2
2	18.4	70.3	48.8
3	0.84	0.01	10.0
4	0.11	0.05	0.9
5	5.1	12.1	12.8

4.5 Discussion

The contribution of this work is a method for identifying damaging events from a road profile to a vehicle system. The first step in identifying damaging events is partitioning the road profile into active and inactive regions. Where inactive regions are the portions of the road profile where the dynamics in all load paths have sufficiently settled. Each damaging event must be separated by an inactive region to ensure consecutive damaging events are unique and separable. Thus it follows that individual active regions can be considered damaging events based on the results from an optimization. Once all active regions are identified an optimization problem is used to identify the best set of active regions to maintain as damaging events in order to achieve a minimum fraction of PD in each load path of interest while minimizing the length of the road profile maintained in damaging events.

In the example, a simple Golden quarter car model with two load paths was used to illustrate the methodology developed herein. However, it should be clear that this work is developed in general terms and applicable to any vehicle model and number of load paths. In this work, the maximum value of the activity function, ψ, is chosen; however, this choice is application specific and is not a limitation of the contribution of this work. In future work, the selection of the activity threshold function, ψ, can be investigated based on the application to understand the affect the activity threshold has on identifying damaging events. Similarly, the objective function chosen is the RMS of the spring and damper load path values. The RMS function was selected to identify an optimal set of damaging events based on the fraction of damage accounted for across all load paths. In future work, the implication of the selection of an objective function, ϕ, can be investigated based on the application. For example, if the objective function was defined as a maximum value then one of the load paths may have driven the selection of an optimal value.

This work is based on pseudo damage (PD) density as a measure for identifying damaging events and the concept of Activity to ensure each damaging event is unique and separable. It is envisioned that the sets of PD densities that comprise each damaging event can be used to characterize and group/cluster events based on their damaging characteristics. When characterizing damaging events from a road surface into a vehicle two important characteristics

are: amount of damage caused and length of event; otherwise stated the damage density of the event. By characterizing and grouping/clustering damaging events, similarities between events can be identified resulting in subgroups of multiple damaging events which have similar effects on the accumulated damage in each of the load paths of interest. In addition, proving ground tests or synthetic events which have similar damaging characteristics to a group of damaging events can be developed.

4.6 Conclusion

The process of identifying unique, separable damaging events is complex because vehicles are causal systems in which the damage accumulated is dependent at varying levels to all prior road excitations. This issue was addressed by identifying damaging events using pseudo damage density, which allows for damage in a load path to be expressed as a function of road location. A second issue that was addressed is that in a vehicle system damage is accumulated through multiple load paths simultaneously and each load path has a set of unique dynamics. In this work, this issue was addressed through partitioning of a road profiled into active and inactive regions, then use of an optimization problem to identify the subset of active regions that are considered damaging events. All theoretical developments in this work are generalized to any desired number of load paths of interest resulting in an event identification method that is scalable based on the application. The formulated optimization problem ensures that a minimum fraction of PD is maintained in each load path while minimizing the length of road maintained in damaging events. When damaging events are identified using this optimization problem the amount of necessary road surfaces to simulate for additional durability analysis is reduced and further analysis can be performed using only events which are damaging to the vehicle (or component).

Chapter 5

Identifying the Similarity Between Damaging Events Using Pseudo Damage Density

5.1 Introduction

When performing vehicle durability analysis and testing it is critical that the damaging events used to model, simulate, and validate vehicle durability performance best match the damaging events that a significant population of the customers experience. However, due to the large population size of customers and large range of public roads a customer could traverse, over 4 million miles in the United States alone [1], it is nearly impossible to capture every event. However, it is possible to know the damage characteristics associated with the current state of a subset of the total number of damaging events. For this subset of damaging events, Mobile Road Mapping Systems (MRMS) can be used to collect measurements of the current state of road surfaces and these measures can be used to identify damaging events and determine the damage characteristics associated with those damaging events. It would be ideal if from the subpopulation of known damaging events that groups of damaging events are formed, where the groups are based on shared similarities in the damage characteristics. By forming these groups of damaging events, further customer usage analysis can be performed based on the groups of the damaging events.

The method of grouping similar damaging events will be different based on the application needs. For example, to best replicate the damaging events a customer experiences during proving ground testing it is desired to identify which proving ground events best replicate the damage characteristics of the damaging events the customer experiences. Then these proving ground events can be used to develop a durability test cycle. In a second example it may be desirable to identify the best grouping between all damaging events, independent of the

origin (*e.g.*, public versus proving ground) of the damaging event. This second application is different from the prior application in that instead of having one proving ground event per group, it is possible that multiple or no proving ground events may exist in each group. However, each group will contain the most similar damaging events.

In this work a statistical analysis for comparing a damaging event to an existing group of damaging events using pseudo damage density in multiple load paths for a vehicle system is developed. In the development of the statistical analysis, two options for similarity metrics are provided. The first option is the Mahalanobis distance and the second option is a likelihood based on a reference distribution from the Mahalanobis distance. The statistical analysis provides the information necessary for determining if a new damaging event is drawn from the same population as an existing group of damaging events. The ability to reject this this hypothesis is application specific. To demonstrate the statistical analysis and the effect of application on on the clustering algorithm and rejection of the hypothesis, two examples are provided that focus on answering the examples posed above.

The remainder of this chapter is organized as follows. First background information in: definition and approach for identifying damaging events, overview of the current state-of-the-art in clustering algorithms, and similarity and statistical evaluation techniques are presented. Next, the statistical analysis is developed in general terms based on a set of indexed damaging events and clusters. Following, two examples using the statistical test on damaging events from real-world road profiles is presented. In the first example, the idea of identifying a damaging event from a set of test events (*e.g.* proving ground events) that is most similar to a damaging event from a public road is identified. In the second example clusters of damaging events are formed using a hierarchical clustering algorithm coupled with the developed statistical assessment. Two potential sets of clusters are formed based a threshold similarity requirement. Lastly, a discussion on future applications of this work are presented along with concluding remarks.

5.2 Background

Mobile Road Mapping Systems (MRMS) can be used to collect measurements of the current state of road surfaces from which longitudinal road profiles can be established. A longitudinal road profile consists of a set of indexed road elevations at discrete locations along the length of a road. Each longitudinal road profile can be used as an input into a vehicle model/simulation resulting in a force-time history through multiple load paths of interest for a vehicle system. Next, the force-time history in each load path can be used to perform pseudo damage (PD) analysis in order to identify separable damaging events. The remainder of the background section is broken up into three section. The first section covers the process of analyzing force-time data to determine pseudo damage and identify damaging events. The second section provides an overview of the current state-of-the-art in clustering analysis. Lastly, an overview of similarity statistics and statistical evaluation tests is presented.

5.2.1 Pseudo Damage and Damaging Events

To identify damaging events the pseudo damage (PD) density at each location in a road profile is needed in order to identify and quantify damaging events. The first step in this process is to estimate the PD rate for each load path. In general, the time rate of change of pseudo damage (pseudo damage rate) for some location indexed by j is defined as $\dot{D}_j$. Using this definition of pseudo damage rate, an approximate integral for the total damage is defined in Eq. (5.1) with respect to a discrete time step, Δt_j.

$$D_T = \sum_{j=0}^{N-1} \dot{D}_j \Delta t_j \tag{5.1}$$

Methods for calculating the PD rate have been established by Bogsjo and Rychlik [63, 64]. Pseudo damage rate provides a relationship between time and total pseudo damage, however it does not explicitly provide a relationship between road location and total pseudo damage. One method for identifying the relationship between pseudo damage and road location with respect to the longitudinal distance traveled is the pseudo damage density. When this relationship is captured using density, the sensitivity of the measure to changes in longitudinal spacing of measurement points is mitigated. For a location indexed by i a corresponding distance traveled, u_i, and a pseudo damage density, d_i, are known. Using the pseudo damage density at each ith location, an approximate integral can be formed for the total pseudo damage as defined in Eq. (5.2) with respect to discrete longitudinal steps, Δu_i.

$$D_T = \sum_{i=0}^{N-1} d_i \Delta u_i \tag{5.2}$$

Any method used to determine the pseudo damage density at each ith road location must posses two properties. First, since vehicles are a causal dynamic system, the pseudo damage density must be causal. Since vehicles are a causal dynamic system, the pseudo damage density must also match that characteristic and be causal. For PD density, that means the PD density at a location must not be affected by any future locations in the road profile. The second property is that for a singular event (*e.g.*, impulse or step) in the absence of other excitations, the PD density must be attributed to only the singular event and no portion of the damage should be attributed to any other location. One method for calculating pseudo damage density that possesses these properties and is used in this chapter is the Localized Pseudo Damage (LPD), which was developed in Chapter 3. The LPD at each location is calculated by comparing two sets of damage rates: one from the original road profile and the second from the road profile with the individual excitation at that location removed. The LPD is a function of the difference in the responses to these two road profiles.

Damaging events are contiguous sequences of locations that cause significant pseudo damage in one or more load paths and are separated by inactive regions, regions where insignificant amount of pseudo damage is accumulated in all load paths for at least a minimum duration.

By knowing the pseudo damage density along a road profile for multiple load paths, inactive regions can be identified amongst all load paths and the remaining locations are considered active regions. One event identification method, used in this work, uses pseudo damage density in multiple load paths to partition the road into active regions and inactive regions (in which the vehicle dynamics have settled). Damaging events are a subset of the active regions and are identified through the solution to an optimization problem. The optimization problem requires that the set of active regions selected to be damaging events will achieve at least a minimum fraction of the total pseudo damage in all load paths. Within this constraint, the optimal solution maximizes the efficiency of the choice of damaging events, maximizing the benefit (fraction of the pseudo damage retained by the damaging events) relative to the cost (the fraction of locations retained as damaging events). The development of the optimization problem and additional details can be found in Chapter 4.

When a damaging event, indexed by r, is identified, the set of pseudo damage density samples for each each load path can be stored in a single matrix, $\mathbf{d}_r$, as shown in Eq. (5.3). Each row in $\mathbf{d}_r$ corresponds to a separate load path of interest and each column corresponds to a unique sample (location) from the damaging event.

$$\mathbf{d}_r = \begin{bmatrix} d_{1,1} & \cdots & d_{1,i} & \cdots & d_{1,N_r} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{p,1} & \cdots & d_{p,i} & \cdots & d_{p,N_r} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{n_p,1} & \cdots & d_{n_p,i} & \cdots & d_{n_p,N_r} \end{bmatrix} \tag{5.3}$$

In Eq. (5.3), the index p denotes the load path and index i denotes the location within the event. It is important to note that each event, E_r, will have a unique number of samples, N_r, that can vary between events. Thus, any evaluation test used to compare the similarity between events must also address the variation in the number of samples between events.

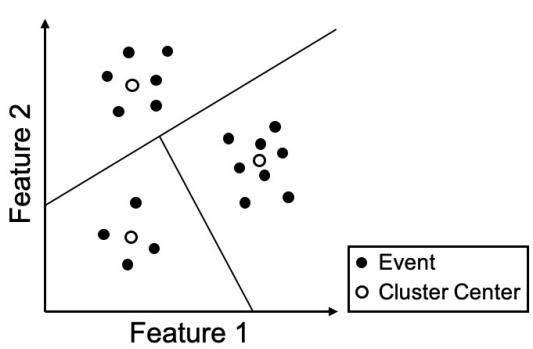
5.2.2 Overview of Clustering Analysis

In applications where multiple events or objects are known and for each event a set of features characterizing the event are known, a clustering algorithm can be used to group/cluster events. Clustering algorithms produce clusters of events such that events with similar features are placed in one cluster and events with dissimilar features are placed in separate clusters. The objective of clustering analysis is to highlight useful information or trends in a set of objects with similar features which may have otherwise gone unnoticed [50]. Over the last decade clustering analysis has experienced great success across a variety of applications.

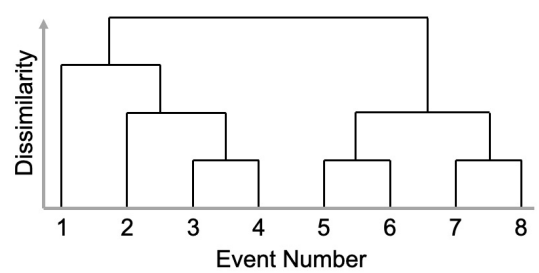
Clustering analysis has four main goals: (1) Development of a typology or classification scheme, (2) investigation of useful conceptual schemes for grouping entities, (3) hypothesis generation through data exploration, and (4) hypothesis testing [51]. Clustering analysis

aims to categorize a set of events with some known features into clusters where each cluster contains high similarities while low similarities exist between clusters. Two main forms of clustering algorithms are partitional and hierarchical clustering. Partitional clustering is the division of a set of events into non-overlapping clusters such that each event is in exactly one cluster and each cluster contains at least one event [52, 53]. A generalized representation of partitional clustering for a set of events with three clusters is provided in Figure 5.1a. In Figure 5.1a, each event is represented by a black dot and characterized by two features. The first feature is referenced along the x-axis and the second feature is referenced along the y-axis. For each of the three clusters, a cluster center is established and referenced by the open circles. The definition of a cluster center is dependent on the clustering algorithm selected. Typically, the cluster center is defined as: the average of all events in the cluster (k-means clustering), median of all events in the cluster (k-medians clustering), or a selected event that is centrally located in the cluster (k-medoids clustering).

Hierarchical clustering is the development of a hierarchy between multiple events [54]. The results of hierarchical clustering are summarized in a dendrogram, such as that shown in Figure 5.1b. A dendrogram provides a tree-like structure that connects similar events together through branches and as you work your way up the branches connections between similar events are formed. The level at which connections are made between events reflects the level of similarity between the events. Thus, events which are connected near the bottom of the dendrogram are more similar than those which are connected near the top. In the dendrogram in Figure 5.1b eight events are clustered and represented by the numbers along the x-axis. The connections formed between similar events is highlighted by the horizontal lines and the corresponding dissimilarity associated with the connection is shown by the y-axis value of the horizontal line. As an example, the dissimilarity between events 3 and 4 is relatively low, while the dissimilarity between event 2 and the cluster of events 3 and 4 is slightly higher.



(a) Generalized Representation of Partitional Clustering.



(b) Generalized Representation of Hierarchical Clustering.

Figure 5.1: Generalized representation of two general types of clustering analysis.

Both partitional and hierarchical clustering operate on the ability to define a measure of similarity between objects. A similarity between objects is commonly established through either a distance between two objects or the probability of two objects being from the same population. After a similarity is defined between objects an objective function is minimized. In many common clustering algorithms the objective is to minimize the dissimilarity between an object and the other objects within the cluster [55]. The specifics of the objective function are dependent on the application. In some applications the objective is to maximize the compactness of clusters along with maximizing the spread of clusters [59].

For all clustering methods, it is important to define the correct similarity between objects to ensure accurate clustering results. In this work it is proposed that if a similarity between damaging events can be defined, then multiple damaging events can be clustered to identify important trends in damaging events. In the following subsection, common similarity metrics and statistical tests for evaluating the similarity between events is provided.

5.2.3 Similarity and Statistical Evaluation Tests

In this work, it is desirable to group similar damaging events together, which is equivalent to rejecting non-similar events from a group. If an event is rejected from a group it can be considered an outlier with respect to the other events in the group. In a univariate data set there are numerous tests for identifying an outlier from other measures. One common method that has an extension to multivariate data is based on the ratio of the difference between the event in question, x_q, and the mean of the group, $\bar{x}$, with respect to the standard deviation of the samples, s, as shown in Eq. (5.4) [87].

$$r_q = \frac{|x_q - \bar{x}|}{s} \tag{5.4}$$

Once the distance, r_q, is calculated, it is compared to a threshold value to decide if the event q is declared an outlier, or not, with respect to the remainder of the data set. This same idea can be extended to multivariate data set where there are n_p variables for each event to be evaluated. The result of this extension is the Mahalanobis Distance Squared (MDS). The MDS uses the difference between each variable in event q with the mean value of each corresponding variable in the data set along with the inverse of the covariance matrix, $\mathbf{\Sigma}$, to define the likelihood of event q being an outlier. The MDS is provided in general terms in Eq. (5.5).

$$r_{mah}^2 = (\mathbf{x}_q - \bar{\mathbf{x}})^\intercal \mathbf{\Sigma}^{-1} (\mathbf{x}_q - \bar{\mathbf{x}}) \tag{5.5}$$

The MDS can be further generalized, to evaluate the similarity/dissimilarity between two random vectors. The extension of the MDS to random vectors is provided in Eq. (5.6) between the random vectors $\mathbf{x}_q$ and $\mathbf{x}_r$. The MDS is a commonly used distance metric between two random vectors because it provides a unitless and scale-invariant distance by weighting the difference in the mean vectors by the inverse of the pooled covariance between

the two events.

$$r_{mah}^2 = (\mathbf{x}_q - \mathbf{x}_r)^\intercal \, \mathbf{\Sigma}^{-1} \, (\mathbf{x}_q - \mathbf{x}_r) \tag{5.6}$$

One potential issue with using the Mahalanobis distance directly is that it can be affected by the population size between the events. Thus, depending on the difference in population sizes (number of samples) between the random vectors being compared it may be necessary to use an addition similarity metric, the likelihood (probability) that the random vectors, $\mathbf{x}_q$ and $\mathbf{x}_r$ are from the same population. A probability is established from a reference distribution associated with the Mahalanobis distances. If the off-diagonal terms of $\mathbf{\Sigma}$ are not significant then the Mahalanobis distance is equivalent to a Euclidean distance of a multivariate standard normal distribution. Then the reference distribution for the Mahalanobis distance is represented by a chi-squared distribution [88]. When the off-diagonal terms in $\mathbf{\Sigma}$ are significant and the standard mean and covariance estimates are used, typically the Mahalanobis distance follows a scaled beta distribution with parameters defined in terms of the number of samples used to define the Mahalanobis distance and the number of features characterizing the objects [89, 90]. Lastly, if the mean and covariance terms are defined independent of the samples, then the reference distribution for the Mahalanobis distance is represented by a F-distribution [91].

Once a reference distribution is established, then a rejection criteria can be used to decide if the null hypothesis, that $\mathbf{x}_q$ and $\mathbf{x}_r$ are from the same population, should be rejected or not. In the following section, a similarity analysis is developed using pseudo damage densities from damaging events. The similarity metric developed is cast in terms of distance or likelihood (probability) to allow for the statistical tests mentioned above to be used to evaluate the similarity/dissimilarity between damaging events.

5.3 Damaging Event Similarity Analysis

When multiple damaging events are identified using pseudo damage density, it is likely that some of the unique damaging events share similar damaging characteristics. The main contribution of this work is the development of an assessment method for identifying the similarity/dissimilarity between multiple damaging events using existing state-of-the-art statistical and clustering analysis tools to group multiple similar damaging events together. The first step in the assessment is defining a method for characterizing damaging events given that each damaging event contains a matrix of PD density samples for all load paths of interest, $\mathbf{d}_r$, as defined in Eq. (5.3). Once a method for characterizing each damaging event is defined a method for defining the similarity between damaging events is presented. Initially this similarity is shown between two damaging events. Then, the extension of the similarity between two damaging events to a cluster of damaging events is provided. Lastly, the process of applying the similarity analysis between clusters is discussed and the effect applications have on selection of clustering algorithms and defining the rejection criteria of two or more events being similar is provided.

5.3.1 Characterizing Damaging Events

For each damaging event, E_r, a set of realizations of PD density in each load path is known and defined by the matrix $\mathbf{d}_r$ in Eq. (5.3). Each row in the matrix corresponds to an individual load path and the PD density in each load path is considered a random variable. For each event, the mean and variation of the PD density is estimated from the samples captured in each row of $\mathbf{d}_r$. Load paths for an event may be dependent on one another, therefore the covariance between PD densities within an event must be identified. The mean values for PD density for a single r^{th} damaging event are summarized by the mean PD density vector $\bar{\mathbf{d}}_r$ in Eq. (5.7).

$$\bar{\mathbf{d}}_r = \begin{bmatrix} \bar{d}_1 & \cdots & \bar{d}_p & \cdots & \bar{d}_{n_p} \end{bmatrix}_r^{\intercal} \tag{5.7}$$

Each element of the vector $\bar{\mathbf{d}}_r$ corresponds to the sample mean of the PD density in a given load path. The variance of PD densities within each load path and amongst each load path is captured by a sample covariance matrix, $\mathbf{\Sigma}_r$. Each element in the sample covariance matrix corresponds to the sample covariance of the PD density from each row in $\mathbf{d}_r$ with respect to another row in $\mathbf{d}_r$.

5.3.2 Similarity Between Two Damaging Events

Consider two events, E_r and E_q, whose mean PD density vectors ($\bar{\mathbf{d}}_r$ and $\bar{\mathbf{d}}_q$) are known along with their sample covariance matrices based on the set of PD density samples for each event ($\mathbf{\Sigma}_r$ and $\mathbf{\Sigma}_q$). It is hypothesized that E_r and E_q are drawn from the same population. To test this hypothesis, the difference in the sample means between E_r and E_q, across all load paths, is used as a similarity measure. Sample means must be considered as a set, with a corresponding covariance matrix, since load paths for the same event are dependent on each other. A reference distribution is formed by assuming the null hypothesis is true. To form this reference distribution, an unbiased pooled estimate of the covariance matrix is formed as a weighted sum of the covariance matrices from the two events. The weighting is dependent on the sample size of the two events, N_r and N_q, for events E_r and E_q, respectively. In Eq. (5.8) the weighted sum to define the pooled covariance matrix is defined using the two known sample covariance matrices.

$$\mathbf{\Sigma}_{q,r} = \left(\frac{N_q - 1}{N_q + N_r - 2}\right)\mathbf{\Sigma}_q + \left(\frac{N_r - 1}{N_q + N_r - 2}\right)\mathbf{\Sigma}_r \tag{5.8}$$

Difference in the mean PD densities across all load paths with respect to the pooled covariance is calculated using the Mahalanobis distance. The Mahalanobis distance squared between E_r and E_q, $\rho_{q,r}$, is defined in Eq. (5.9) using the mean PD density vectors and the pooled covariance matrix defined in Eq. (5.8).

$$\rho_{q,r} = \left(\bar{\mathbf{d}}_q - \bar{\mathbf{d}}_r\right)^{\intercal} \mathbf{\Sigma}_{q,r}^{-1} \left(\bar{\mathbf{d}}_q - \bar{\mathbf{d}}_r\right) \tag{5.9}$$

The result for the Mahalanobis distance squared (MDS), $\rho_{q,r}$, is a distance that is unitless and scale-invariant. Since MDS is calculated using a weighted pool covariance matrix, the effect of population size on the variation is taken into account when determining the relative distance between the two events. The probability that E_r and E_q are drawn from different populations is calculated using the reference distribution for MDS, which is not only a function of the distance but also the number of degree of freedom in each of the events. Recall, each damaging event has a unique number of samples. In Eq. (5.9) a sample size weighted covariance matrix is used to address the difference in population size in terms of the variance. However, one MDS value cannot accurately be compared to another MDS value between events with a different sample size. To address this concern, $\rho_{q,r}$ is converted to a probability that the two events compared are from the same population. Typically, the Mahalanobis distance follows a scaled beta distribution when the standard mean and covariance estimates are used [89, 90]. The two parameters for the beta distribution are defined in terms of the number of load paths, n_p and a total population sample Λ that accounts for estimation of the average values defined as $\Lambda = N_q + N_r - 2$. The resulting Beta function similarity between $\rho_{q,r}$ and the distribution of Mahalanobis distance values, $P_{q,r}$, is provided in Eq. (5.10).

$$\frac{\Lambda}{(\Lambda-1)^2} P_{q,r} \sim f_{Beta}\left(\frac{\Lambda}{(\Lambda-1)^2} \rho_{q,r} \middle| \frac{n_p}{2}, \frac{\Lambda - n_p - 1}{2} \right) \tag{5.10}$$

Using the cumulative Beta distribution the probability that E_r and E_q are drawn from the same population of damaging events is known and referenced as $p_{q,r}$. Thus, two measures of similarity between E_r and E_q have been shown. The first is a distance measure provided in Eq. (5.9) and the second is a probability based on a Beta distribution provided in Eq. (5.10). The selection of the appropriate similarity metric is application specific. The MDS and probability based on the Beta distribution are introduced as options.

5.3.3 Extension to Multiple Damaging Events

Next define a cluster of events $\mathbb{E}_C$, as a set of damaging events, shown in Eq. (5.11), where the indices for the events in the cluster form the set C.

$$\mathbb{E}_C = \left\{ E_r | \quad r \in C \right\} \tag{5.11}$$

Each event in cluster $\mathbb{E}_C$ has a known mean PD density vector and covariance matrix. Now consider a new damaging event, E_q. It is hypothesized that E_q and all damaging events in the cluster $\mathbb{E}_C$ are drawn from the same population of damaging events. To test this hypothesis, an overall similarity between E_q and the cluster $\mathbb{E}_C$ is needed. To establish an overall similarity, the similarity between each rth event in $\mathbb{E}_C$ and E_q is investigated individually. For each rth event, a pooled covariance matrix, $\mathbf{\Sigma}_{q,r}$ is formed with E_q using Eq. (5.8) and a distance between the rth event and E_q, $\rho_{q,r}$, is calculated using Eq. (5.9).

The set of all distances between E_q and and the events in cluster $\mathbb{E}_C$ are stored in the set, $\boldsymbol{\rho}_{q,C}$, as illustrated in Eq. (5.12).

$$\boldsymbol{\rho}_{q,C} = \left\{\rho_{q,r} | \quad r \in C\right\} \tag{5.12}$$

For each distance in the set $\boldsymbol{\rho}_{q,C}$ a likelihood of the rth event in the cluster and E_q being drawn from the same population, $p_{q,r}$, is established using Eq. (5.10). Each calculated likelihood value is stored in the set $\mathbf{p}_{q,C}$, provided in Eq. (5.13).

$$\mathbf{p}_{q,C} = \left\{p_{q,r} | \quad r \in C\right\} \tag{5.13}$$

Thus, two sets of similarity measures between E_q and the cluster $\mathbb{E}_C$ are presented. The first similarity is in terms of distance, see Eq. (5.12), and the second similarity is in terms of probability, see Eq. (5.13). To establish an overall similarity between $\mathbb{E}_C$ and the new event E_q a method for linking all of the similarity metrics together (independent of the similarity type) is needed. An overall measure of similarity is established using a linkage criteria. A linkage criteria provides a single relative measure, $S_{q,C}$, between E_q and the cluster $\mathbb{E}_c$ using all similarity metrics between the new event, E_q, and the cluster, $\mathbb{E}_c$. Several linkage criteria currently exist such as: complete-linkage, single-linkage, and average linkage [92, 93, 94]. One of these linkage methods can be selected based on the application and used to define the respective similarity between an event and a cluster. In general, $S_{q,C}$ can be written as a function equivalent to the selected linkage criteria function, L, and the set of similarity metrics between E_q and the cluster $\mathbb{E}_c$, as defined in Eq. (5.14). The set of similarities, $\mathbf{s}_{q,C}$, in Eq. (5.14) corresponds to either distance, $\boldsymbol{\rho}_{q,C}$, as defined in Eq. (5.12) or probability, $\mathbf{p}_{q,C}$, as defined in Eq. (5.13).

$$S_{q,C} = L\left(\mathbf{s}_{q,C}\right) \tag{5.14}$$

Once a linkage criteria is selected and $S_{q,C}$ for the cluster is determined, the ability to reject the null hypothesis, that E_q and $\mathbb{E}_C$ are drawn from the same population, can be evaluated using $S_{q,C}$ along with an additional rejection criteria based on the application needs. The decision to reject the null hypothesis is dependent on the application along with other clustering constraints (*e.g.*, number of clusters, maximum dissimilarity, maximum variation in clusters, etc.). Taking into account the application needs, a rejection criteria can be established. An example of a rejection criteria is a specified threshold on the likelihood metric. Another less rigorous rejection criteria when more than one cluster is known is a minimax criteria (*e.g.*, minimum distance or maximum likelihood). For the minimax criteria the null hypothesis is rejected for all clusters that are not the solution to the cost function, and for the cluster that does solve the minimax cost function, the null hypothesis is not rejected. The minimax criteria is useful in applications where it is desirable to find the closest matching events even if the likelihood of the closest matching events is not necessarily considered statistically significant.

In application, a clustering algorithm is used to develop multiple clusters for a given set of damaging events. A review of current state-of-the-art clustering algorithms is provided

in Section 2.7, any of these algorithms can be implemented using the similarity between damaging events, defined in terms of distance or probability. In addition, a linkage criteria and rejection criteria for rejecting the hypothesis that two or more events are from the same population, as specified by the user, can be used to construct the a set of optimal clusters. The selection of a clustering algorithm and rejection criteria is application specific and should be based on the desired conclusions to be drawn or questions that need to be answered. To illustrate the application effect on clustering method and rejection criteria two examples are provided in the following section. It should be mentioned that the goal if this work is not to address all applications of the similarity analysis, it is to provide the method for establishing the similarity between damaging event and how current clustering methods can be leveraged to provide answers to important vehicle durability testing questions.

5.4 Examples

In this section two examples are presented to illustrate the applicability of the developed similarity analysis between clusters and illustrate the importance of selecting the correct clustering algorithm based on application. For both examples 25 damaging events are used. All of the events are identified from collected measurements of public highways across the United States. All road segments can be found in the Federal Highway Administration (FHWA) Long-term Pavement Program (LTPP) database [86]. To illustrate an additional point related to vehicle durability testing, in these examples the complete set of 25 events is subcategorized into two sets: Set 1 contains 5 damaging events and Set 2 contains the remaining 20 damaging events. This subcategorization of the damaging events is done to illustrate how the similarity analysis can be used to group multiple damaging events from different sets when both sets contain the same load paths of interest. In the examples presented herein, all events are from public roads. However, in application Set 1 could be considered damaging events from a private vehicle proving ground, while Set 2 is damaging events corresponding to a customer usage test on public roads. A table of the main characteristics for each of the damaging events used in this work is presented in Appendix C.

In the first example, Set 1 is considered a test set and Set 2 is considered the set of damaging events from a customer usage test. The goal of this example is to find the damaging event in Set 1 that best replicates the damage characteristics of each damaging event in Set 2. To perform this analysis, each damaging event in Set 1 is considered a cluster center. Therefore, there are five possible clusters that each event from Set 2 can be potentially grouped into. The likelihood of each event in Set 2 with respect to each event in Set 1 is calculated. Then, a minimax rejection criteria is used to identify the closest matching damaging event from Set 1 for each damaging event in Set 2. For reference, this example is analogous to having a set of damaging events from public roads (Set 2) that need to be replicated on a known set of proving ground events (Set 1).

In the second example, the complete set of 25 damaging events is investigated for partitioning

into multiple clusters based on a maximum likelihood threshold. The goal in this example is to find the damaging events within the two sets that are most similar to each other, independent of the set in which they belong. To achieve this goal a hierarchical clustering algorithm is used with a complete linkage criteria. The results of this example reveal groups of damaging events that are most similar to each other even if those groups contain multiple or no damaging events from one of the two sets. For reference, this example is analogous to having a set of damaging events from public roads and a set of damaging events from a private proving ground and needing to identify how similar the two sets are to each other. By performing this analysis it can reveal clusters where there are repetitive or missing damaging events from the proving ground set.

5.4.1 Example 1: Specifying Cluster Centers

In this example, a k-medoids clustering algorithm is used with an initial condition of five clusters with the medoid of each cluster corresponding to a single event from Set 1 (*e.g.*, Cluster 1 has a medoid with the damage characteristics of Event 1 from Set 1). After each cluster medoid is defined, the likelihood of an event from Set 2 being drawn from the same population as each damaging event from Set 1 is calculated using Eq. (5.9) and a corresponding Beta reference distribution presented in Eq. (5.10). Thus, for each event in Set 2, five different overall cluster similarity values, $S_{q,C}$, are calculated. To determine if the hypothesis that the event from Set 2 was drawn from the same population as each corresponding event in Set 1, the rejection criteria of maximum likelihood was used. Therefore, the hypothesis was rejected for the four events in Set 1 that did not have the highest likelihood. For the fifth cluster, since a greater likelihood value was not present, the hypothesis was not rejected and therefore the event from Set 2 was clustered with that Set 1 event. The resulting clusters of events of Set 2 with respect to individual events from Set 1 are summarized in Table 5.1.

Each column of Table 5.1 corresponds to one of the five resulting clusters. In the second row of Table 5.1 the corresponding event from Set 1 that characterized the cluster initially is presented and in the subsequent rows the events from Set 2 that were identified as most similar to the characterizing event from Set 1 are presented. By identifying these five clusters, if it is necessary to recreate an event from Set 2 using the events from Set 1, then find the column in Table 5.1 in which the Set 2 event is located and the corresponding event from Set 1 in that column is the event that will most closely replicate the damage characteristics. For example, to reproduce the damage characteristics of E2-5 using the events in Set 1, E1-1 should be used. It should be noted, that by clustering using k-medoids with a minimax rejection criteria, the clustering results in Table 5.1 are optimal, in that each event from Set 2 is clustered with the most similar event from Set 1 based on pseudo damage density. In the following example a different clustering objective is selected to illustrate how the damaging events can be clustered differently to provide a solution to a different set of questions.

Table 5.1: Summary of Event Grouping

Cluster	1	2	3	4	5
Set 1:	E1-1	E1-2	E1-3	E1-4	E1-5
Set 2:	E2-5	E2-19	E2-18	E2-1	E2-11
	E2-6			E2-2	E2-12
	E2-7			E2-3	E2-17
	E2-8			E2-4	
	E2-9			E2-15	
	E2-10				
	E2-13				
	E2-14				
	E2-16				
	E2-20				

5.4.2 Example 2: Clustering Using Likelihood Threshold

In the second example, the complete set of 25 events is used in the clustering analysis and no distinction is made between the subset in which each event belongs. This allows for clusters be defined independent of the subset in which the events originated from. When the damaging events are clustered this way, the most similar damaging events are clustered together even if those events are from the same set or if a cluster lacks a damaging event from one of the subsets. For the clustering algorithm, no initial condition was provided regarding the number of clusters or the characteristics of the clusters. To define the clusters a hierarchical clustering algorithm was used to iteratively include damaging events in clusters based on highest probability (or lowest dissimilarity). For the clustering algorithm, the likelihood of two events being drawn from the same population is calculated using Eq. (5.10). When comparing a new damaging event to an existing cluster of damaging events a complete linkage criteria was used to define the linkage function in Eq. (5.14). A complete linkage criteria adds a new damaging event to a cluster based on the maximum dissimilarity between an event and each event already in the cluster.

The connections formed between similar damaging events on each iteration in a hierarchical clustering algorithm are best illustrated through a dendrogram. A dendrogram contains branches that connect similar events together and highlight the linkages between damaging events through horizontal lines. The resulting dendrogram for the 25 damaging events is provided in Figure 5.2. In Figure 5.2 the order of the damaging events across the x-axis corresponds to the damaging events that are most similar to each other. The solid horizontal lines indicate a linkage between events/clusters. In the dendrogram in Figure 5.2 the level of dissimilarity associated with the linkage between each new damaging event and an existing cluster is represented by the y-axis value of the horizontal line. For example, the linkage

between event E1-3 and E2-18 is much lower than the linkage between E2-9 and E2-10. The vertical lines are used to help show the connection between each event along the x-axis and each corresponding linkage.

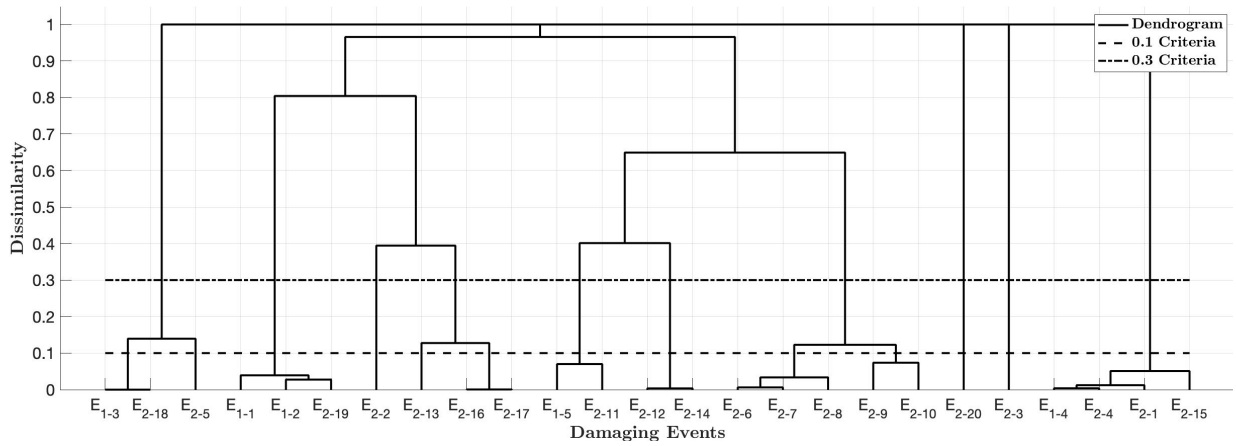


Figure 5.2: Dendrogram illustrating the clustering results for Example 2.

To decide how to separate the damaging events into clusters a maximum dissimilarity threshold is used. This maximum threshold is illustrated as a horizontal line across the dendrogram. In Figure 5.2, two thresholds are illustrated. The first threshold is a low level of dissimilarity of 0.1 (10%) illustrated by the dashed line in Figure 5.2. The second threshold is a higher dissimilarity value of 0.3 (30%), illustrated by the dash-dot line in Figure 5.2. Each branch in the dendrogram is cut at a level corresponding to the threshold to produce clusters that contain dissimilarities between events that are no greater than the corresponding threshold. By cutting a branch in a dendrogram, all linkages (horizontal lines) in the dendrogram above the threshold do not satisfy the rejection criteria for defining clusters and are therefore insignificant and rejected from defining new events as members of clusters.

When a maximum allowable dissimilarity threshold of 0.1 is selected 13 unique clusters are identified. The events that are considered components of each of these clusters is provided in Table 5.2. By selection of this threshold it follows that all events within each cluster have a dissimilarity that is no greater than 0.1. When analyzing the data, the user can see that in Cluster 1 has one event each from Set 1 and Set 2. Therefore, to reproduce events E2-18, event E1-3 from Set 1 should be used. However, this is not the case for all resulting clusters. For example Clusters 4, 5, 6, 7, and 8 contain only events from Set 2, indicating that no damaging event from Set 1 can be used to replicate the damage characteristics from those events. Lastly, in Cluster 3 it can be seen that two events from Set 1, E1-1 and E1-2 are grouped together. This indicates that two of the events from Set 1 are very similar and either one can be used to replicate the damage characteristics of the other or to replicate the damage from E2-19.

Table 5.2: Clustering for 0.1 Criteria

Cluster	Events in the Cluster			
1	E1-3	E2-18		
2	E2-5			
3	E1-1	E1-2	E2-19	
4	E2-2			
5	E2-13			
6	E2-16	E2-17		
7	E1-5	E2-11		
8	E2-12	E2-14		
9	E2-6	E2-7	E2-8	
10	E2-9	E2-10		
11	E2-20			
12	E2-3			
13	E1-4	E2-4	E2-1	E2-15

When a maximum allowable dissimilarity of 0.3 is selected the number of unique clusters is reduced to 9. The resulting damaging events in these 9 clusters is presented in Table 5.3. The reduction in number of clusters is a result of combining similar clusters from Table 5.2. For example, Clusters 1 and 2 from Table 5.2 are combined into a single cluster, Cluster 1 in Table 5.3. While the likelihood threshold is increased, there are still clusters formed which do not have a damaging event from both sets. For example, Cluster 3 in Table 5.3 contains six damaging events from Set 2, but none from Set 1. Therefore, no event in Set 1 can replicate the damage from these six events. If it is desirable to add a new damaging event to Set 1, the damage characteristics associated with Cluster 3 would be useful for defining the new event that should be added. Lastly, one notable feature shared between Table 5.2 and Table 5.3 is that E 2-20 is not clustered with any other events even at both thresholds. This could indicate that this event is a potential outlier and does not exist often or that the damage characteristics are unique and should be used to influence the addition of damaging events to Set 1.

Table 5.3: Clustering for 0.3 Criteria

Cluster	Events in the Cluster				
1	E1-3	E2-18	E2-5		
2	E1-1	E1-2	E2-19		
3	E2-2				
4	E2-13	E2-16	E2-17		
5	E1-5	E2-11			
6	E2-12	E2-14			
7	E2-6	E2-7	E2-8	E2-9	E2-10
8	E2-20				
9	E2-3				
10	E1-4	E2-4	E2-1	E2-15	

The clusters provided in Table 5.2 and Table 5.3 highlight the importance of selecting the correct threshold criteria for deciding if an event belong to a cluster or not. For this example, a reduction of four clusters was achieved when the threshold was relaxed, however in a larger dataset this effect could be amplified. Thus, it is important to consider what level of similarity is necessary for the application.

5.5 Discussion

The contribution of this work is development of an assessment method for identifying the similarity/dissimilarity between multiple damaging events and using the similarity to group multiple events together into a cluster. To perform the similarity analysis it is assumed that for each damaging event, a set of PD density samples are known for a specified number of load paths of interest. To perform the similarity analysis each damaging event is characterized by an average PD density vector, containing the average PD density of each load path, and a sample covariance matrix, capturing the variability and correlation between the PD density samples in each load path. Next, the similarity between a damaging event and a cluster of damaging events is developed to cast the problem in terms of distance or probability. By calculating the similarity any existing clustering algorithm can be used to determine an optimal clustering of damaging events. The development of this work is a statistical test, therefore the null hypothesis that an event is drawn from a cluster of damaging events can never be accepted, it can only be rejected based on proof through a clustering analysis.

When defining the similarity between events an important factor to consider is that the sample size between events being compared is not consistent. The differences in sample size will cause irregularities in Mahalanobis distance because of the change in number of degrees of freedom (*e.g.*, a Mahalanobis distance of 1.0 in from a sample size of 100 is not

the same as a Mahalanobis distance of 1.0 from a sample size of 10). However, by converting the Mahalanobis distance to a probability, the sample size is taken into account allowing for varying sample sizes to be compared. In applications where the sample size is not significantly different between events being grouped, the Mahalanobis distance is an acceptable similarity metric. When the sample size varies significantly then a reference distribution should be used to define a likelihood.

In this work, two examples using the same set of 25 damaging events are presented. Between the two examples, two different clustering algorithms and criteria for rejecting the hypothesis that two or more damaging events are from the same cluster are used. While the same set of damaging events are used in both examples, the resulting cluster solutions and conclusions drawn are different between the two examples. However, both examples provide meaningful solutions based on the desired questions to answer. The two example presented herein highlight that the application of interest should inform the type of clustering analysis used and the criteria for rejecting an event from a cluster.

In the examples presented herein, a hard clustering algorithm was used to determine the optimal clusters between damaging events. In future work, one area of development is the extension of this work to fuzzy clustering. In fuzzy clustering it is possible that a damaging event would be considered a component of multiple clusters. The similarity metric developed in this work would allow for the methodology developed to be extended to fuzzy clustering.

It is envisioned that the results from this work can be extended to development and improvement of durability tests. The results from a clustering analysis like those presented in this work can be used to influence the damaging events that a typical customer's vehicle must be able to survive. Specifically, the events in a proving ground needed to replicate the damage from public roads can be identified. In addition, through clustering, any repeat or missing types of damaging events from a proving ground can be identified and addressed in future improvements of the proving grounds. Lastly, by grouping damaging events based on the damage characteristics a durability testing cycle can be formed based on type of damaging events experienced, along with a measure of the number of times a customer is likely to experience each type of damaging event.

5.6 Conclusion

The process of assessing the similarity between damaging events is difficult for two reasons. First, the overall length of damaging events is not consistent between events. In this work, this issue is overcome by using PD density to assess the similarity between damaging events. Second, the accumulation of damage is not consistent from load path to load path. This second issue is addressed by using discrete samples of PD density in multiple load paths of interest. The results of this work is a method for defining a similarity, in terms of distance or probability, between two or more damaging events. Then, this similarity metric can be

used with any current state-of-the-art clustering algorithm to provide solutions to important vehicle durability testing questions. The process of clustering damaging events is important because it allows for additional characterization of damaging events and comparison of damaging events between two sets of data (*e.g.*, public road events and proving ground events). By developing connections between unique damaging events and investigating the relationship between sets of events a vehicle durability cycle can be developed based on clusters/types of damaging events.

Chapter 6

Summary, Conclusions and Future Work

6.1 Summary

The main objective of this dissertation is to identify, characterize, and group damaging events based on their damage characteristics across multiple load paths in a vehicle system. Based on existing literature, theoretical developments have been formulated in general terms to allow for scalability and variation between vehicle models and classes. In this work the following objectives have been achieved:

1. Theoretical development of pseudo damage density, a measure that allows for unique separable damaging events to be identified from a known road profile.
2. Development of an optimization problem for identifying damaging events that identifies the optimal balance between pseudo damage accounted for and the length of road required.
3. Development of a statistical analysis for testing the hypothesis that two or more separable damaging events are from the same population and using current state-of-the-art clustering algorithms to draw conclusions based on the application needs.

6.2 Conclusions

The main conclusions from this work are:

- Identification of the contribution a single excitation makes to the total accumulated pseudo in load path of a vehicle is complex because a vehicle is dynamic system and

therefore reacts at varying levels to all prior road excitations. This complex issue was overcome by modeling the Localized Pseudo Damage (LPD) as a linear combination of the relative contributions a single excitation makes to every other location along the road profile.

- A theoretical development of pseudo damage density for any load path in a vehicle system is developed. The theoretical development in this work does not make any vehicle model or road profiles assumptions ensuring that it is applicable to a wide range of models.

- The developed LPD model is proven to demonstrate causality and differentiation between damage inducing and damage reducing excitations. The causality condition is presented through a theoretical proof involving an impulse and step excitation and the the differentiation between types of excitations is demonstrated through a dual impulse example.

- A partitioning method that is insensitive of the number of load paths is developed to separate a road profile into active and inactive regions.

- Unique and separable damaging events are identified from a road profile (more specifically the set of active regions) based on an optimization function that maximizes the choice of damaging events based on maximizing the benefit (fraction of pseudo damage retained by the damaging events) relative to the cost (fraction of locations retained as damaging events).

- A null hypothesis is proposed that two separate damaging events are from the same population and a statistical test is developed to evaluate the null hypothesis based on the mean PD density values and the variation of PD densities within a damaging event.

- A null hypothesis is proposed that a damaging event is from the same population as a cluster of multiple damaging events and an extension of the statistical tests developed for assessing two damaging events is provided to evaluate the null hypothesis.

- An example distance similarity metric is developed using the Mahalanobis distance and an example probability metric is developed using a Beta reference distribution on the possible values of the Mahalanobis distance.

- The effect of application on the selection of a clustering algorithm and a rejection criteria for the null hypothesis is demonstrated through two clustering examples using pseudo damage density in multiple load paths.

6.3 Future Work

The following research topics are proposed as areas of future work:

- **Development of a Durability Test Using Types of Events**
  Using the similarity analysis developed in Chapter 5 along with a clustering algorithm, a durability test based on event type can be formed. When clustering of damaging events is performed such as that shown in the second example in Chapter 5 (see Section 5.4.2) a set of customer events from public roads are grouped with private proving ground events and 1 of 3 resulting classifications of clusters, with respect to the number of proving ground events in the cluster can be established. When one proving ground event is grouped in the cluster the number of times to repeat this event can be identified using the other customer events in the cluster. When more than one proving ground event is grouped in the cluster, a sub clustering can be used to identify which customer events should be paired with each proving ground event in the cluster. Then the resulting sub clustering can be used to determine the number of times to repeat each respective proving ground event. Lastly, when no proving ground events are present in the cluster then additional analysis must be performed to identify the best set of event(s) and number of repeats to replicate each customer event in that cluster. The number of times to repeat a proving ground event is established using the likelihood of a customer experiencing each customer event. This likelihood can be established using a measure like the Annual Average Daily Traffic (AADT) value.

- **Pseudo damage density sensitivity to vehicle parameters and vehicle types**
  As a method for validating the theoretical work developed in this dissertation, a Golden quarter car model with two load paths of interest was used. While this model is representative of typical passenger cars it does not accurately model all types of production vehicles. Therefore, it would be beneficial to identify the sensitivity vehicle models and parameters have on the set of identified damaging events. Changes in vehicle parameters can potentially affect the location and severity of damaging events to the vehicle. Therefore, a sensitivity analysis should be performed to identify the sensitivity of the location and magnitude of damaging events to changes in vehicle parameters and models. This investigation may lead to development of a vehicle dependent damaging event index (*e.g.*, an identified event affects a sedan in a certain way and the same event affects a truck in a different way).

- **Damaging events based on road type**
  Public roads that customers drive on, are typically classified by road type. When the Federal Highway Administration assess roads, they are classified into one of the following road types: interstate, principal arterial, freeway/expressway, minor arterial, major collector, minor collector, and local. If multiple road profiles are known from each road type then the similarity analysis presented in Chapter 5 can be used to identify types of damaging events for each each road type. Once clusters are formed

for each type of road, then a comparison can be made between damaging event clusters to identify any similar damaging event clusters between road types and any damaging event clusters that are road type dependent.

- **Development of representative damaging events per cluster**
  In Section 5.4.2 it was demonstrated that clusters of damaging events can be formed which do not contain corresponding durability test events (*e.g.*, proving ground events). For these clusters it may be desirable to develop a new proving ground event. Similarly, it may desirable to create representative events for each damaging event cluster that can be used in computer simulations or lab testing. The representative events will match the pseudo damage density characteristics of the cluster and be representative of all events in the cluster. One component that will need to be addressed when developing the representative events is the duration of the representative event. When performing durability analysis it is desirable to minimize the duration of durability test but still maintain the same damage characteristics as the customer usage events. Therefore, a method of achieving the shortest representative event should be addressed in this future work.

- **Development of road surface pseudo damage density**
  One of the contributions of this dissertation is the development of a pseudo damage density for vehicle systems. Through the developments in this work local regions which have have a strong contribution to the damage accumulated in one or more load paths for a vehicle system are identified. However, the principals in this development can be extended to road surfaces using vehicle tire forces. A road surface pseudo damage density can be developed using the tire forces to identify local regions of excessive damage from vehicles. To ensure the density is representative of the real-world the tire forces and corresponding road surface pseudo damage density should be identified for multiple types of vehicles (*e.g.*, passenger cars, small/medium trucks, and large trucks). Then the overall road surface pseudo damage density will be based on a weighted combination of these vehicle dependent road pseudo damage densities. The development of a road surface PD density would lead to identifying locations of the road that strongly contribute to the damage in the road at other locations.

- **Real proving ground events**
  In Chapter 5, five damaging events that were identified from roads in the FHWA LTPP database were used to represent durability test events (*i.e.*, proving ground events). In future work it would be beneficial to perform the same analysis in Chapter 5 but define the events in Set 1 using events from a proving ground.

- **Comparison to instrumented vehicles**
  To validate the theoretical development in this work, publicly available road profiles from the FHWA LTPP database were used and coupled with a golden quarter car vehicle model. It would be beneficial to perform experimental validation of the de-

velopments in this work by comparing the results from a measured road to the data collected from an instrumented vehicle driven over the road profile.

- **Evaluation of run-to-run variance**
  When performing physical durability test there is run-to-run variance when performing repeat runs of a damaging event. It would be beneficial to understand how this run-to-run variance affects identified events along with the characteristics of the identified events.

Research in other areas:
In Appendix D the idea of using Delaunay triangles to estimate the transverse and longitudinal spacing between road surface measurements is presented. This idea was originally developed during a sponsored Federal Highway Administration (FHWA) Transverse Pavement Profiler assessment, however due to time constraint the methodology was never completely developed and had to be put on hold.

- Finalize the methodology for using Delaunay triangulation to determine road surface measurement spacing in the transverse, longitudinal, and vertical directions for any provided road surface point cloud.
- Use of Delaunay triangulation to identify the potential of an individual road surface measurement being an outlier. When point cloud measurements are acquired of a road surface all measurements are linked together in that they are measurements of a single surface. It is hypothesized that for a point cloud measurement of a road surface a Delaunay triangulation can be established and the length and orientation of the Delaunay edges can be used to identify any potential outliers.

Bibliography

[1] U.S. Department of Transportation, *National Transportation Statistics*. U.S. Department of Transportation Bureau of Transportation Statistics, 2018.

[2] K. Bogsjo, "Development of analysis tools and stochastic models of road profiles regarding their influence on heavy vehicle fatigue," *Vehicle Mechanics and Mobility*, 2006.

[3] J. B. Ferris and G. L. Larsen, "Establishing chassis reliability testing targets based on road roughness," *International Journal of Maerials and Product Technology*, 2002.

[4] E. Zamora-Alvarez, "A discrete roughness index for longitudinal road profiles," Master's thesis, Virginia Tech, 2015.

[5] F. Hveem, "Devices for recording and evaluating pavement roughness," *Highway Research Board Bulletin*, vol. 264, pp. 1–26, 1960.

[6] T. Gillespie, M. Sayers, and L. Segel, "Calibration and correlation of response-type road roughness measuring systems," tech. rep., National Cooperative Highway Research Program Transportaiton Research Board National Research Council, 1980.

[7] J. Kern and J. Ferris, "Development of a 3-d terrain measurement system part i: Equipment setup," in *Proceedings of ISTVS: Innovations in Terrain and Vehicle Systems, ISTVS*, 2007.

[8] S. Wagner, J. Kern, W. Israel, and J. Ferris, "Development of a 3-d vehicle-terrain measurement system part ii: Signal processing and validation," in *Proceedings of ISTVS: Innovations in Terrain and Vehicle Systems, ISTVS*, 2007.

[9] M. W. Sayers, "Two quarter-car models for defining road roughness. iri and hri," *Transportation Research Record*, 1989.

[10] M. W. Sayers, "On the calculation of international roughness index from longitudinal road profile," *Transportation Research Record*, 1995.

[11] H. M. Chemistruck, Z. R. Detweiler, J. B. Ferris, A. A. Reid, and D. J. Gorsich, "Review of current developments in terrain characterization and modeling," in *Modeling and Simulation for Military Operations IV*, vol. 7348, pp. 1–12, 2009.

[12] R. A. C. Capuruço, T. Hegazy, S. L. Tighe, and S. Zaghloul, "Full-car roughness index as summary roughness statistic," *Transportation Research Record*, vol. 1905, no. 1, pp. 148–156, 2005.

[13] O. Kropáč and P. Múčka, "Indicators of longitudinal unevenness of roads in the usa," *International Journal of Vehicle Design*, 2008.

[14] O. Kropáč and P. Múčka, "Non-standard longitudinal profiles of roads and indicators for their characterisation," *International Journal of Vehicle Design*, 2004.

[15] M. W. Sayers and S. M. Karamihas, "Estimation of rideability by analyzing longitudinal road profile," *Transportation Research Record*, 1996.

[16] A. T. Papagiannakis and B. Raveendran, "International standards organization-compatible index for pavement roughness," *Transportation Research Record: Journal of the Transportation Research Board*, 1998.

[17] J. G. Howe, "Quarter car model stress analysis for terrain/road profile ratings," *International Journal of Vehicle Design*, vol. 36, no. 2-3, pp. 248–269, 2004.

[18] O. Kropáč and P. Múčka, "Be careful when using the international roughness index as an indicator of road unevenness," *Journal of Sound and Vibration*, vol. 287, no. 4-5, pp. 989–1003, 2005.

[19] K. Bogsjö and I. Rychlik, "Vehicle fatigue damage caused by road irregularities," *Fatigue and Fracture of Engineering Materials and Structures*, 2009.

[20] I. Zaabar and K. Chatti, "New mechanistic-empirical approach for estimating the effect of roughness on vehicle durability," *Transportation Research Board*, 2011.

[21] ISO, "Mechanical vibration - road surface profiles - reporting of measured data," Standard 8608: 1995(E), International Organization for Standardization, Geneva, Switzerland, 1995.

[22] Y.-L. Lee, J. Pan, R. Hathaway, and M. Barkey, *Fatigue Testing and Analysis Theory and Practice*. Elsevier, 2005.

[23] I. Rychlik, "Simulation of load sequences from rainflow matrices: Markov method," *International Journal of Fatigue*, 1996.

[24] R. M. Wetzel, *A method of fatigue damage analysis*. PhD thesis, University of Waterloo, Ontario, Canada, 1971.

[25] S. D. Downing and D. F. Socie, "Simple rainflow counting algorithms," *International Journal of Fatigue*, 1982.

[26] ASTM International, "Standard practices for cycle counting in fatigue analysis," *ASTM Standard E1049*, 1985.

[27] R. C. Rice and *et al.*(Eds.), *Fatigue Design Handbook*. Society of Automotive Engineers, 3 ed., 1997.

[28] C. Amzallag, J. P. Gerey, J. L. Robert, and J. Bahuaud, "Standardization of the rainflow counting method for fatigue analysis," *Fatigue*, 1993.

[29] D. Schütz, H. Klätschke, H. Steinhilber, P. Heuler, and W. Schütz, "Standardized load sequences for car wheel suspension components - car loading standard carlos," *LBF Report No FB-191, 1990/IABG Report No TF-2695, LBF/IABG, Germany*, 1990.

[30] G. E. Leese and D. Socie, "Multiaxial fatigue: Analysis and experiments - ae 14," *Society of Automotive Engineering*, 1988.

[31] R. J. Anthes, "Modified rainflow counting keeping the load sequence," *International Journal of Fatigue*, 1997.

[32] W. Ramberg and W. R. Osgood, "Description of stress–strain curves by three parameters," *Technical Note No. 902, NACA*, 1943.

[33] H. J. French, "Fatigue and hardening of steels," *Transactions, American Society of Steel Treating*, pp. 899–946, 1933.

[34] B. F. Langer, "Fatigue failure from stress cycles of varying amplitude," *ASME Journal of Applied Mechanics*, 1937.

[35] J. B. Kommers, "The effect of overstressing and understressing in fatigue," *Proceedings, American Society for Testing and Materials*, pp. 249–268, 1938.

[36] A. Fatemi and L. Yang, "Cumulatie fatigue damage and life prediction theories: a survey of the state of the art for homogeneous materials," *International Journal of Fatigue*, 1997.

[37] L. C. Lim, Y. K. Tay, and H. S. Fong, "Fatigue damage and crack nucleation mechanisms at intermediate strain amplitudes," *Acta Metallurgica et Materialia*, 1990.

[38] L. F. Coffin, "Design aspects of high-temperature fatigue with particular reference to thermal stresses," *Transactions of the ASME*, 1956.

[39] E. E. Baldwin, G. J. Sokol, and L. F. Coffin, "Cyclic strain fatigue studies on aisi 347 stainless steel," *Proceedings, American Society for Testing and Materials*, 1957.

[40] K. J. Miller, "An experimental linear cumulative-damage law," *Journal of Strain Analysis*, 1970.

[41] F. E. Richart and N. M. Newmark, "A hypothesis for the determination of cumulative damage in fatigue," *Proceedings, American Society for Testing and Materials*, 1948.

[42] S. M. Marco and W. L. Starkey, "A concept of fatigue damage," *Transactions of the ASME*, 1954.

[43] E. Kobayashi, H. Mochizuki, H. Doi, T. Yoneyama, and T. Hanawa, "Fatigue life prediction of biometical titanium alloys under tensile/torsional stess," *Materials Transations*, vol. 47, no. 7, pp. 1826–1831, 2006.

[44] C. J. Dodds and J. D. Robson, "The description of road surface roughness," *Journal of Sound and Vibration*, vol. 31, no. 2, pp. 175–183, 1973.

[45] B. Bruscella, V. Rouillard, and M. Sek, "Analysis of road surface profiles," *Journal of Transportation Engineering*, 1999.

[46] V. Rouillard, B. Bruscella, and M. Sek, "Classification of road surface profiles," *Journal of Transportation Engineering*, 2000.

[47] O. Kropáč and P. Múčka, "Specification of obstacles in the longitudinal road profile by median filtering," *Journal of Transportation Engineering*, 2011.

[48] J. B. Ferris, "Characterising road profiles as markov chains," *International Journal of Vehicle Design*, 2004.

[49] H. Chemistruck and J. B. Ferris, "Compact models of terrain surfaces," in *ASME 2010 Dynamic Systems and Control Conference*, pp. 125–132, American Society of Mechanical Engineers, 2010.

[50] W. Frawley, G. Piatetsky-Shapiro, and C. Matheus, "Knowledge discovery in databases: An overview," *AI Magazine*, vol. 13, no. 3, pp. 57–57, 1992.

[51] M. S. Aldenderfer and R. K. Blashfield, *Quantitative Applications in the Social Sciences: Cluster Analysis*. Thousand Oaks, CA: SAGE Publications, 1984.

[52] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification and Scene Analysis*, vol. 3. Wiley New York, 1973.

[53] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Prentice-Hall, Inc., 1988.

[54] Y.-L. Zhou, N. Maia, R. Sampaio, and M. A. Wahab, "Structural damage detection using transmissibility together with hierarchical clustering analysis and similarity measure," *Structural Health Monitoring*, vol. 16, no. 6, pp. 711–731, 2017.

[55] X. Jin and J. Han, *Encyclopedia of Machine Learning.* Springer US, 2010.

[56] A. Ahmad and L. Dey, "A k-means clustering algorithm for mixed numeric and categorical data," *Data & Knowledge Engineering*, vol. 63, no. 2, pp. 503–527, 2007.

[57] C. Sammut and G. Webb, *Encyclopedia of Machine Learning.* Springer Science+Business Media LLC, 2010.

[58] M. R. Rezaee, B. P. Boudewijn, and J. H. Reiber, "A new cluster validity index for the fuzzy c-mean," *Pattern Recognition Letters*, vol. 19, no. 3-4, pp. 237–246, 1998.

[59] M. Halkidi and M. Vazirgiannis, "Clustering validity assessment: Finding the optimal partitioning of a data set," in *Proceedings - IEEE International Conference on Data Mining, ICDM*, pp. 187–194, 2001.

[60] Y. Liu, Z. Li, H. Xiong, X. Gao, and J. Wu, "Understanding of internal clustering validation measures," in *2010 IEEE International Conference on Data Mining*, pp. 911–916, IEEE, 2010.

[61] C. Altmann and J. Ferris, "Theoretical development of localized pseudo damage," *SAE International Journal of Passenger Cars - Mechanical Systems*, vol. 13, no. 1, 2020.

[62] C. Altmann and J. Ferris, "Customer usage based on pseudo damage," in *Proceedings of the ASME 2017 Dynamic Systems and Control Conference*, vol. 3, 2017.

[63] I. Rychlik, "On the 'narrow-band' approximation for expected fatigue damage," *Probabilistic Engineering Mechanics*, vol. 8, 1993.

[64] I. Rychlik, "A new definition of the rainflow cycle counting method," *International Journal of Fatigue*, vol. 9, 1987.

[65] L.-H. Zhao, J.-W. Yu, T. Chen, J. Li, and S.-L. Zheng, "Development of accelerated durability tests for rear suspension components under failure-correlated load," *Journal of Testing and Evaluation*, vol. 46, no. 5, pp. 1862–1873, 2018.

[66] A. Halfpenny, "Methods for accelerating dynamic durability tests," in 9th *International Conference on Recent Advances in Structural Dynamics*, 2006.

[67] L. Wang, R. Burger, and A. Aloe, "Considerations of vibration fatigue for automotive components," *SAE International Journal Commercial Vehicles*, 2017.

[68] D. Chindamo, M. Gadola, and F. Marchesin, "Reproduction of real-world road profiles on a four-poster rig for indoor vehicle chassis and suspension durability testing," *Advances in Mechanical Engineering*, vol. 9, 2017.

[69] W. Ferreira, T. Meehan, V. Cardoso, and N. Bishop, "A comparative study of automotive system fatigue models processed in the time and frequency domain," tech. rep., SAE Technical Paper 2016-01-0377, 2016.

[70] S. Pereira, *On the utilization of Simultaneous Localization and Mapping(SLAM) along with vehicle dynamics in Mobile Road Mapping Systems.* PhD thesis, Virginia Tech, 2019.

[71] C. Koch, K. Doycheva, B. Akinci, and P. Fieguth, "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure," *Advanced Engineering Informatics*, vol. 29, pp. 196–210, April 2015.

[72] I. Pitas, *Digital Image Processing Algorithms and Applications.* John Wiley and Sons, 2000.

[73] S. Li, C. Yuan, L. D., and H. Cai, "Integrated processing of image and gpr data for automated pothole detection," *Journal of Computing in Civil Engineering*, vol. 30, no. 6, 2016.

[74] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Advanced Engineering Informatics*, vol. 25, pp. 507–515, August 2011.

[75] M. Jahanshahi, F. Jazizadeh, S. Masri, and B. Becerik-Gerber, "Unsupervised approach for autonomous pavement-defect detection and quantification using an inexpensive depth sensor," *Journal of Computing in Civil Engineering*, vol. 27, no. 6, 2013.

[76] Z. Zhang, X. Ai, C. Chan, and N. Dahnoun, "An efficient algorithm for pothole detection using stereo vision," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014.

[77] M. Fischler and R. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Association for Computing Machinery*, vol. 24, no. 6, 1981.

[78] A. Mikhailiuk and N. Dahnoun, "Real-time pothole detection on tms320c6678 dsp," in *IEEE International Conference on Imaging Systems and Techniques (IST)*, 2016.

[79] E. Zamora-Alvarez, J. Ferris, D. Scott, and E. Horn, "Development of a discrete roughness index for longitudinal road profiles," *International Journal of Pavement Engineering*, vol. 19, no. 12, pp. 1043–1052, 2018.

[80] M. Chaika, D. Gorsich, and T. Sun, "Some statistical tests in the study of terrain modelling," *International Journal of Vehicle Design*, 2004.

[81] P. Andrén, "Power spectral density approximations of longitudinal road profiles," *Vehicle Design*, vol. 40, pp. 2–14, 2006.

[82] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, "Real time pothole detection using android smartphones with accelerometers," in *International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)*, IEEE, 2011.

[83] M. Ghadge, D. Pandey, and D. Kalbande, "Machine learning approach for predicting bumps on road," in *International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, IEEE, 2015.

[84] F. Serag, B. van der Zwaag, A. Dilo, T. Luarasi, and P. Havinga, "Roads: A road pavement monitoring system for anomaly detection using smart phones," in *Proceedings of the 1st International Workshop on Machine Learning for Urban Sensor Data, SenseML*, 2014.

[85] J. Ren and D. Liu, "Pads: A reliable pothole detection system using machine learning," in *Qiu M. (eds) Smart Computing and Communication. SmartCom 2016*, 2017.

[86] U.S. Department of Transportation Federal Highway Administration, "Ltpp infopave."

[87] K. Worden, G. Manson, and N. Fieller, "Damage detection using outlier analysis," *Journal of Sound and Vibration*, vol. 229, no. 3, pp. 647–667, 1999.

[88] B. Manly, *Multivariate Statistical Methods: A Primer*, vol. 3. Chapman & Hall / CRC Press, 2005.

[89] J. Hardin and D. Rocke, "The distribution of robust distances," *Journal of Computational and Graphical Statistics*, vol. 14, no. 4, pp. 928–946, 2005.

[90] D. Ververidis and C. Kotropoulos, "Gaussian mixture modeling by exploiting the mahalanobis distance," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2797–2811, 2008.

[91] C. J. McLachlan, "Mahalanobis distance," *Resonance*, vol. 4, no. 6, pp. 20–26, 1999.

[92] D. L. Massart, "The interpretation of analytical chemical data by the use of cluster analysis," tech. rep., 1983.

[93] G. N. Lance and W. T. Williams, "A general theory of classificatory sorting strategies: 1. hierarchical systems," *The Computer Journal*, vol. 9, no. 4, pp. 373–380, 1967.

[94] SAS System, *SAS/STAT(R) 9.2 User's Guide*, 2nd ed., September 2009.

[95] J. B. Kommers, "The effect of overstress in fatigue on the endurance life of steel," *Proceedings, American Society for Testing and Materials*, 1945.

[96] J. A. Bennett, "A study of the damaging effect of fatigue stressing on x4130 steel," *Proceedings, American Society for Testing and Materials*, 1946.

[97] D. L. Henry, "A theory of fatigue damage accumulation in steel," *Transactions of the ASME*, 1955.

[98] R. R. Gatts, "Application of a cumulative damage concept to fatigue," *ASME Journal of Basic Engineering*, 1961.

[99] R. R. Gatts, "Cumulative fatigue damage with random loading," *ASME Journal of Basic Engineering*, 1962.

[100] J. I. Bluhm, "A note on fatigue damage," *Materials Research and Standards*, 1962.

[101] H. T. Corten and T. J. Dolon, "Cumulative fatigue damage," *Proceedings of the International Conference on Fatigue of Metals*, 1956.

[102] A. M. Freudenthal and R. A. Heller, "On stress interaction in fatigue and a cumulative damage rule," *Journal of the Aerospace Sciences*, 1959.

[103] A. M. Freudenthal, "Physical and statistical aspects of cumulative damage," *Colloquium on Fatigue*, 1955.

[104] R. Spitzer and H. T. Corten, "Effect of loading sequence on cumulative fatigue damage of 7071-t6 aluminum alloy," *Proceedings, American Society for Testing and Materials*, 1961.

[105] S. S. Manson, A. J. Nachigall, and J. C. Freche, "A proposed new relation for cumulative fatigue damage in bending," *Proceedings of the American Society for Testing and Materials*, vol. 61, pp. 679 – 703, 1961.

[106] S. S. Manson, A. J. Nachigall, C. R. Ensign, and J. C. Freche, "Further investigation of a relation for cumulative fatigue damage in bending," *ASME Journal of Engineering for Industry*, 1965.

[107] S. S. Manson, "Interfaces between fatigue, creep, and fracture," *International Journal of Fracture Mechanics*, 1966.

[108] S. S. Manson and G. R. Halford, "Practical implementation of the double linear damage rule and damage curve approach for treating cumulative fatigue damage," *International Journal of Fracture*, 1981.

[109] F. R. Shanley, "A theory of fatigue based on unbonding during reversed slip, report p-350," *The Rand Corporation, Santa Monica*, 1952.

[110] S. R. Valluri, "A unified engineering theory of high stress level fatigue," *Aerospace Engineering*, 1961.

[111] S. R. Valluri, "A theory of cumulative damage in fatigue. report no. arl 182," *Aeronautical Research Laboratory, Office of Aerospace Research, United States Air Force*, 1961.

[112] T. D. Scharton and S. H. Crandall, "Fatigue failure under complex stress histories," *ASME Journal of Basic Engineering*, 1966.

[113] T. Bui-Quoc, "An interaction effect consideration in cumulative damage on a mild steel under torsion loading," *Proceedings of the 5th International Conference on Fracture*, 1981.

[114] T. Bui-Quoc, "Cumulative damage with interaction effect due to fatigue under torsion loading," *Experimental Mechanics*, 1982.

[115] A. Biron and T. Bui-Quoc, "Cumulative damage concepts with interaction effect consideration for fatigue or creep; a perspective," in *Transactions of the 6th International Conference on Structural Mechanical Reaction Technology, Paris, France*, 1981.

[116] T. Bui-Quoc, "A simplified model for cumulative fatigue damage with interaction effects," in *Proceedings of the 1982 Joint Conference on Experimental Mechanics, Society for Experimental Stress Analysis, Brookfield Center, CT*, 1982.

[117] J. Newman, "A crack closure model for predicting fatigue crack growth under aircraft spectrum loading. in methods and models for predicting fatigue crack growth under random loading," *ASTM STP 748*, 1981.

[118] A. Vašek and J. Polak, "Low cycle fatigue damage accumulation in armco-iron," *Fatigue and Fracture of Engineering Materials and Structures*, vol. 14, no. 2-3, pp. 193–204, 1991.

[119] Z. Azari, M. Lebienvenu, and G. Pluvinage, "Functions of damage in low-cycle fatigue.," *Advances in Fracture Research (ICF 6)*, 1984.

[120] B. Delaunay, "Sur la sphere vide," *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, vol. 7, no. 793-800, pp. 1–2, 1934.

[121] National Digital Elevation Program (NDEP), *Guidelines for Digital Elevation Data*, 1 ed., 2004.

[122] S. A. Levin, "The problem of pattern and scale in ecology," *Ecology*, vol. 73, no. 6, pp. 1943–1967, 1992.

[123] G. T. Raber, J. R. Jensen, M. E. Hodgson, J. A. Tullis, B. A. Davis, and J. Berglund, "Impact of lidar nominal post-spacing on dem accuracy and flood zone delineation," *Photogrammetric Engineering and Remote Sensing*, vol. 73, no. 7, pp. 793–804, 2007.

[124] P. T. Shih and C.-M. Huang, "Airborne lidar point cloud density indices," in *AGU Fall Meeting Abstracts*, 2006.

[125] R. B. Myerson, *Game Theory.* Harvard University Press, 2013.

[126] E. Rasmusen, *Games and Information: An Introduction to Game Theory.* Blackwell, 2006.

[127] D. M. Kreps, *Game Theory and Economic Modelling*. Oxford University Press, 1990.

[128] R. Aumann and S. Hart, *Handbook of Game Theory with Economic Applications*. North-Holland Amsterdam, 1992.

Appendices

A Rainflow Counting Algorithms

Schematics illustrating the process of performing cycle counting for common Rainflow counting methods are shown below.

A.1 Three Point Rainflow Cycle Counting

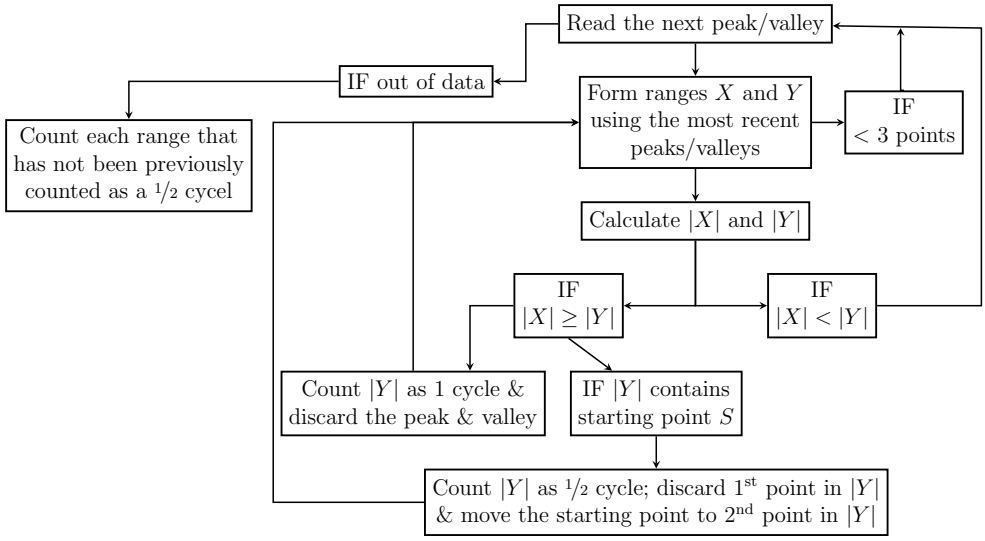


Figure A.1: Three point cycle counting outline

A.2 Four Point Rainflow Cycle Counting

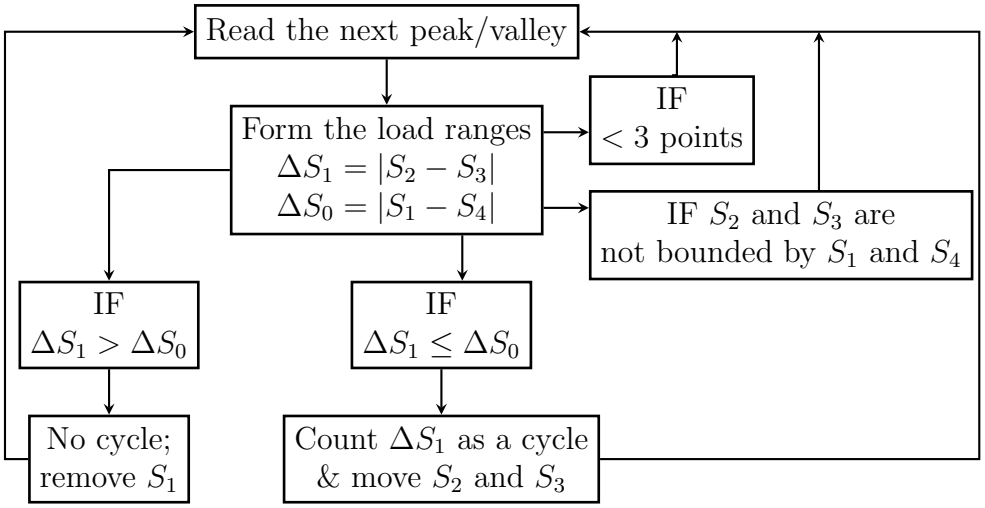


Figure A.2: Four point cycle counting outline

A.3 Rainflow Cycle Counting - Keeping the Load Sequence

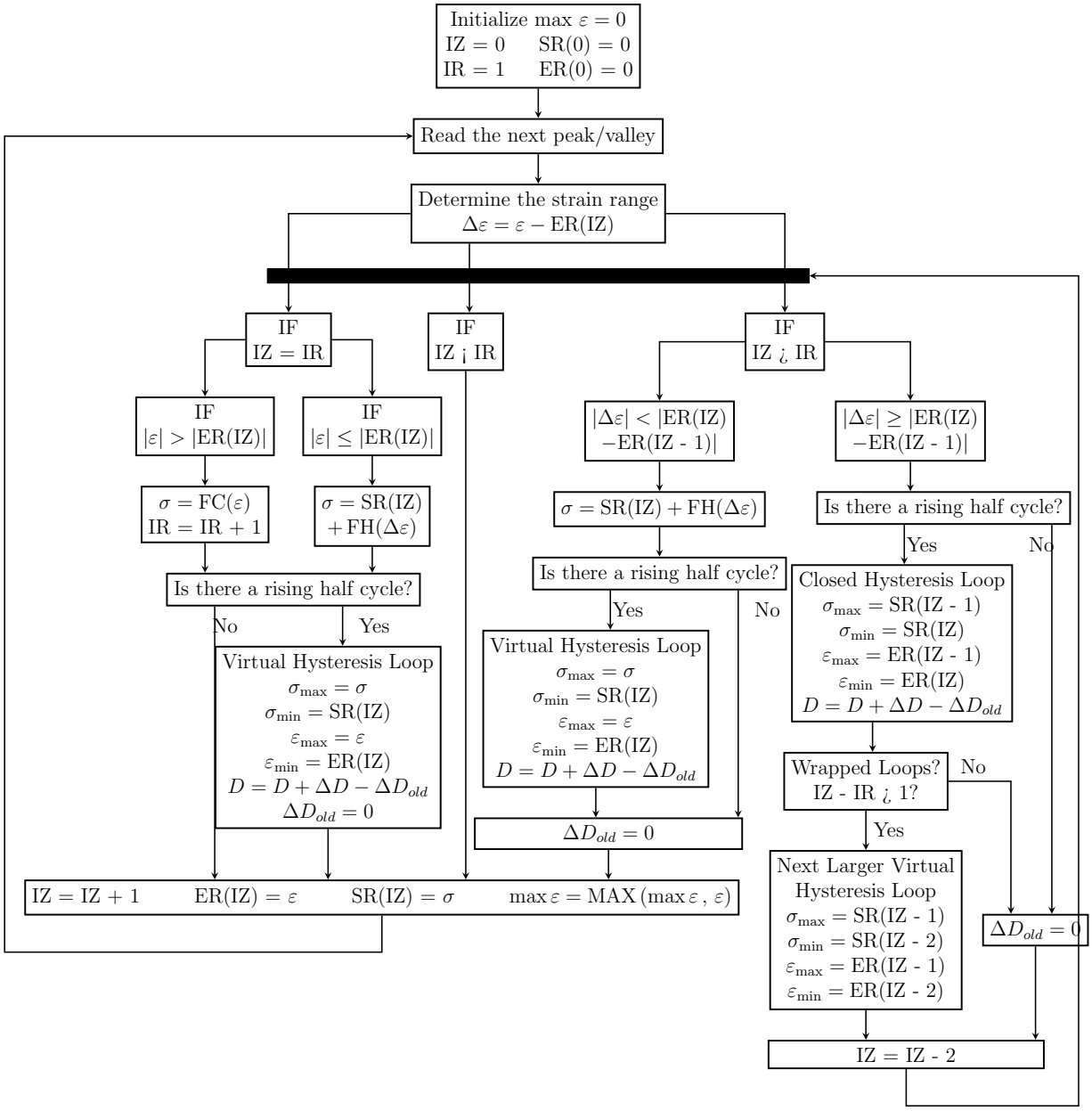


Figure A.3: Rainflow cycle counting designed to keep the load sequence

B Stress/Strain Damage & Life Prediction Theories

This section covers the history and advancement of stress and strain based damage and life prediction methods. For each method covered the positives and negatives associated withe the method are provided. Section B.1 covers early work in damage theories and the development of five types of damage analysis. Section B.2 covers hybrid damage theories, which have been developed to model stress and strain controlled damage analysis. In Section B.3, advanced crack growth concepts are provided. In Section B.4 an overview of some newer theories involving life curve modification are presented. Next, an overview of energy based damage theories is provided in Section B.5 and following an introduction to continuum damage mechanics is provided in Section B.6. Lastly, other damage theories that do not conform to the conventional damage categories is presented in Section B.7.

B.1 Early Work in Damage Theories

In general, models that were developed in the early stages of damage theory, before the 1970s, were based on directly looking at structures experiencing loading while work since the 1970s has been more analytical [36]. In general, early work in damage and life prediction theories can be categorized into one of the following five groups [36]:

1. Endurance limit based approach
2. S-N curve modification approach
3. Two-stage damage approach
4. The damage curve approach (DCA)
5. Crack growth-based approach

The damage curve approach (DCA) and its relationship to Miner's Linear Damage Rule (LDR) is covered in Section 2.5.1. Initial work performed in the remaining four groups is covered in below.

Endurance Limit Reduction Damage Prediction
The idea that an endurance limit is affected by prestress became an important influence on further damage model development. Kommers and Bennett used a two-level step loading scenario to investigate the effect of prestressing on the endurance limit [95, 96]. The results from these investigations suggest that the reduction in the endurance limit could potentially be used as a damage measure, but Kommers and Bennett did not perform any correlation to the life fraction. After some time, in 1955, Henry was the first to form a correlation between the reduction in endurance limit and the life fraction [97]. Later, Gatts and Bluhm also identified a similar correlation in testing they conducted [98, 99, 100].

S-N Curve Modification
When variable amplitude loading is present in a component, load interactions can occur depending on the cycle counting and damaging model selected. Thus, the ability to estimate

the damage in a component is subject to the damage model being able to handle load-interactions. Some early theories to handle load-interaction effects were developed by Corten-Dolon and Freudenthal-Hellar. Both of these methods shared a common method. This method involved using a modification of the S-N diagram. This modification was in the form of a clockwise rotation of the original S-N line about a reference point [36]. In the Corten-Dolon model the rotation point was selected to be the point corresponding to the highest level in the load history [101]. In the Freudenthal-Hellar method the reference point was chosen based off of the stress level which corresponded with the life of 10^3 - 10^4 cycles [102, 103]. Spitzer and Corten tried to improve the Corten-Dolon method by using the average results achieved from performing a two-step block test to determine the slope of the modified S-N line [104]. Manson *et al.* suggested that a point corresponding to a life of 10^2 - 10^3 cycles on the original S-N live could be selected as the reference point for the rotation to be performed around [105, 106]. Manson *et al.* approach was a major improvement because it was able to provide a method for predicting the reduction of the endurance limit due to precycling and it not only accounted for load interaction effects, but also accounted for small cycle damage [36].

Two-Stage Linear Damage Approach

The two-stage linear damage approach was developed in order to fix the issues associated with the LDR approach, which were mentioned above, but still contain the simplicity of the LDR approach. The two stages associated with constant amplitude stress are:

1. Damage due to crack initiation
2. Damage due to crack propagation

In either of these stages the current approach is to apply the LDR to determine the damage. In 1966 Manson proposed the double linear damage rule (DLDR). In Manson's proposed method the two stages were separated by Eq. (B.1), where P is a coefficient which corresponds to the second stage [107].

$$\begin{aligned} N_1 &= N_f - PN_f^{0.6} \\ N_2 &= PN_f^{0.6} \end{aligned} \tag{B.1}$$

The DLDR method presented in Eq. (B.1) can be defined by linearizing the damage curves from the Damage Curve Approach, covered below. When the DLDR method is redefined in this form a knee point separating the two stages is identified by performing a regression analysis on experimental data to define a set of material constants [36].

Damage Curve Approach (DCA)

In order to improve the DLDR method the damage curve approach was developed using the physics of crack growth. Crack growth can be broken down into four different portions: 1) dislocation agglomeration, 2) subcell formation, 3) multiple micro-crack formation, and 4) the independent growth of different cracks that grow to form a single crack that leads to failure. Manson and Halford realized these portions and formulated the 'effective crack growth' model. This model shown in Eq. (B.2) as a function of the normalized crack length, r, accounts for these effects by calculating the instantaneous crack length, a, using the initial

($r = 0$) crack length, a_o; the final ($r = 1$) crack length, a_f; and a exponent q that is a function of the number of cycles and material constants [36, 108].

$$a = a_o + (a_f - a_o)r^q \tag{B.2}$$

Typically, the initial crack length, a_o, is zero allowing for simplification of Eq. (B.2). In addition, damage is equivalent to the ratio of the instantaneous crack length, a, to the final crack length, a_f as shown in Eq. (B.3) power [108]. Based on the assumption of zero initial crack length the damage is equivalent to the crack length raised to the q power.

$$D = \frac{a}{a_f} = r^q \tag{B.3}$$

The DCA can further extended to a Double Damage Curve Approach (DDCA) by adding a linear term into the development of Eq. (B.3) [36].

Crack Growth Approach

A damage theory that based the damage measure off of the crack length was introduced by Shanley in 1952. This theory suggested that the rate of crack growth was either linearly or exponentially related to the applied stress level [109]. In 1961, Valluri presented a model in the form of a differential equation. This differential equation was formed using the concepts of dislocation theory combined with macroscopic elasto-pllastic fracture theory [110, 111]. The differential equation is similar to that developed in Linear Elastic Fracture Mechanics (LEFM) and is provided in Eq. (B.4)

$$\frac{da}{dN} = Cf(\sigma)a \tag{B.4}$$

In Eq. (B.4) a is the crack length, C is a material constant, and $f(\sigma)$ is a function that depends on the material and loading configuration [36]. In 1966, Scharton and Crandall created another crack growth damage theory shown in Eq. (B.5) using a material constant m [112].

$$\frac{da}{dN} = a^{m+1}f(\sigma_{ij}) \tag{B.5}$$

B.2 Hybrid Theory

The hybrid damage theory was first developed for stress-controlled fatigue and later adapted to strain controlled fatigue. The theory for these two models was based off the hybridization of four prior damage models that were developed by Henry, Gatts, Shanley, and Valluri [36]. These four models provided a means for modeling: the crack growth rate using a power rule, a relationship between crack growth and cyclic stress range, and the instantaneous value of the endurance limit. Both stress and strain-controlled theories were able to account for the reduction in the strain endurance limit due to prior strain cycles. In addition the hybrid

theory was able to provide an improved damage prediction over the LDR method, but since it did not account for load interaction effects there was still a deviation from collected data.

In order to account for the load interaction effects Bui-Quoc developed two different approaches. The first approach was the fictitious load approach ([113, 114, 115]) and the second approach is the cycle ratio modification approach ([114, 116]). The factitious load method was developed only for the two-step load cycling [36]. IN the cycle ratio modification approach, a load-interaction parameter was introduced. By introducing this parameter the model is able to be extended beyond two-step load cycling to multi-step load cycling based on altering the value of the load-interaction parameter.

B.3 Theories Involving the 'Crack Growth Concept'

Since crack growth is directly related to damage accumulation, methods involving crack growth have become widely accepted. In the early 1970s numerous macro crack growth models were developed based on the concepts of Linear Elastic Failure Mechanics (LEFM) [36]. One popular macro crack growth method was developed by Wheeler. In Wheeler's model the rate of the crack grow is related to the interaction of multiple crack-tip plastic zones under residual compression stresses due to overloading forces [36]. Since Wheeler's model other macro fatigue crack growth models have been developed, one notable one being a model developed by Newman. Newman's model is notable because it focused on determining the opening stress associated with crack closure, an issue with prior macro crack growth models, using an iterative solution based on finite element programs [117].

In order to model the accumulation of damage between the formation of the initial crack and stage 1 growth, Miller and Zachariah formed an exponential relationship between the crack length and the elapsed life for each of the crack formation phases. This model is commonly referred to as the double exponential law [36]. The algorithm was then reformatted by Ibrahim and Miller. The reformatted algorithm was based on the small crack growth mechanism, where stage 1 crack propagation was expressed in a similar method to the LEFM method that is used for stage 2 crack growth.

Miller *et. al* decided to investigate the behavior of short cracks. From their investigation they found that cracks immediately occur in metal during fatigue loading and throughout the entire fatigue lifetime cracks propagate from some initial defect. Miller *et. al* renamed the initial and stage 1 crack propagation as Microstructurally Short Crack (MSC) growth and Physically Small Crack (PSC) growth, respectively. Models for both of these crack stages can be developed through experimental observation, however further analysis may need to be performed to verify the validity of the method depending on the application [36]. Ma and Laird developed the Ma-Laird model, which is similar to the short crack theory developed by Miller *et. al.* In the Ma-Laird model a the crack population is defined as a linear relationship to the strain amplitude applied to the component and amount of life used. Therefore, the crack population can be used as an indicator for damage. The benefit of the Ma-Laird model

is that it is capable of accounting for the load interaction effects [36].

A last, notable crack growth concept is one developed by Vasek and Polak. After performing exerpiments Vasek and Polak found there to be two regimes that model the crack formation: crack initiation and crack propagation. For the crack initiation regime a constant crack growth rate model was proposed and for the crack propagation regime a linear approximation was used [118].

B.4 Recent Theories Based on Life Curve Modifications

Prior to the 1970s theories based on life curve modifications were simple and relatively effective. It was after the 1970s the models were modified such that they are load-level dependent and can account for load sequence effects. Below, some of the more well known life curve modification methods are covered.

Subramanyan's Knee Point Approach
Based on experimentation and observations Subramanyan created a knee point-based approach. The knee point-based approach was done by forming a set of isodamage lines that would converge to the S-N curve endurance knee point. The damage was then found by finding the ratio of of the isodamage line slope to slope of the original S-N curve [36]. With this method it is assumed that the endurance limit of the material remains constant throughout the loading sequence. In addition, this this method is not valid for stress levels that are near the material fatigue limit because of the singularity associated with the knee point and the nonlinearity that occurs in a log-log plot of the S-N curve nearby the fatigue limit [36].

Hasin-Rotem Model
Hasin and Rotem proposed two forms of convergence to the S-N line. The first approach, is based on the idea that all lines pass through the point where the original S-N line intersections the S axis (usually referred to as the static ultimate) [36]. This method is convenient because by using it a convergence point no longer needs to be found, like Subramanyan's method required. The second approach was to use the endurance limit as the convergence limit [36]. This second method is basically the same as Subramanyan's knee point approach discussed above.

Ben-Amoz's Boundary Theory
Ben-Amoz realized that damage is a statistical phenomenon and because of this results are not always the same. This lead to them creating the concept of using bands on the residual fatigue life instead of using a definite damage rule like most other models utilize [36]. Ben-Amoz's boundary theory is based on the premise that the residual life lines would fall in the upper and lower bounds of the original life line [36]. Ben-Amoz used the Miner LDR and Subramanyan's theory to represent the two bounds as an initial approximation. These bands were then narrowed by including other information from the fatigue damage (i.e. crack initiation and propagation). This method was then further improved by considering all of

the parameters to be functions of random variables N_1, N_2, and N_e. The boundary theory was also modified to predict the creep residual time in two-stage expose to stress (σ_1 and σ_2) at a fixed elevated temperature [36]. The boundary theory was again modified to model creep-fatigue interactions.

Leipholz's Approach
Leipholz agreed with Freudenthal's and Heller's opinion in that the error in the LDR was not due to its linear summation but was instead due to the assumption of damage-rate independence of loading levels [36]. Leiphoz replaced the original S-N curve with a modified S'-N' curve to account for load interaction effects. Experimental results showed that Leipholz's model can accurately predict the fatigue life when repeated block loading occurs [36].

B.5 Energy Based Damage Theories

Inglis found that there was connection between hysteresis energy and fatigue behavior, due to this relationship studies have been carried out based on the energy method [36]. It wasn't until the 1960s that Morrow and Halford developed strain energy based failure criteria, but in the last two decades strain energy based cumulative damage theories were developed. The benefit of using an energy-based damage is that it can unify the damage caused by different loadings (i.e. thermal cycling, creep, and fatigue). In addition, energy-based damage models can have load profiles based on mean stress and multiaxial loads because strain energy-based models have multiaxial fatigue parameters [36]. For a complete review of energy-based damage models see [36].

B.6 Continuum Damage Mechanics Approach

Continuum damage mechanics (CDM) is a relatively new form of damage modeling and is based on the deterioration of the material at a continuum scale [36]. CDM damage modeling was originally based off of the concepts developed by Kachanov and Rabotnov's methods for modeling creep damage. CDM has proven to be successful and has lent to expanding this approach to ductile plastic damage, creep-fatigue interaction, brittle fracture, and fatigue damage.

B.7 Other Damage Theories

While the prior methods of approximating the damage experienced by a component are commonly used and in some cases been improved since they were first conceived, there are still other damage theories that have been recently developed. These other damage models are unique and do not fall under the common damage types presented in Sections B.1 through B.6. Some of these not as commonly used damage theories are outlined below.

Kramer's Surface Layer Stress Model

Surface quality is critical to the fatigue characteristics of a material and a part. Kramer realized the importance of understanding the surface and therefore introduced the concept of surface layer stress in order to model damage. The basis of Kramer's model is that during fatigue cycling the surface layers of the specimen with harden. This hardening can be attributed to higher dislocation density than the interior portion of the specimen. This means that in order to achieve a certain plastic strain must be imposed than if the hardened layers did not exist [36].

Internal and Effective Stress Model

It was discovered that the average dislocation velocity is proportionally related the the effective (resolved) stress at a dislocation. Since the average dislocation velocity is proportional to the effective stress that means the plastic strain rate is also proportionally related to the stress at the dislocation. Ikai *et al* performed experimentation with constant and variable amplitude stress cycling. The conclusion of these experiments is that internal stress is representative of the fatigue resistance of a material. In addition, when the effective stress is above a critical value then fatigue damage occurs in the material. But, when the effective stress is below the critical value it is believed that the material reacts in two manners. The first manner is that it exhibits damage as a consequence of effective stress. The second manner is that the part is strengthened proportionally to the internal stress. One issue with this method is that it cannot be used to analyze crack propagation (i.e. where internal stress measurement is not useful).

Overload Damage Model

Overloads commonly occur on parts undergoing real world loading scenarios. These overloads can have an effect on the damage of the part, therefore Brose *et al* performed some testing. In their testing they used small, smooth specimens and used fully reversed strain-controlled fatigue testing. The specimen was loaded in two different methods: ten completely reversed initial overstrain cycles or one fully reversed periodic overstrain cycle at 10^5 cycle intervals [36]. The damage model presented in Eq. (B.6) was created based on the experiment results.

$$D = \sum D_{OL} + \sum D_{SS} + \sum D_{INT} \tag{B.6}$$

In Eq. (B.6) D_{OL} is the overload damage, D_{SS} is the smaller amplitude cycle damage (i.e. steady state damage in the absence of overloads), and D_{INT} is the interactive damage from the smaller cycles that come after the overload. This damage model is promising because it accounts for the interactive damages when either a tensile or compressive overload is present [36].

Plastic Strain Evolution Model

In an effort to define a relationship between damage and changes in the mechanical properties of a material Azari *et al* created a general damage function. During experimental testing, the strain range, $\Delta\epsilon$, was controlled and the plastic strain range, $\Delta\epsilon_p$, was selected as an evaluation property for the model. Thus, the accumulated damage in the component tested

is defined using the initial plastic strain, $\Delta\epsilon_{po}$; the final plastic strain, $\Delta\epsilon_{pf}$; and the fatigue constant, C.

$$D = \sum_{i=1}^{n} \left(\frac{\Delta\epsilon_p - \Delta\epsilon_{po}}{\Delta\epsilon_{pf} - \Delta\epsilon_{po}} \right)_i^{-1/C} \tag{B.7}$$

For a constant amplitude strain range, a damage evolution curve can be generated by plotting D as a function of the number of cycles. Then, the area under the damage evolution curve can be found. During Azari *et al* experiments it was found that when the area under the curve was divided by the total number of cycles to failure at the applied strain a constant value (approximately 0.55) was achieved [119]. Thus, Azari concluded that fatigue life under complex loading can be modeled using only constant amplitude tests. The damage from complex loading is simply the summation of the area under the damage evolution curve for each load divided by the number of cycles to failure at that load. While this damage model can be used to analyze complex loading regimes caution should be used because it is unable to account for load interaction effects and small cycle damage.

C Damaging Event Characteristics

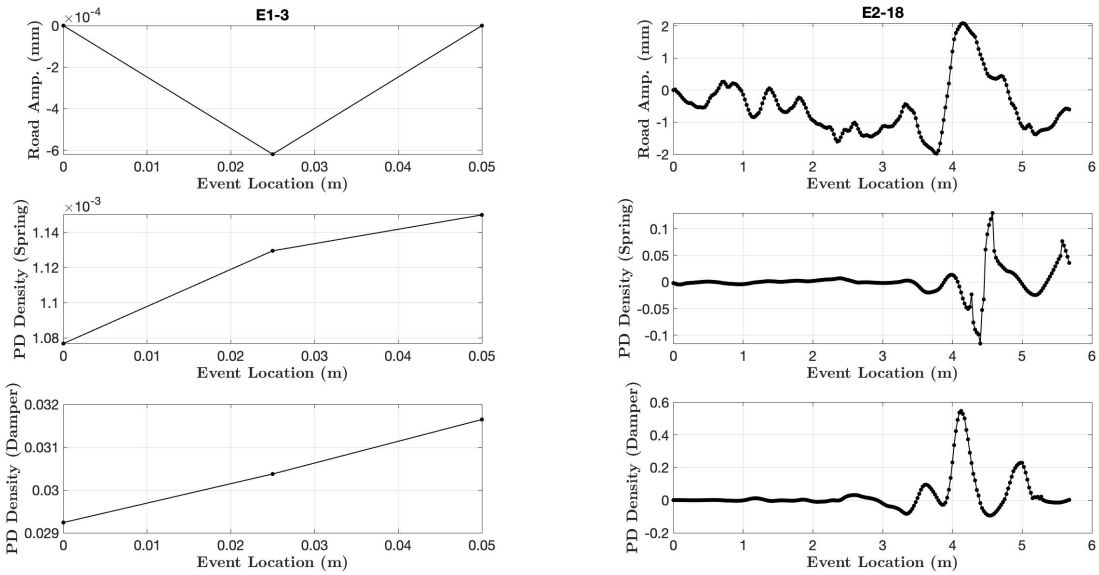
The following figures in this appendix show the twenty-five damaging events that are grouped into thirteen clusters in the second example in Chapter 5, summarized in Table 5.2. Each figure corresponds to one of the thirteen clusters and each of the sub-figures corresponds to an individual damaging event that is a component of the cluster. For each damaging event, a plot of the original road profile amplitudes along with the pseudo damage density in the spring and damper load paths are provided. Each event contains a unique length ranging from approximately 0.1 m to 28 m, therefore a unique number of samples of PD density and road elevations are known for each damaging event. All events were identified from FHWA LTPP road profiles using the identification method presented in Chapter 4 with parameters matching the example presented in Section 4.4. For all events, the disturbances in the road that resulted in the damaging event can be clearly seen. For example, event E2-18 in Cluster 1 (Figure C.1) contains a sharp bump in the road around the 4 m location resulted in significant damage in both the spring and damper load paths; this characteristic is highlighted by the relatively high level of PD density in the load paths between the 3.5 and 5 m locations.

Some damaging events used in the similarity analysis are composed of multiple road disturbances that have an effect on one or more load paths. For example, event E1-1 in Cluster 3 (Figure C.3a) contains a sharp dip in the road between the 4 and 6 m locations of the event resulted in high levels of damage accumulation in both load paths. Then a secondary dip in the road profile around the 10m location resulted in damage in the damper load path, but not the spring load path. These two locations mentioned are in the same event because the pseudo damage density did not sufficiently settle between the peaks for both load paths. Similar causal characteristics can be seen between the road amplitude plot and the PD density in one or both load paths for each event presented in Figure C.1 thru Figure C.13.

In Cluster 3 (Figure C.3), event E1-2 (Figure C.3b) was clustered with event E2-19 (Figure C.3c). Event E1-2 is formed by a bump and a dip in the road, while E2-19 is formed by a road surface with a natural roughness followed by a bump with a similar height to that found in E1-2. In both road profiles the bump results in a set of PD density values in the spring and damper load paths that are similar between the two events resulting in the two events being clustered together. A similar instance of two events having similar road profiles with similar PD density distributions being grouped together is seen in Figure C.13 between E2-1 (see Figure C.13c) and E2-4 (see Figure C.13b). These two events have a similar road profile shape that produced similar PD density distributions for both the spring and damper load paths resulting in the events being paired together. However, in some clusters, events do not have similar profiles. For example, Cluster 8 in Figure C.8 contains two events, E2-12 and E2-14. When looking at the road profile between the two events there does not appear to be similarity between the two events. However, looking at the PD density in the spring and damper load paths between the two events similarities can be seen. Thus, while the

events contained different features and road profile characteristics, they had similar damage characteristics to the vehicle based on the distribution of PD density in both load paths.

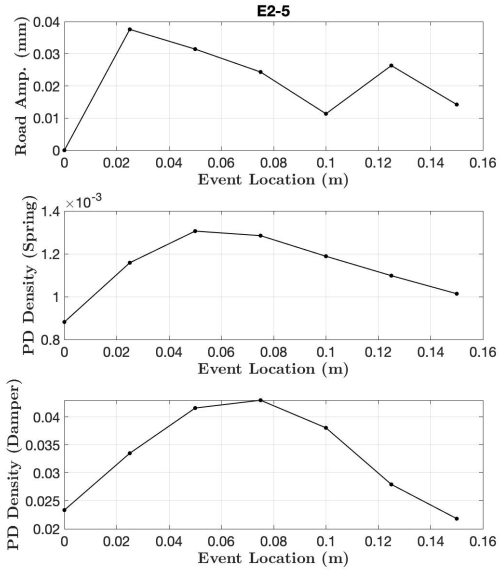
Clusters 2 (Figure C.2), 4 (Figure C.4), 5 (Figure C.5), 11 (Figure C.11), and 12 (Figure C.12) are formed from only a single event because they did not contain a dissimilarity with any other event below the 0.1 threshold selected. In Chapter 5, a higher dissimilarity threshold of 0.3 was used resulting in a reduced number of clusters. Using the higher dissimilarity threshold of 0.3, E2-5 in Cluster 2 (Figure C.2) is determined to be similar to the events in Cluster 1 (Figure C.1). Similarly, Clusters 5 (Figure C.5) and 6 (Figure C.6) are grouped together, and then clusters 9 (Figure C.9) and 10 (Figure C.10) are grouped together. However, even with the higher dissimilarity E2-2 found in Cluster 4 (Figure C.4), E2-20 found in Cluster 11 (Figure C.11), and E2-3 found in Cluster 12(Figure C.12) are still not considered similar to any of the other 25 events. These three events contain PD density distributions that are unique and not similar to any of the events used in this study and are therefore clustered individually.



(a) Set 1 - Event 3.

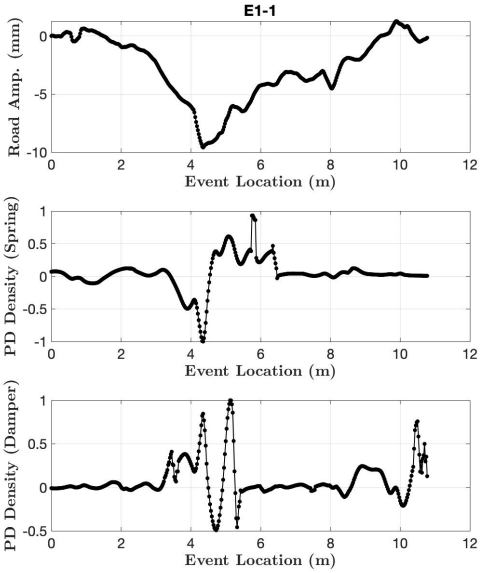
(b) Set 2 - Event 18.

Figure C.1: Damaging events for cluster 1 in Table 5.2.

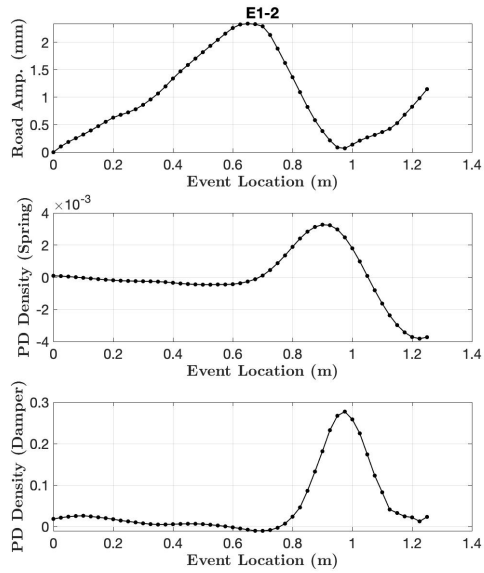


(a) Set 2 - Event 5.

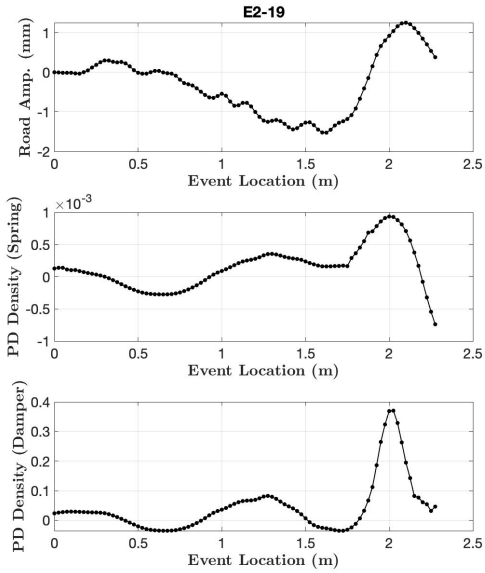
Figure C.2: Damaging events for cluster 2 in Table 5.2.



(a) Set 1 - Event 1.

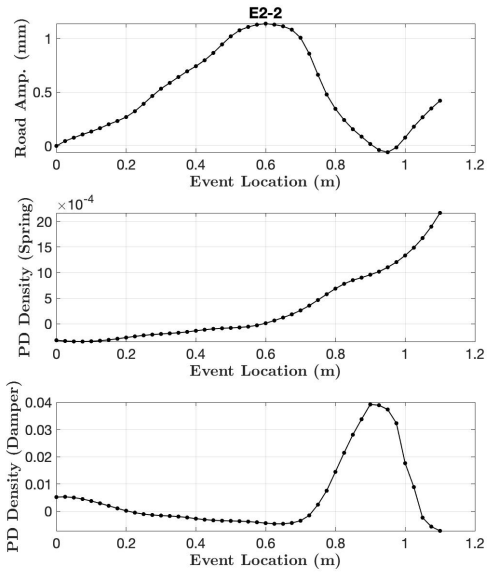


(b) Set 1 - Event 2.



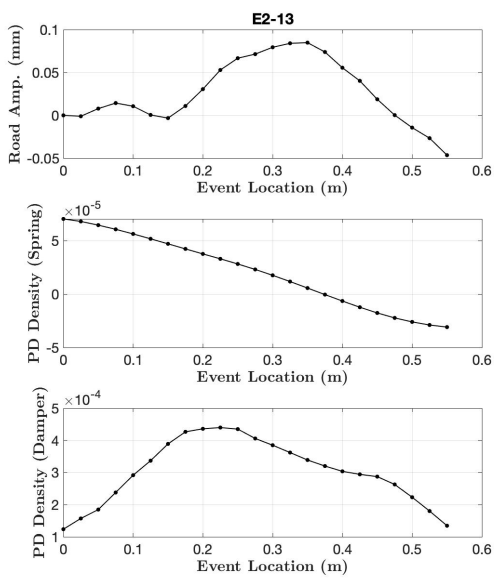
(c) Set 2 - Event 19.

Figure C.3: Damaging events for cluster 3 in Table 5.2.



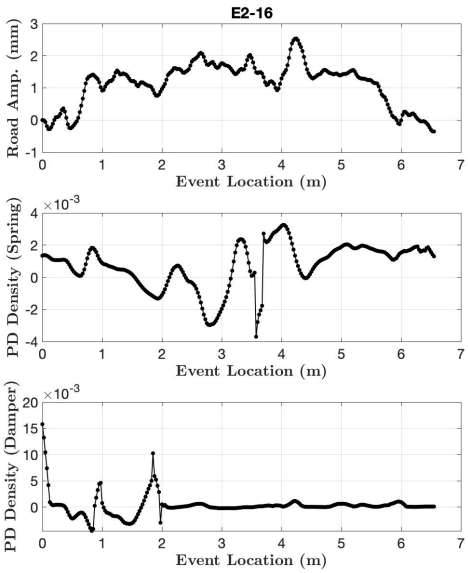
(a) Set 2 - Event 2.

Figure C.4: Damaging events for cluster 4 in Table 5.2.

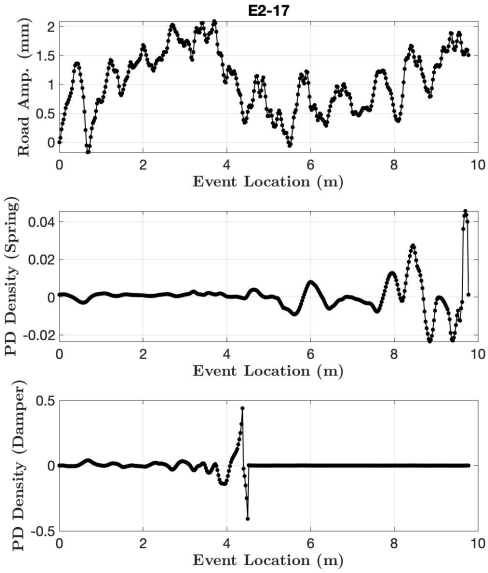


(a) Set 2 - Event 13.

Figure C.5: Damaging events for cluster 5 in Table 5.2.

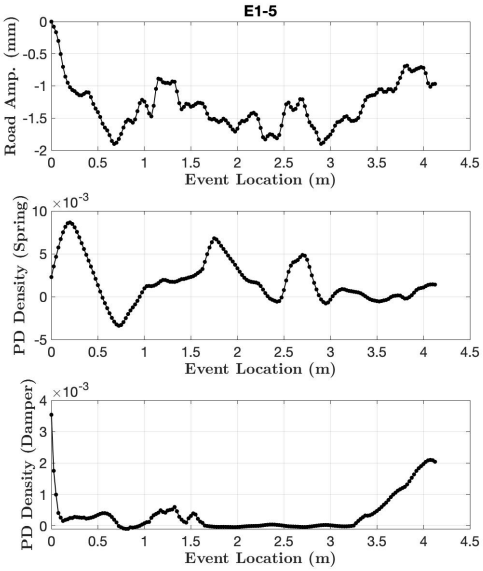


(a) Set 2 - Event 16.

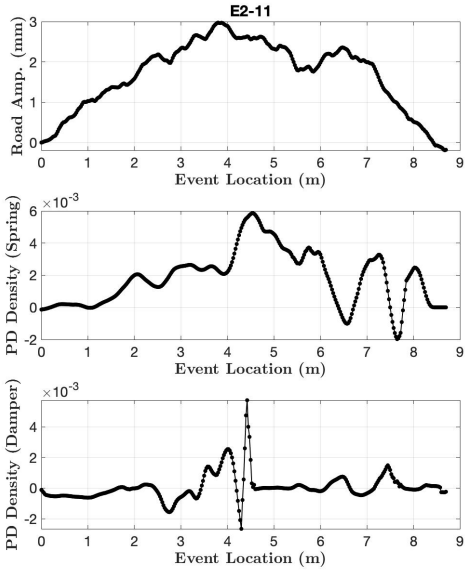


(b) Set 2 - Event 17.

Figure C.6: Damaging events for cluster 6 in Table 5.2.

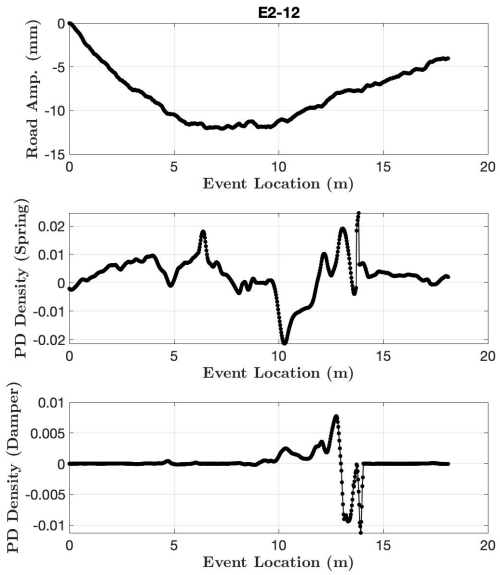


(a) Set 1 - Event 5.

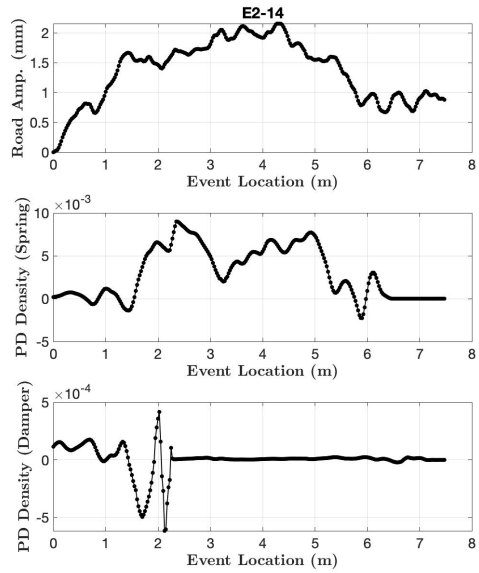


(b) Set 2 - Event 11.

Figure C.7: Damaging events for cluster 7 in Table 5.2.

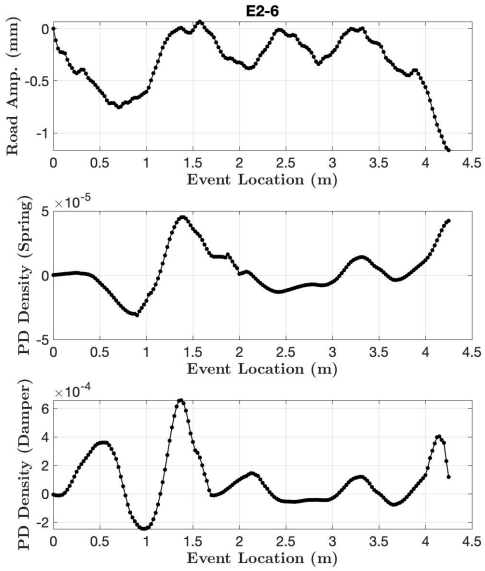


(a) Set 2 - Event 12.

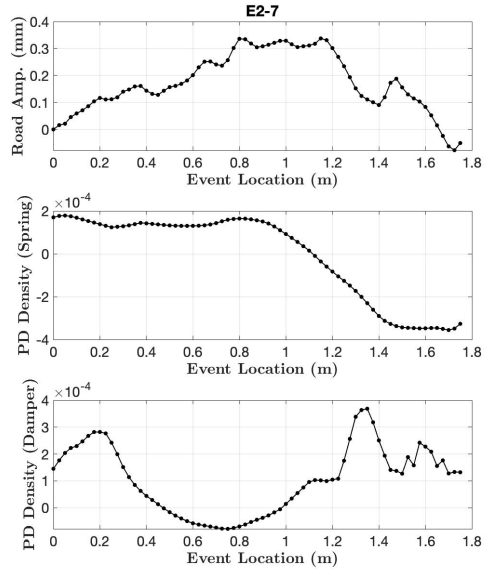


(b) Set 2 - Event 14.

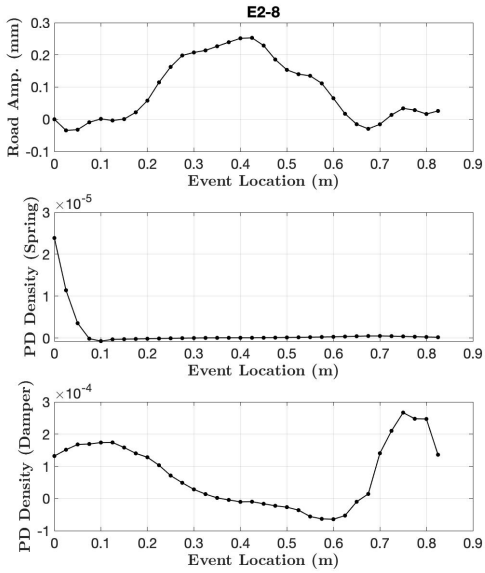
Figure C.8: Damaging events for cluster 8 in Table 5.2.



(a) Set 2 - Event 6.

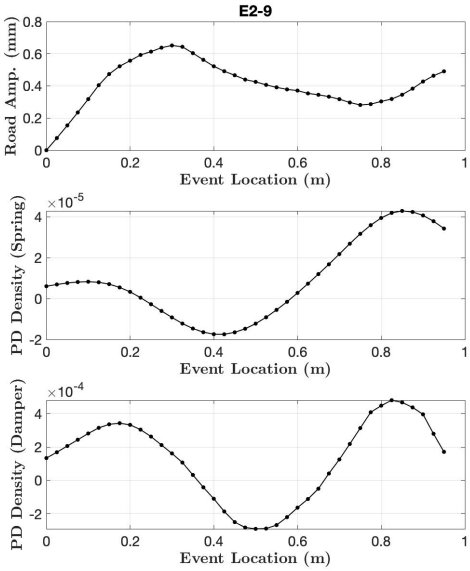


(b) Set 2 - Event 7.

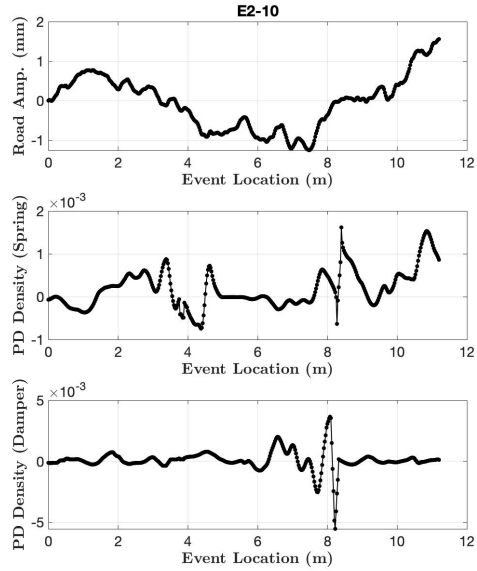


(c) Set 2 - Event 8.

Figure C.9: Damaging events for cluster 9 in Table 5.2.

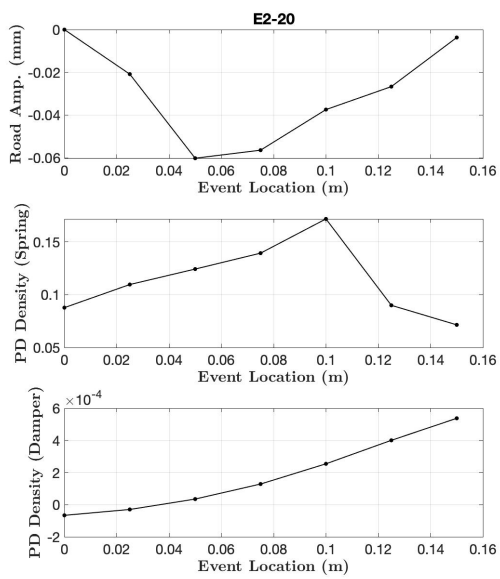


(a) Set 2 - Event 9.



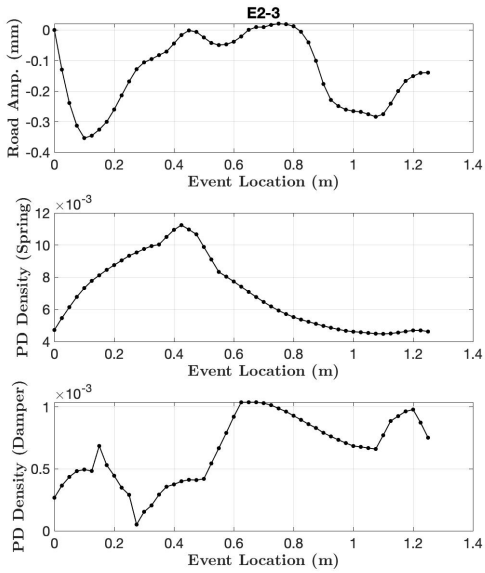
(b) Set 2 - Event 10.

Figure C.10: Damaging events for cluster 10 in Table 5.2.



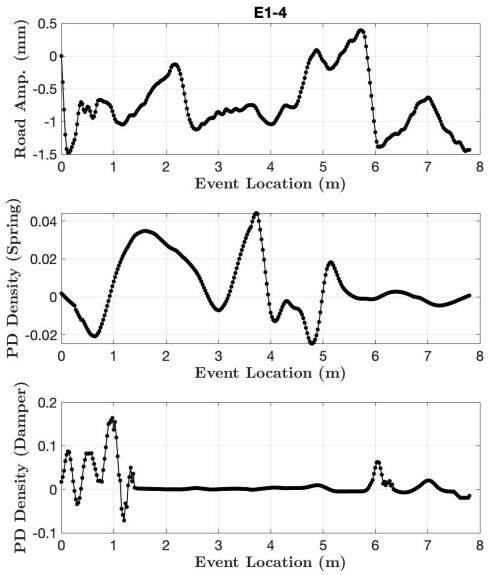
(a) Set 2 - Event 20.

Figure C.11: Damaging events for cluster 11 in Table 5.2.

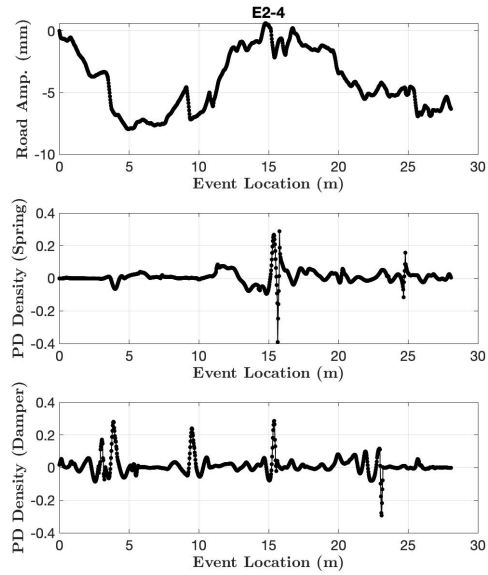


(a) Set 2 - Event 3.

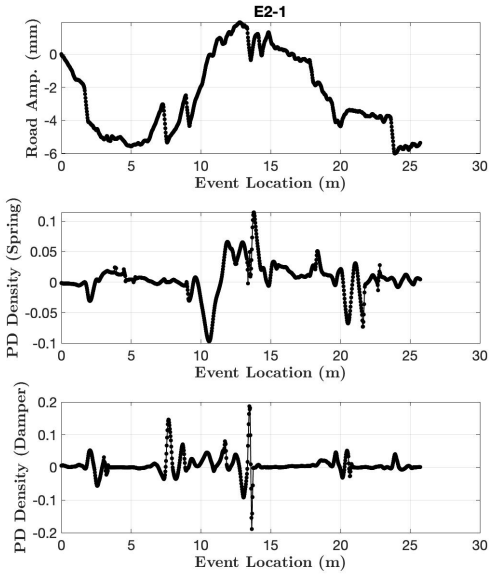
Figure C.12: Damaging events for cluster 12 in Table 5.2.



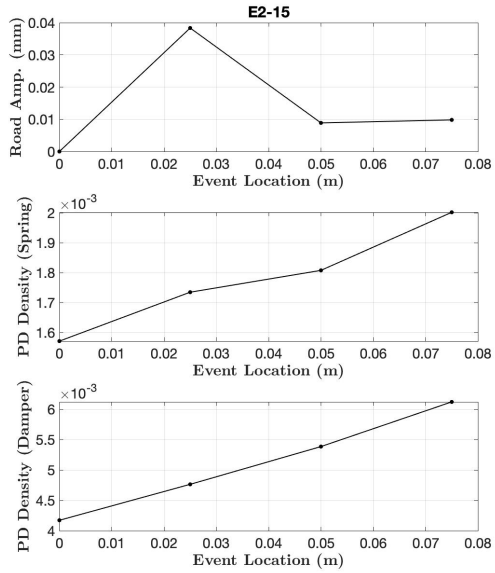
(a) Set 1 - Event 4.



(b) Set 2 - Event 4.



(c) Set 2 - Event 1.



(d) Set 2 - Event 15.

Figure C.13: Damaging events for cluster 13 in Table 5.2.

D Defining Road Surface Measurement Spacing Through Delaunay Triangulation

While performing my doctoral research presented in the main chapters of this dissertation I had the privilege of working on a Federal Highway Administration project focused on the assessment of systems that collect transverse measurements of road surfaces. Transverse profiles consists of a set of road surface elevations that are perpendicular to the direction of travel along a road profile. Figure D.1 illustrates the conventional axes for defining transverse profiles. Figure D.1 clearly illustrates that the transverse direction is constantly changing based on the direction of travel.

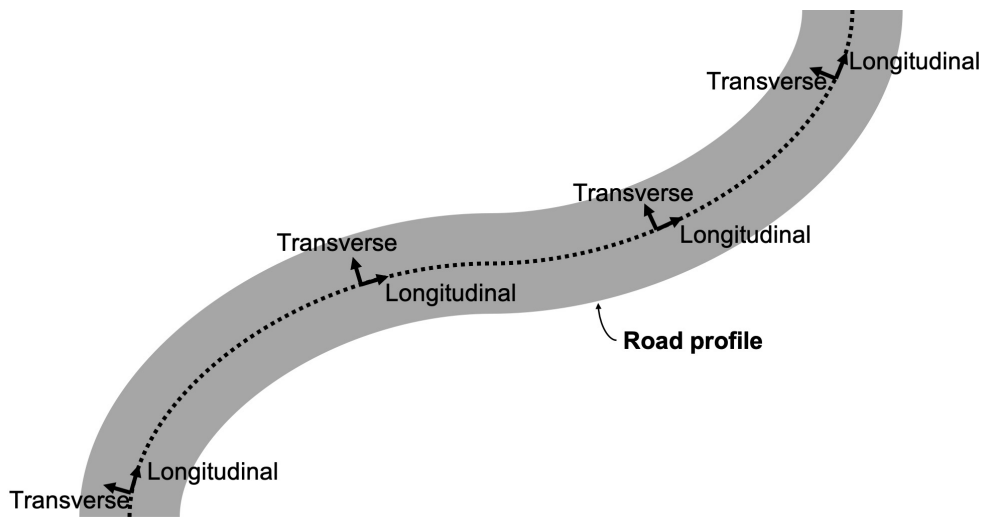
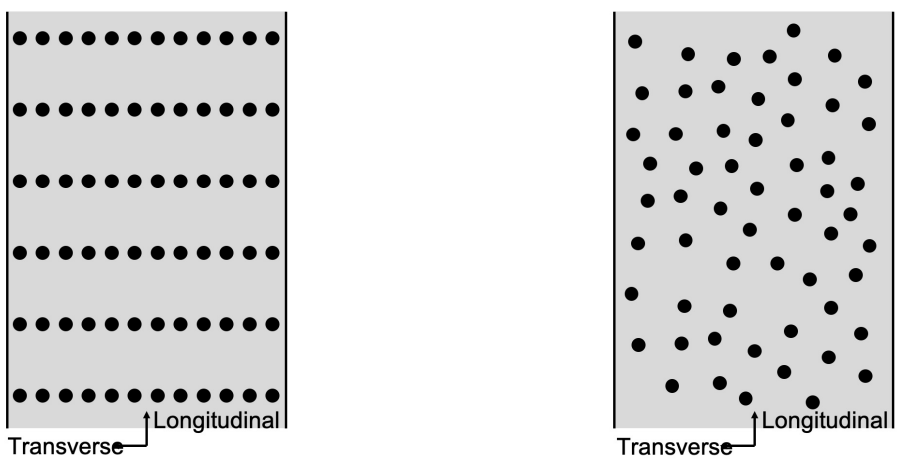


Figure D.1: Generalized road surface illustrating that longitudinal road profiles follow the direction of travel along a road surface and transverse profiles are perpendicular.

While working on the project one of the assessment that needed to happen was identifying the transverse and longitudinal spacing a transverse pavement profiler is capable of producing at highway speeds. While working on this project I made the realization that while discrete measurements from a transverse profiler are discrete, there is some connection between these measurements and a method called Delaunay triangulation can be used to generate a set of proposed connections between the discrete measurements. Unfortunately, due to time constraints of the project the idea of using Delaunay Triangulation to solve for transverse and longitudinal measurement spacing of a transverse pavement profilers was never completed. However, this appendix is included to serve as a summary of my current work in this area to allow for future development, and hopefully a refined process for using Delaunay triangulation to estimate the measurement spacing.

D.1 Introduction

Typical measurement systems that are used to collect transverse pavement profiles are comprised of complex electro-mechanical systems that collect elevation heights at regular sampling rates. These systems use a variety of technologies to measure the road surface elevations and to account for the movement of the system used to collect the measurements (*e.g.*, roll, pitch, and yaw of the host vehicle). However, in the end these systems are able to represent a road surface has a three-dimensional point cloud as illustrated in Figure D.2. In Figure D.2a a regularly spaced point cloud is provided. A regularly spaced point cloud is typically achieved through additional post-processing of the measurements to achieved a desired uniform spacing throughout the reported measurements. In Figure D.2b a set of irregularly spaced point cloud measurements is provided. These irregularly spaced measurements do not always have a uniform appearance and at a work case can appear similar to random data points.



(a) Regularly spaced point cloud measurements of a road surface.

(b) Irregularly spaced point cloud measurements of a road surface.

Figure D.2: Two-dimensional representation of a point cloud measurements of a road surface.

Independent of the point cloud type, a set of connections can be made between neighboring points. One method for defining these connections is Delaunay triangulation. Delaunay triangulation allows for a set of triangles to be formed from a set of points such that no points lie inside a circum-hypershpere of any triangle/tetrahedron in the set [120]. Thus, all edges of triangulations formed between points in a point cloud are unique and do not cross over other points in the point cloud.

For collected ground surface measurements it is necessary to identify the spacing between

measurements constituting the three-dimensional point cloud. For road surface measurements it is necessary to analyze the measurements in the horizontal plane in both the transverse (direction perpendicular to the centerline) and longitudinal (in line with the centerline) directions, therefore a single distance measurement will not provide adequate information regarding spacing of the data in each of the principle components. In this work a novel method of establishing the anisotropic (transverse and longitudinal) spacing for a three-dimensional point cloud of a road surface is established. The developed method is tested to ensure data at an angle, and data with random noise does not affect the results.

D.2 Background

When three-dimensional measurements are collected of a surface it is often necessary to identify the resolution of the measurements to understand the smallest features which can be accurately extracted. To identify the resolution of a point cloud, two nominal measurements are defined: point spacing and point density. Point spacing is the shortest distance between two neighboring points in a point cloud [121]. Point density is related to point spacing, in that the higher the point density the lower the spacing between points and the lower the point density the greater the spacing between points. When either measurement characterization is selected, a nominal measurement can be established under the assumption that the point cloud is isotropic.

The process for identifying the spacing between measurement points in a point cloud is captured through the Nominal Point Spacing (NPS) of the point cloud. The NPS is a representative distance between points in a point cloud and is used to define the resolution of the point cloud. The NPS can be calculated through a variety of methods. Three common methods are: 1) box counting, 2) Voronoi diagram, and 3) Delaunay triangulation. These three methods are illustrated for an example region of a point cloud in Figure D.3. The first two methods calculate the the NPS using an estimated density of the point cloud and the third approach calculates the NPS using the edge length of the Delaunay triangles.

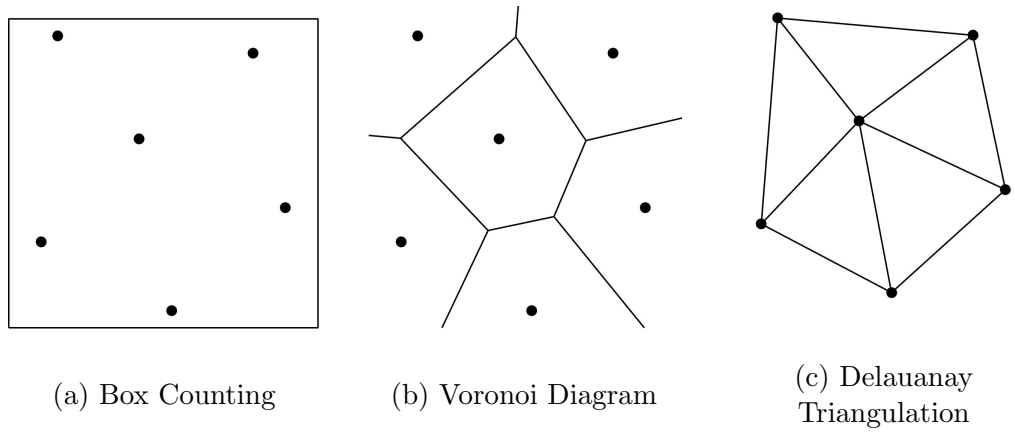


Figure D.3: Three typical methods for calculating the Nominal Point Spacing (NPS) of a point cloud.

The box counting approach uses a specified area (typically a square), illustrated by the black square in Figure D.3a, and identifies the number of points that lie inside this area. Then, the measurement density is simply the number of points divided by the area of the square. The issue with the box counting method is deciding on the shape and size of the polygon along with the location within the point cloud [122, 123]. To address these concerns, Shih and Huang proposed a Triangular Irregular Networks (TIN) based approach to identify the point density for each point in a point cloud by generating Voronoi diagrams based off of a Delaunay triangulation [124]. The objective of a Voronoi diagram is to partition a surface into multiple polygons. In this application, at the center of each polygon is a single measurement point from the point cloud. The edges that define each polygon are such that the area within each polygon is defined to be closest to center point. When a Voronoi diagram is determined for a point cloud, a measurement density is established for each point individually as the inverse of the Voronoi polygon area.

However, the issue with the average window and Voronoi diagram approach is that an average density is calculated to estimate the spacing distance between points. To address the issue of using an average density, the edges that connect neighboring measurement points from a Delaunay triangulation can be used to establish the point spacing. Based on the edges defined through the Delaunay triangulation process, the NPS is calculated as the average of all Delaunay edge lengths. The issue with the NPS is that it provides only a single representative measure of the spacing between measurement points. Therefore, when an anisotropic point cloud is provided the NPS does not accurately capture the measurement spacing in the two directions of interest.

D.3 Delaunay Triangulation of a Road Surface

An important characteristic of road surface measurements is that the measurements are collected of a single surface (a road surface). This characteristic is useful because it means that when point cloud measurements from a road surface are viewed in the transverse-longitudinal orientation as shown in Figure D.2 a two-dimensional Delaunay triangulation can be formed using only the transverse and longitudinal measurements associated with each point cloud. To illustrate this an example point cloud for a regularly gridded road profile are provided in Figure D.4. In Figure D.4 both 10 transverse profiles and 10 longitudinal profiles are present. The transverse spacing is defined in general terms by the variable t and the longitudinal spacing is defined by the variable L.

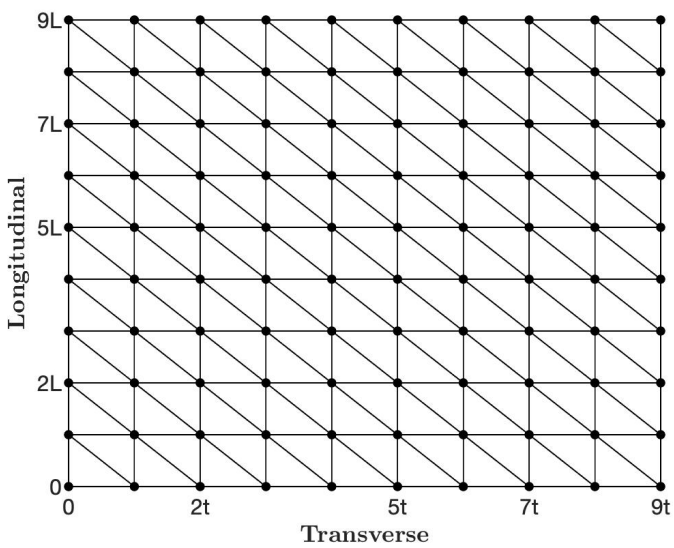


Figure D.4: Delaunay triangulation for a regularly gridded road surface with .

In Figure D.4 each edge of a Delaunay triangle (every black line connection each neighboring dot) can be used to determine an estimate of the measurement spacing in the transverse and longitudinal directions. These estimates can be achieved by decomposing the edge length onto transverse and longitudinal components as illustrated in Figure D.5. In Figure D.5 two edges are shown E_1 and E_2, both edges are not in line with a principal axis and therefore have relative contributions in both the transverse and longitudinal directions, that are illustrated by the dashed lines. It is hypothesized that when all Delaunay edges are decomposed into transverse and longitudinal components the resulting set of transverse distances for all edges in the point cloud can be used to estimate the transverse measurement spacing the resulting set longitudinal distances for all edges in the point cloud can be used to estimate the longitudinal measurement spacing.

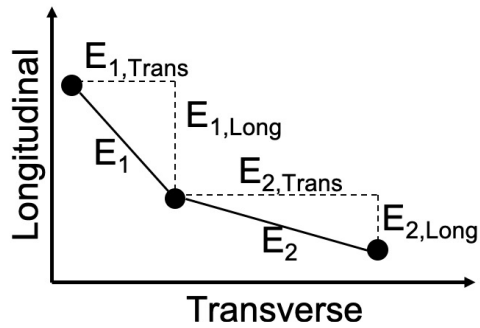


Figure D.5: Projection of Delaunay edges in the transverse and longitudinal directions.

D.4 Distribution of Delaunay Edges

When a Delaunay triangulation is formed each point (or node) in the point cloud contains a set of edges that connect it with neighboring points. Figure D.6 shows an example of a typical node and the expected six edges that connect the node with neighboring points. In general the edges for a node can be placed into one of three categories: longitudinal, transverse, or angled. The longitudinal category is for edges which only contain a distance in the longitudinal direction, similarly transverse edges contain only a distance in the transverse direction. Lastly angled edges contain components of both transverse and longitudinal directions.

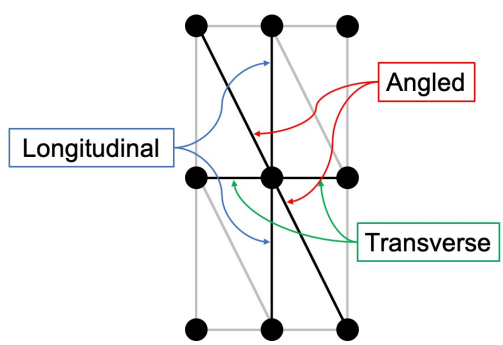


Figure D.6: Edge types for a conventional regularly spaced Delaunay triangulation.

Depending on 'randomness' of the point cloud it can be difficult to directly categorize an edge into one of these categories initially. However, this leads to question in this work that has yet to be answered, should an edge be allowed to contribution to determining the measurement spacing in only one direction or should it contribute to both transverse and longitudinal directions? Two brief investigations of identifying measurement spacing for two types of point clouds is presented below.

Regular Spacing

For the regularly spaced point cloud provided in Figure D.4, the transverse and longitudinal distances associated with each Delaunay triangulation edge was calculated and the results are summarized by the cumulative distributions provided in Figure D.7. The cumulative distribution for transverse spacing is formed from all Delaunay edge transverse distances and the cumulative distribution for the longitudinal spacing is formed from all Delaunay edge longitudinal directions. In both cumulative distributions a crossover from 0 spacing to t/L occurs at 33.3%. This percentage matches up with the 2 out of 3 edges being in a direction that results in one of the principal components of the edge being 0.

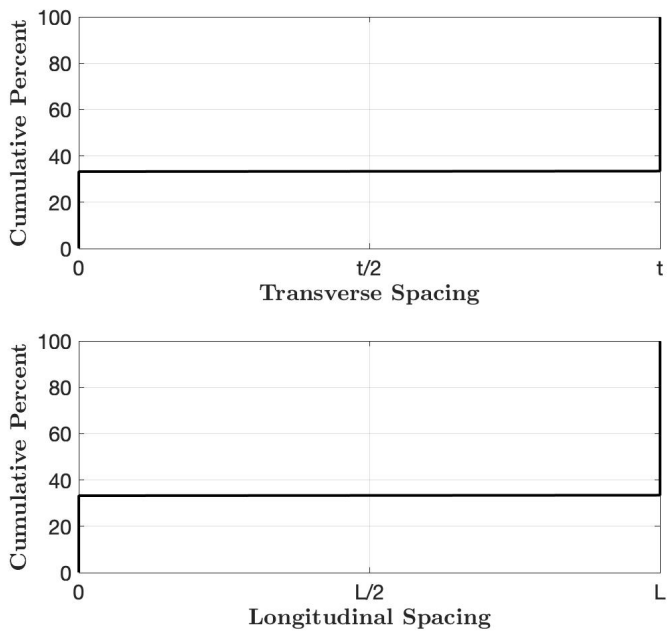


Figure D.7: Cumulative distributions for Delaunay edge lengths in the transverse and longitudinal directions for a regularly spaced point cloud.

For the regularly spaced point cloud, if the cumulative distributions in Figure D.7 were used to define the measurement spacing a criteria of 33.3% or higher would need to be used to ensure the correct spacing values is selected.

Hexagon Closest Packed

The second point cloud type tests was a hexagon closest packed point cloud as shown in Figure D.8. In this structure, a single point is bounded by six other points such that each

of the 6 points lies a distance r away fro the center point. Two critical measurements are indicated in Figure D.8. The first is a transverse distance of $d/2$ and the second is a longitudinal distance of $r/2$. For a hexagon closest packed structure r and d are related as shown in Eq. (D.1).

$$r = \frac{d}{\sqrt{3}} \tag{D.1}$$

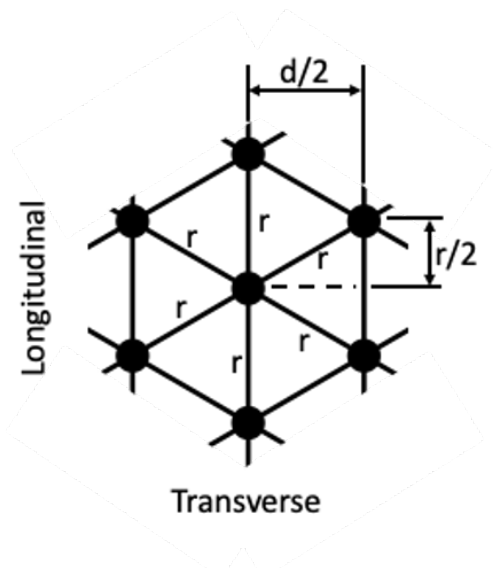


Figure D.8: Delaunay triangulation for a hexagonal closest packed structure.

When the structure provided in Figure D.8 is repeated multiple times for a road profile the transverse and longitudinal components of each Delaunay triangulation edge can be calculated and represented in a cumulative distribution as shown in Figure D.9. For the transverse measurement spacing it can be seen that at 33% a cross over from 0 to $d/2$ transverse spacing is present. Thus, if any percentile greater than 33% is selected a transverse spacing of $d/2$ will be identified. For the longitudinal measurement spacing it can be seen that at 66% a cross overs from $r/2$ to r is seen. Thus, if a percentile lower than 66% is selected a longitudinal measurement spacing of $r/2$ will be identified, however if a percentile greater than 66% is selected a longitudinal measurement spacing of r will be identified.

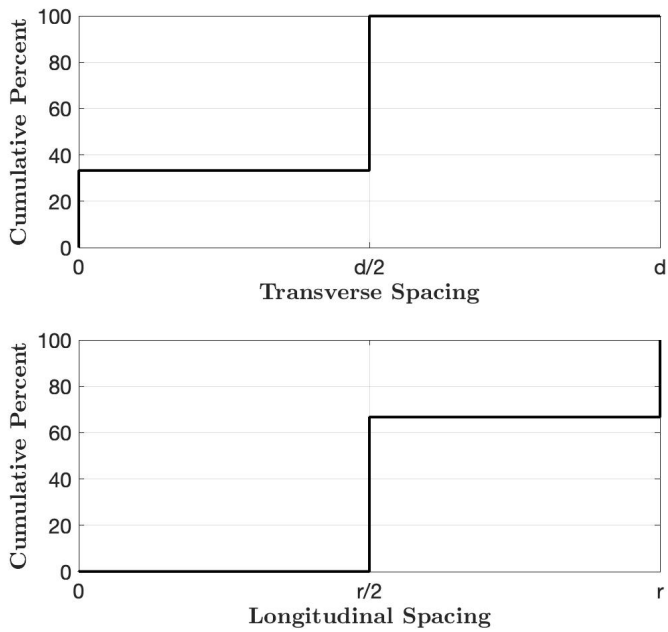


Figure D.9: Cumulative distributions for Delaunay edge lengths in the transverse and longitudinal directions for a hexagonal closet packed point cloud.

E Connection to Game Theory

Game theory is used to develop mathematical models of interactions among rational decision maker [125]. This framework has been implemented in a variety of fields where the objective is to model a logical decision making process. The term 'game' is an umbrella term used to define any interaction between multiple individuals where the outcome for each individual is dependent on the decisions of others. To define a game the following three pieces of information are needed [126, 127, 128]:

1. Players - the individuals participating in the game
2. Information and actions - the knowledge and operations each player can perform throughout the decision making process
3. Payoffs - a numerical value associated with each outcome

There are two common divisions of game theory: cooperative and non-cooperative (sometimes referred to as competitive). In a non-cooperative games each player is in a competition with every other player, where as in cooperative games coalitions of players work together to achieve a collective payoff. In this work presented in this dissertation it is desirable to identify the affect each individual road excitation has on the total pseudo damage in a load path, therefore the connection with the developed work and cooperative game theory is presented. In cooperative game theory, when the objective is to determine what action you (Player i) should make in the game and you know the action of every other player, then it is easy for you to select the appropriate action. In mathematical notation, the action of every other player (except for yourself, location i) can be represented by the set, $\mathbf{a}_{-i}$, provided in Eq. (E.1).

$$\mathbf{a}_{-i} = \left\{ a_1 \quad \dots \quad a_{i-1} \quad a_{i+1} \quad \dots \quad a_N \right\} \tag{E.1}$$

The complete set of all actions in the game is defined as $\mathbf{a}$ which is the union of $\mathbf{a}_{-i}$ and your action, a_i, as presented in Eq. (E.2)

$$\mathbf{a} = \mathbf{a}_{-i} \cup a_i \tag{E.2}$$

In reality Player i does not know the response of all other $N-1$ players, however game theory provides a method for identifying the best response of all players to inform the action of Player i. The best response of all players is defined to be the Shapley Value which uses the gains/costs associated with the actions of players to determine the individual contributions of each player. This individual contribution is determined by removing each player from the game and identifying the marginal contribution of the player (the relative gain/lost the player provides to the coalition).

In Chapter 3 of the dissertation the margin contribution of each road profile excitation to the pseudo damage in a load profile is identified by removing the road excitation and identifying the relative affect on the pseudo damage density. This process is analogous to the removal of a player from a game to identify the marginal contribution of the player and consequently the Shapley value for all players. In Chapter 5 damaging events are grouped by finding the

individual event effect to a cluster of damaging events. This process can be thought of in game theory notation by placing an event (player) into the cluster (coalition) in which it provides the best marginal gain.