

NORTHWESTERN UNIVERSITY

Analysis of Bridge Performance
Using Structural Health Monitoring and Weigh-in-Motion Data

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Civil and Environmental Engineering

By

Fiorella Mete

EVANSTON, ILLINOIS

September 2018

ABSTRACT

Understanding bridge performance and detecting structurally deficient components are of increasing concern to modern infrastructure owners and managers. Concerns over possible effects of increased truck weights led to the opportunity to monitor a highway bridge, regularly subjected to heavy traffic due to logging activities typical of its locale. A weigh-in-motion and a structural health monitoring system were installed to collect long-term data on both traffic and bridge response. This thesis aims to tackle different aspects of the analysis of the data generated by these systems, to promote a gauge of bridge performance. After assessing the conditions of the bridge at the moment of the data collection, the first goal is to develop new metrics for bridge response based on strain, the second goal is to model and predict bridge response based on traffic and finally, the third goal is to classify the most common bridge responses and categorize the trucks which cause them. By achieving these tasks, this thesis aims to provide data-driven decision-making support for policymakers, bridge managers, and owners, to assure timely and effective interventions, to improve long-term durability and extended serviceability, to increase safety, and eventually to save money.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor Prof. Corr for this incredible opportunity, and for his precious support during these years. I always felt lucky to be able to have meetings in the comfortable and spontaneous environment he was always able to create.

I would like to thank the rest of my Ph.D. committee Prof. Chen, Prof. Durango-Cohen, Dr. Kosnik and Prof. Stathopoulos. They all provided invaluable guidance in different but complementary aspects of my Ph.D. career. They positively impacted with insightful comments and suggestions (from the very beginning to the end) both my research and my curriculum here at Northwestern.

I would like to thank my officemates for the joys and frustrations we inevitably shared during these years. Many thanks also to the helpful and friendly staff of our Department and all the professors for whom I had the pleasure to serve as TA.

I would like to thank my late father-in-law, who, on Christmas day, lent me his civil engineering manuals to let me compute the influence line, because if you have something to do and you do it with passion, it does not matter what day it is.

I would like to thank my parents and sister, always my biggest supporters, and my grandparents for always asking me the hardest question of all “what it is what you actually do?”.

I would like to dedicate this Ph.D. thesis to my husband. He always believed in me more than I ever done myself. I learned a lot by silently (kind of) observing him and following his example. Countless times I re-turned my computer and accomplished my unfinished task after looking at him still working after a long day. His encouragement and love got me here and I will never thank him enough for this.

LIST OF ABBREVIATIONS

3DOBS	3D Optical Bridge-evaluation System
AASHTO	American Association of State Highway and Transportation Officials
ACAT	Advanced Condition Assessment Technology
AISI	American Iron and Steel Institute
ANN	Artificial Neural Network
ARTBA	American Road and Transportation Builders Association
ASCE	American Society of Civil Engineers ³
AX#	Axle Weight of axle #
BVRCS	Bridge Viewer Remote Camera System
CAR	Center for Automotive Research
CV	Cross Validation
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DIC	Digital Image Correlation
ESA	European Space Agency
EO	Earth Observation
FHWA	Federal Highway Administration
GDF	Girder Distribution Factor
GNSS	Global Navigation Satellite System
GVW	Gross Vehicular Weight
ICAN	Inductive Classifying Artificial Network
KmL	K-Means Longitudinal

IL	Influence Line
IOA	Influence Observations Analysis
InSAR	Interferometric SAR
LFD	Load Factor Design
LRFD	Load and Resistance Factor Design
MDA	Moment Distance Area
MDOT	Michigan Department of Transportation
M-LiDAR	Mobile Light Detection and Ranging
MTTI	Michigan Tech Transportation Institute
MOT	Maintenance of Traffic
MRL	Multilinear Regression
MSE	Error Mean of Squares
MST	Total Mean of Squares
NBIS	National Bridge Inspection Standards
NCHRP	National Cooperative Highway Research Program
NU-ITI	Northwestern University Infrastructure Technology Institute
R2SHM	RADARSAT-2 Structural Health Monitoring
RMSE	Root Mean Square Error
RT	Regression Tree
SDA	Strain Distance Area
SDOT	South Carolina Department of Transportation
SHM	Structural Health Monitoring
SSE	Sum of Squared errors

SSR	Regression Sum of Squares
SST	Total Sum of Squares
SXY	Spacing between Axle X and Axle Y
ThIR	Thermal Infrared Imagery
UWBIRS	Ultra-Wide Band Imaging RADAR System
VIF	Variance Inflation Factor
W IO	With Influential Observations
W/IO	Without Influential Observations
WIM	Weigh-in-Motion
WisDOT	Wisconsin Department of Transportation
ZTP	Zero-to-Peak

TABLE OF CONTENTS

1	BRIDGE MAINTENANCE: A NON-SUSTAINABLE TREND TO	
	CURB 21
	1.1 CURRENT SITUATION 22
	1.2 IMPORTANCE OF MONITORING 24
	1.3 VISUAL INSPECTION AND SHM: A COMPARATIVE ANALYSIS	
	OF COSTS AND BENEFITS 26
	1.3.1 Case Study 1 28
	1.3.2 Case Study 2 30
	1.3.3 The South Carolina DOT Saving Strategy 33
	1.3.4 Summary 34
	1.4 SUBJECT BRIDGE DESCRIPTION AND DATA ACQUISITION	
	SYSTEMS 34
	1.4.1 WIM System 38
	1.4.2 SHM System 42
	1.4.3 Dataset 47
	1.5 FUTURE OF BRIDGE MONITORING 50
	1.5.1 Monitoring Bridge from Space: the R2SHM Project 50
	1.5.2 Monitoring Bridge from Space: the GeoSHM Project 54
2	MONITORING BRIDGE PERFORMANCE 57
	2.1 INTRODUCTION 58

2.1.1	Literature Review	58
2.2	PRELIMINARY ANALYSIS	60
2.3	GIRDER DISTRIBUTION FACTOR (GDF)	67
2.3.1	AASHTO Design GDF	67
2.3.2	Experimental GDF	68
2.4	RESULTS	70
2.4.1	AASHTO Values in Other Studies	74
2.4.2	Influence of Truck's Characteristics	75
2.5	LONG-TERM AGGREGATE ANALYSES	79
2.6	CONCLUSIONS	90
3	METRICS FOR BRIDGE PERFORMANCE		91
3.1	INTRODUCTION	92
3.2	MEASURE OF BRIDGE RESPONSE	93
3.2.1	Strain-Distance-Area (SDA)	93
3.2.2	Data Pre-processing	98
3.2.3	Benefits of SDA	98
3.2.4	Bridge Flexibility Over Time	102
3.3	CONCLUSIONS	109
4	PREDICTING BRIDGE RESPONSE	111
4.1	INTRODUCTION	112
4.1.1	Technical Background	113

4.1.2	Data Pre-processing	116
4.2	METHODOLOGY	117
4.2.1	Multilinear Regression (MLR)	117
4.2.2	Artificial Neural Network (ANN)	119
4.2.3	Regression Tree (RT)	121
4.2.4	Influential Observations Analysis (IOA)	122
4.3	RESULTS	123
4.3.1	Multilinear Regression (MLR)	123
4.3.2	Artificial Neural Network (ANN)	128
4.3.3	Regression Tree (RT)	129
4.3.4	Influential Observations Analysis (IOA)	133
4.4	DISCUSSION	133
4.5	CONCLUSIONS	136
5	CLUSTERING BRIDGE RESPONSE AND TRUCK CLASSIFICATION	138
5.1	INTRODUCTION	139
5.1.1	Technical Background	140
5.1.2	Longitudinal K-Means (KmL) in Other Fields	141
5.2	METHODOLOGY	142
5.2.1	Data Pre-processing	142
5.2.2	Longitudinal Clustering	146

	10
5.2.3 Truck traffic classification 149
5.2.3.1 Image processing 149
5.2.3.2 Classification Tree 150
5.3 ANALYSIS STRATEGY (PROOF OF CONCEPT) 151
5.4 RESULTS 167
5.4.1 Monthly clustering 169
5.4.1.1 Sensitivity Analysis 186
5.4.2 Monthly Classification Tree 200
5.5 CONCLUSIONS 214
6 CONCLUSIONS 216
6.1 OVERVIEW 216
6.2 CONTRIBUTION 216
6.3 LIMITATIONS AND FUTURE WORK 220

LIST OF TABLES

Table 1.1	case study 1, bridge estimated inspection costs	29
Table 1.2	case study 1, wired and wireless SHM costs for case study bridge	30
	case study 2, sample of bridge inspection costs for selected		
Table 1.3	transportation agencies	31
Table 1.4	case study 2, summary of costs per individual technology	32
Table 3.1	α_e values from fitting line	107
	chronological list of MLR and ANN applications in structural		
Table 4.1	engineering	115
Table 4.2	estimated coefficients of MLR model	124
	ANN model calibration parameters: decay λ and size, based on the		
Table 4.3	best R^2 and lowest	129
Table 4.4	mean values of R^2 after cross-validation	133
Table 4.5	ranking criteria, 1=Best	135
Table 5.1	43 trucks clustered	155
Table 5.2	summary of trucks characteristics within each cluster.	157
Table 5.3	comparison of 18 truck's characteristics	159
Table 5.4	reassignment (4 cluster-option)	164
Table 5.5	percentages of trucks in each cluster	170
Table 5.6	logging trucks probabilities	170

Table 5.7	maximum strain values of each cluster	175
Table 5.8	August 2010	202
Table 5.9	September 2010	203
Table 5.10	October 2010	204
Table 5.11	November 2010	205
Table 5.12	December 2010	206
Table 5.13	March 2011	207
Table 5.14	April 2011	208
Table 5.15	March 2012	209
Table 5.16	April 2012	210
Table 5.17	June 2012	211

LIST OF FIGURES

	state map of the 2018 structurally deficient bridges report,		
Figure 1.1	(ARTBA)	23
	structurally deficient bridges on the U.S. National Highway		
Figure 1.2	System, (ARTBA)	23
Figure 1.3	bridge view	28
Figure 1.4	sensors layout	28
Figure 1.5	Hurley Bridge (Photo credits David Kosnik)	35
	a) bridge map location; b) aerial view of the bridge, Montreal River		
Figure 1.6	and US Highway 2 (Google Maps)	35
Figure 1.7	plan view	36
Figure 1.8	cross section with truck position visualization	36
Figure 1.9	Logging truck (photo credits David Kosnik)	38
Figure 1.10	sawmills and lumber store close to the bridge	38
Figure 1.11	WIM site overall layout (courtesy of David Kosnik)	39
	WIM details westbound lanes, dimensions in ft (courtesy of David		
Figure 1.12	Kosnik)	40
Figure 1.13	photo of the WIM site (Google Maps)	40
Figure 1.14	two types of WIM's output with axle configuration information	41
	location of SHM sensors on the bridge (photo credits David		
Figure 1.15	Kosnik)	43
Figure 1.16	bridge configuration and location of SHM sensors on the bridge	43

Figure 1.17	dynamic responses of the bridge caused by truck crossing in lane 1	44
Figure 1.18	dynamic responses of the bridge caused by truck crossing in lane 2	44
	hourly data since November 2009, (left) strains, (right)		
Figure 1.19	displacements, air temperature, and relative humidity	46
Figure 1.20	combined output of the WIM-SHM system	48
Figure 1.21	scheme of the concurrent events not included in the study	49
	cumulative displacement map over greater Vancouver (Jan		
Figure 1.22	2010- Feb 2012)	52
Figure 1.23	interferogram of the Lions Gate Bridge	53
Figure 1.24	GeoSHM sensor locations on the Forth Road Bridge	55
Figure 2.1	waveform generated by the truck crossing and ZTP identification	61
Figure 2.2	scatterplots of ZTP vs GVW, by girder (March 2011)	64
Figure 2.3	normalized ZTP vs GVW, by girder (March 2011)	67
	AASHTO _{GDF} , 3-year averages (μ) of ZTP _{GDF} , \pm std. dev. a) lane 1;		
Figure 2.4	b) lane 2	71
	“Load Sharing among girders based on dynamic strain		
Figure 2.5	measurements 1997” (Tennyson et al., 2001)	72
Figure 2.6	load percentages carried by each girder, by lane	73
Figure 2.7	AASHTO _{GDF} and calculated ZTP _{GDF} maxima	73
	3-year averages, maximum values and AASHTO specifications. a)		
Figure 2.8	lane 1; b) lane 2	74

Figure 2.9	ZTP _{GDF} vs GVW, girder 3 and 4, lane 1; girder 2 and 3, lane 2	76
Figure 2.10	ZTP _{GDF} vs length, girder 3 and 4, lane 1; girder 2 and 3, lane 2	77
Figure 2.11	ZTP _{GDF} vs speed, girder 3 and 4, lane 1; girder 2 and 3, lane 2	78
Figure 2.12	1-day crossing, Saturday, 20 February 2010	82
Figure 2.13	daily average, lane 1	83
Figure 2.14	daily average, lane 2	84
Figure 2.15	1-week crossing, 20-26 February 2010	85
Figure 2.16	weekly average, lane 1	86
Figure 2.17	weekly average, lane 2	87
Figure 2.18	monthly average, lane 1	88
Figure 2.19	monthly average, lane 2	89
	a) strain vs time; b) strain vs distance of two trucks with similar		
Figure 3.1	GVW, length, and speed, (girder 4).	95
Figure 3.2	theoretical bending moment influence line (M_{IL}).	96
	measured strain waveform (average girder 3 and 4); b) calculated		
	M_{IL} at mid- span, generated by a 331.4 kN (74.5 kips), 22 m (72 ft)		
Figure 3.3	truck	97
	a) normalized ZTP vs GVW; b) normalized SDA vs GVW. Values		
Figure 3.4	are averages of girder 3 and 4, Lane 1, March 2011.	99
	an example of underestimated ZTP due to the presence of another		
Figure 3.5	vehicle approaching the bridge	100

Figure 3.6	a) ZTP_{GDF} ; b) SDA_{GDF} . Monthly average (March 2011)	101
Figure 3.7	a) ZTP_{GDF} ; b) SDA_{GDF} , girder 4 (March 2011)	102
Figure 3.8	cross-section of steel internal girder and concrete deck	103
Figure 3.9	a) one-month (March 2011), class 9 (5-axle) trucks only; b) three years, from class 9 to class 13 (up to 9-axle) trucks	105
Figure 3.10	M_{IL} at mid-span, truck considered as a point load	106
Figure 3.11	a) SDA-MDA and b) ZTP_{ϵ} - ZTP_{IL} IL calculated with axle weights; c) SDA-MDA and d) ZTP_{ϵ} - ZTP_{IL} IL calculated with GVW	109
Figure 4.1	actual SDA against predicted SDA a) for MLR with (W IO); b) for ANN with (W IO); c) for MLR without (W/O IO); for ANN without (W/O IO). All SDA values are standardized ranging from 0 to 1	126
Figure 4.2	a) scheme of MLR; b) scheme of ANN	127
Figure 4.3	deviance vs tree size (or complexity parameter) λ	130
Figure 4.4	RT with 8 nodes	131
Figure 4.5	RT with 6 nodes	132
Figure 5.1	selection of the number of clusters according to the Calinski- Harabatz quality criterion	149
Figure 5.2	day-time photo and night-time photo, March 2012, lane 2	150
Figure 5.3	clustering output, 3 clusters	153

	from left to right, Calinski-Harabatz criterion (y-axis is the $C(k)$ of Eq.47, the x-axis represents the 20 iterations); 3 clusters and		
Figure 5.4	waveform; waveforms color-coded by cluster	154
Figure 5.5	sample images of trucks by cluster	157
Figure 5.6	two identical logging trucks belonging to two different clusters	158
	(above) March 29 at 23:19:18, cluster B; (below) March 28 at		
Figure 5.7	22:05:07, cluster C	160
Figure 5.8	clustering output, 4 clusters	161
	from left to right, Calinski-Harabatz choice method, 4 clusters,		
Figure 5.9	waveforms color-coded by cluster	162
Figure 5.10	cluster reassignment	163
Figure 5.11	classification tree, 3 clusters	165
Figure 5.12	classification tree, 4 clusters	165
Figure 5.13	flow chart of the two-step strategy	166
Figure 5.14	visual summary of trucks classification	168
Figure 5.15	monthly and yearly average percentages of trucks by cluster	171
	monthly and yearly average percentages of trucks by cluster, with		
Figure 5.16	C&D representing logging trucks	172
	monthly and yearly average percentages of trucks by cluster where		
Figure 5.17	B&C&D represents heavy trucks	173
Figure 5.18	August 2010, clusters (left) waveforms sorted by cluster (right)	176
Figure 5.19	September 2010, clusters (left) waveforms sorted by cluster (right)	177

Figure 5.20	October 2010, clusters (left) waveforms sorted by cluster (right)	178
Figure 5.21	November 2010, clusters (left) waveforms sorted by cluster (right)	179
Figure 5.22	December 2010, clusters (left) waveforms sorted by cluster (right)	180
Figure 5.23	March 2011, clusters (left) waveforms sorted by cluster (right)	181
Figure 5.24	April 2011, clusters (left) waveforms sorted by cluster (right)	182
Figure 5.25	March 2012, clusters (left) waveforms sorted by cluster (right)	183
Figure 5.26	April 2012, clusters (left) waveforms sorted by cluster (right)	184
Figure 5.27	June 2012, clusters (left) waveforms sorted by cluster (right)	185
	August 2010, 3 clusters, (left) basic normalization, (right)		
Figure 5.28	time- window normalization	187
	September 2010, 4 clusters, (left) basic normalization, (right)		
Figure 5.29	time- window normalization	188
	October 2010, 5 clusters, (left) basic normalization, (right) time-		
Figure 5.30	window normalization	189
	November 2010, 3 clusters, (left) basic normalization, (right)		
Figure 5.31	time- window normalization	190
	December 2010, 4 clusters, (left) 4-rule outliers detection, (right)		
Figure 5.32	ttc-normalization	191
	March 2011, 4 clusters, (left) basic normalization, (right)		
Figure 5.33	time- window normalization	192
	April 2011, 4 clusters, (left) basic normalization, (right)		
Figure 5.34	time- window normalization	193

	March 2012, 4 clusters, (left) basic normalization, (right)		
Figure 5.35	time- window normalization	194
	April 2012, 4 clusters, (left) basic normalization, (right)		
Figure 5.36	time- window normalization	195
	June 2012, 4 clusters, (left) basic normalization, (right)		
Figure 5.37	time- window normalization	196
	<i>Normalization 2</i> , 3 cluster, dataset 1, non-normalized (left) and		
Figure 5.38	dataset 2, normalized (right)	198
	<i>Normalization 3</i> , 3 clusters, dataset 3, peak alignment (left) and		
Figure 5.39	dataset 4, entering-time alignment (right)	199
Figure 5.40	classification tree, August 2010	202
Figure 5.41	classification tree, September 2010	203
Figure 5.42	classification tree, October 2010	204
Figure 5.43	classification tree, November 2010	205
Figure 5.44	classification tree, December 2010	206
Figure 5.45	classification tree, March 2011	207
Figure 5.46	classification tree, April 2011	208
Figure 5.47	classification tree, March 2012	209
Figure 5.48	classification tree, April 2012	210
Figure 5.49	classification tree, June 2012	211
Figure 5.50	comparison of classification trees between two consecutive months	212

comparison of classification trees between two months in two

Figure 5.51 consecutive years

..... 213

1 BRIDGE MAINTENANCE: A NON-SUSTAINABLE TREND TO CURB

This thesis starts with an overview chapter which aims to explain the importance of bridge monitoring and how this work aims to contribute to the field. A comparison of costs and benefits between visual inspection and structural health monitoring is reported as well as some case-studies which focus on the implementation of bridge monitoring from space, as a potential solution for the future. Finally, the subject bridge and its components are described as well as the systems deployed to measure the bridge response (structural health monitoring) and the characteristics of the trucks crossing the bridge (weigh-in-motion).

1.1 CURRENT SITUATION

Highways and bridges are commonly referred to as the “backbone” of the US transportation system and serve as the nation’s central artery of commerce and economic activity. Critical infrastructures are defined by Egan (2007) as “systems that provide critical support services to a country, geographic area or a corporate entity; when they fail, there is potentially a large cost in human life, the environment or economic markets”. The American Society of Civil Engineers (ASCE) gave the overall condition of US infrastructure a grade of D+ in its most recent report (ASCE, 2017). According to the 2018 Deficient Bridge Report (ARTBA, 2018) (Fig.1.1 and 1.2), 54,300 structurally deficient U.S. bridges are crossed 174 million times daily across the nation. In 2017, out of 612,677 total bridges, 54,259 are rated as structurally deficient, meaning that 1 in 9 of the nation’s bridges is structurally deficient (data from the 2017 National Bridge Inventory ASCII files, released in January 2018 by the Federal Highway Administration (FHWA)). The average age of a structurally deficient bridge is 67 years, compared to 40 years for non-deficient bridges. The replacement and rehabilitation costs associated with those bridges are \$21.5 billion and \$14.5 billion, respectively (Halsey III, TWP, 2016). Not all deficient bridges are in danger of collapse, however, there are consequences that come with their decayed state. One of them is the need to impose weight restrictions which can have an impact on the routes of heavy trucks, which remain the highest volume mode in the US business supply chain. That can cause delays and those delays, ultimately, may cost the average consumer money. For example, the average daily delay for a UPS truck, due to substandard road conditions, congestion, and capacity issues is five minutes and that translates to \$105 million in additional annual cost to this one company alone (Monga, WSJ, 2015). Other carriers suffer similarly, and the increased cost is passed on to consumers.

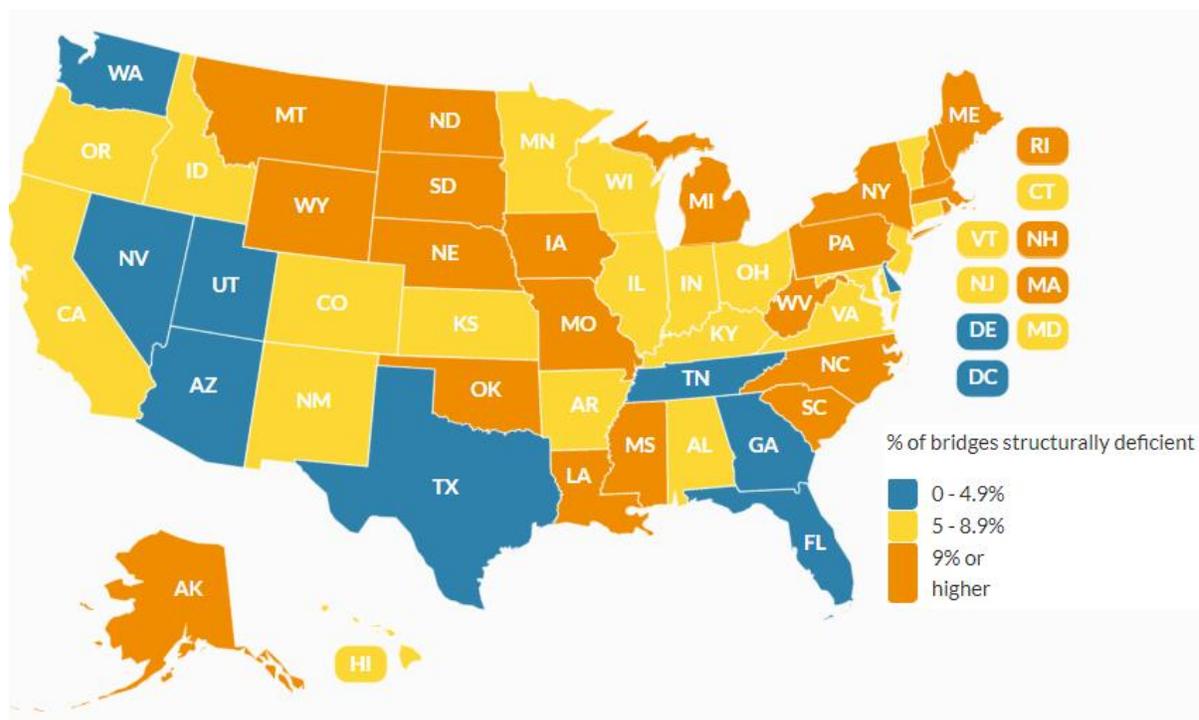


Figure 1.1: state map of the 2018 structurally deficient bridges report, (ARTBA).

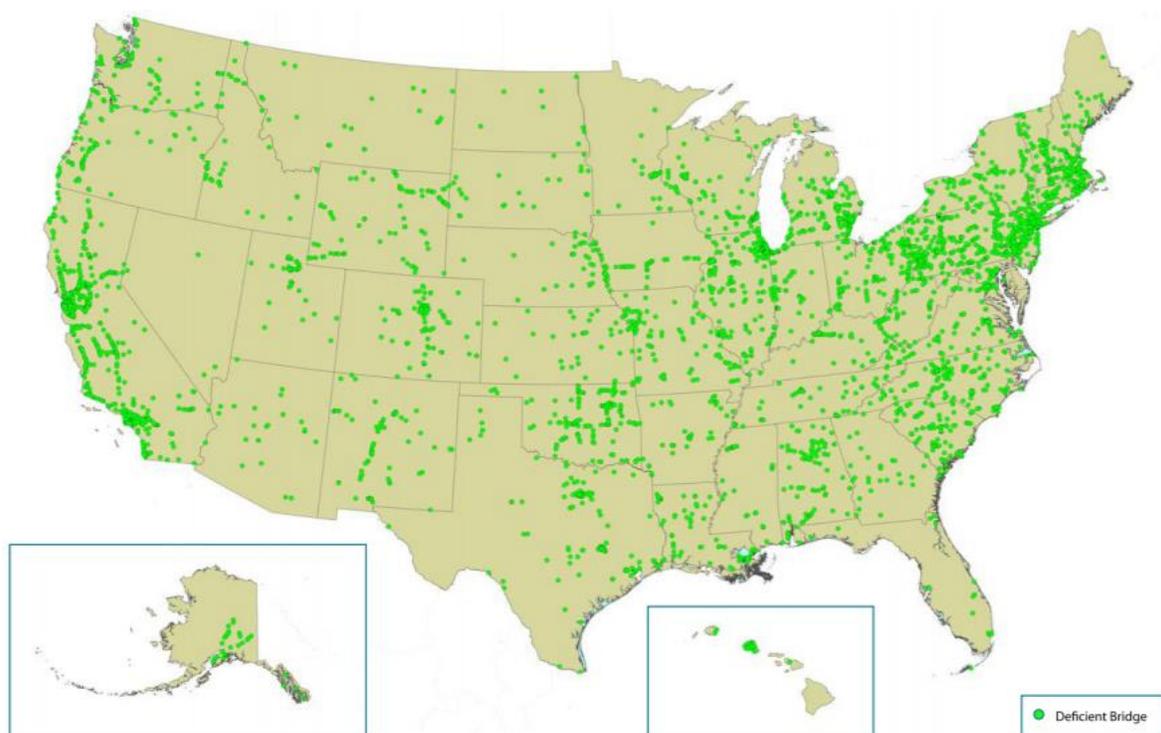


Figure 1.2: structurally deficient bridges on the U.S. National Highway System, (ARTBA).

Given that almost all economic activity, to some degree, depends on the mobility of our highways, roads, and bridges, keeping America's roads and bridges in a state of good repair can positively contribute to a robust economy, improve labor productivity and favorably impact the quality of life of all Americans, including improved public health and higher energy efficiency (U.S. Treasury Department, 2012).

1.2 IMPORTANCE OF MONITORING

America's highway network is underperforming, causing a significant impact on the economy by preventing goods and people to move efficiently. It has proven challenging for state and local governments to keep pace with the growing nation's bridge needs. Consequently, understanding bridge performance and detecting structurally deficient components are crucial steps to assure timely and effective interventions to improve long-term durability and serviceability (DeWolf et al., 2002). The sooner repairs are made, the cheaper they are. According to Kahn and Levinson (2011), every \$1 in preventive maintenance saves between \$4 and \$10 in future repairs (Jaffe, 2015). Bridges are structures designed to display ductile failure modes in case of collapse, and consequently, most of them will display early warnings when the structure is under extreme loads, providing enough time for remedial actions. With this in mind, structural health monitoring (SHM), intended as the practice of monitoring a structure to ensure that its structural integrity and safety remain satisfactory, is ideal for detecting early warnings of possible trends which can lead to failures or need for rehabilitation and strengthening. The broad objectives of SHM are to measure bridge condition, evaluate in-service performance, detect deterioration, determine required maintenance, and estimate remaining service life (Cusson et al., 2012). Adopting SHM techniques on critical bridges can contribute to addressing some of today's challenges and

improving inspection, repair, and rehabilitation methods and reducing traffic disruption. For instance, a more accurate knowledge of the life cycle performance of a bridge network through SHM can provide more complete and timely information to decision makers for an improved management of maintenance and rehabilitation of highway bridges. As a result, timely identification of potential problems can help mitigate their impact on structural health and reduce bridge rehabilitation costs, extending the service life and minimizing life-cycle cost of bridge networks.

With aging structures and increased user demands, proper maintenance and monitoring of bridges are more of a national priority than they have ever been. This thesis aims to provide a contribution to the bridge condition assessment practice to improve the decision-making process by addressing the following issues:

- Preliminary monitoring of bridge performance (Chapter 2);
- Development of bridge response metrics (Chapter 3);
- Modeling and predicting bridge response (Chapter 4);
- Clustering bridge response and trucks' classification (Chapter 5).

Each chapter addresses a different challenge to improve various aspects of the bridge monitoring practice. Chapter 2 shows the preliminary analysis which is usually developed with the data collected using strain gages. The calculation of girder distribution factors (GDF) and aggregate analysis at different level are presented to point out potential trends over time. Chapter 3 proposes new metrics to measure bridge response and bridge integrity over time based on the influence line theory. Chapter 4 provides a comparative study of three machine learning methods to model and

predict bridge response using the metrics developed in chapter 3, and an alternative strategy based on the influential observations analysis is proposed. Chapter 5 presents, for the first time, an application of longitudinal clustering to structural engineering. The strain responses of the subject bridge to heavy trucks are clustered and the characteristics of the trucks which produce each response are extracted. The information obtained can be used by decision-makers, policy-makers and even during design, to understand what kind of traffic is crossing (or it is expected to cross) the bridge and flag those trucks which could cause more harm to the structure.

1.3 VISUAL INSPECTION AND SHM: A COMPARATIVE ANALYSIS OF COSTS AND BENEFITS

The most common technique of bridge monitoring is visual inspection. However, monitoring bridge with sensors is becoming more and more common due to numerous advantages. Inspection and monitoring are distinct methods, each with its own pros and cons but not mutually exclusive. Their functional differences can be leveraged for a complementary approach to bridge monitoring resulting in potential savings, efficiency of assessment and better decision-making support. Below, two studies (Agdas et al., 2015 and CAR et al., 2012), which developed the cost assessment of visual inspection and structural health monitoring, are shown.

Visual inspection is the default bridge inspection methodology governed by National Bridge Inspection Standards (NBIS) published by the FHWA. It is required every two years, however, flexibility is given to the state and federal agencies to increase the frequency when necessary. The inspection is done mostly visually, more advanced inspections may include the execution of destructive and non-destructive testing. Monitoring for corrosion and scour is

common practice but also fracture critical members of the bridge, defined by the FHWA as steel members in tension, or with a tension element, need to be inspected because in the event of a steel member's failure there is no path for the transfer of weight being supported by the member, causing potential failure of a portion of or the entire bridge to collapse. Visual inspection has certainly some limitations such as the inspection frequency, the timing of the inspection becomes particularly important for deficient bridges; interpretability because visual inspection depends on the inspector's subjective assessment; accessibility is one of the most important shortcomings of the visual inspection, all internal problems which are not visible from surface will not be identified and flagged, and also some areas of the bridge might be difficult to access.

The term structural health monitoring encompasses a range of methods and practices designed to assess the condition of a structure based on a combination of measurement, modeling, and analysis. While non-destructive evaluation approaches seek to discover flaws at the material level resulting in a local damage assessment, SHM is a more global approach to assess the structure. The size and complexity of civil structures like bridges often require this type of approach. SHM data are generated by sensors installed on the bridge need to be processed and analyzed to capture the structure's response and any trend which can lead to potential anomalous behavior. This approach can also provide the benefits of real-time assessment. The SHM system heavily depends on the number of sensors and their locations. A plethora of sensors are now available in the market, strain gages, accelerometers widely have been used for decades to measure structure response, more recently, optical fibers sensors are used to measure strain, temperature, and vibration. SHM has also some limitations such as complexity: size, complexity of the structure and surrounding environment; system maintenance: hardware and software failure, so maintenance

is required for long-term operation; automated data analysis: dedicated personnel is required for monitoring and analyzing the system outputs; costs and implementation: SHM requires a significant initial investment.

1.3.1 Case Study 1

The first case study (Agdas et al., 2015) presents a typical pre-stressed concrete girder bridge in a coastal region, with three 65-ft spans (Fig.1.3) and 8-ft girder spacing and a 56-ft wide deck (Fig. 1.4). The bridge presents known issues with corrosion of pre-tensioning steel and scour.

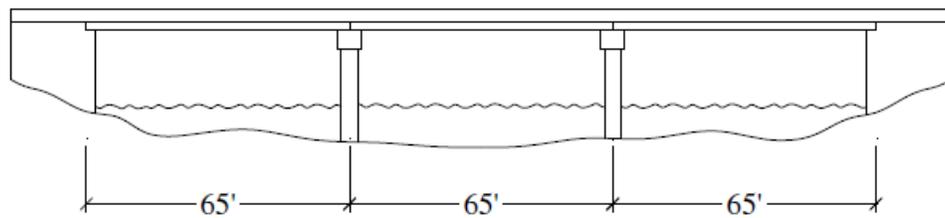


Figure 1.3: bridge view.

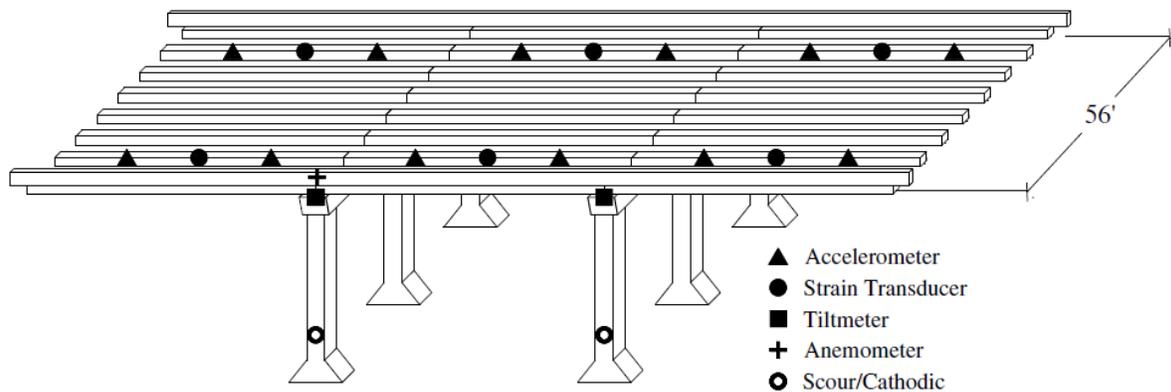


Figure 1.4: sensors layout.

Visual inspection costs

Table 1.1 shows all the items which add up to a total cost: routine field inspection, which are the costs associated with the actual inspection which is expected to take one day, maintenance of traffic (MOT) are expenses due to traffic control for one day, underwater inspection is performed due to the possibility of scour, a snooper, necessary to access the underside of the bridge.

Table 1.1: case study 1, bridge estimated inspection costs.

Cost item	Unit cost	Unit	Quantity	Case study cost
Routine field inspection	\$232	Eq. span ^a	4.91	\$1,140
Routine inspection report	\$155	Eq. span	4.91	\$762
Underwater routine and sub-marine cable	\$185	Eq. span	4.91	\$909
MOT	\$1,500	Day	1	\$1,500
Snooper	\$2,500	Day	1	\$2,500
QA bridge inspection	\$63	Eq. span	4.91	\$310
Snooper mobilization	\$1,285	Each	1	\$1,285
Safety boat	\$80	h	5	\$400
Total	—	—	—	\$8,806

^aEquivalent span is a dimensionless measure of bridge size that includes superstructure and substructure with all incidentals.

SHM costs

The estimation of SHM costs is complicated due to the high number of potential composition of the system, methods, parameters to monitor and other aspects. The bridge is equipped with corrosion and scour sensing sensors, strain gages to track load sharing between girders, accelerometers to investigate changes in the modal properties of the structure over time (Fig.1.4). Wired and wireless systems are considered, with the latter showing a 40% reduction in the initial

costs (Table 1.2), this cost-benefit will increase with the bridge length. The life expectancy of typical system components, with proper maintenance, is approximately 10 years.

Table 1.2: case study 1, wired and wireless SHM costs for case study bridge.

Initial	Hardware	Unit Cost	Unit	Quantity	Wireless	Wired	
	Wireless Processing Unit w/ Embedded Accelerometer	\$ 600	Node Location	14	\$ 8,400	-	
	Accelerometers	\$ 750	Sensor	14	-	\$ 10,500	
	Strain Gauge	\$ 550	Sensor	6	\$ 3,300	\$ 3,300	
	Anemometer	\$ 2,600	Sensor	1	\$ 2,600	\$ 2,600	
	Cathodic Protection	\$ 5,450	Bent	2	\$ 10,900	\$ 10,900	
	Scour	\$ 7,000	Bent	2	\$ 14,000	\$ 14,000	
	Base Station	\$ 6,500	System	1	\$ 6,500	\$ 6,500	
	Software License	\$ 1,000	System	1	\$ 1,000	\$ 1,000	
	Installation & Power						
	Wired Installation	\$ 20,000	Bent	2	-	\$ 40,000	
	Wireless Installation	\$ 8,000	Bent	2	\$ 16,000	-	
	Conduit	\$ 1,020	Span	3	-	\$ 3,060	
	AC Power	\$ 6,240	Span	3	-	\$ 18,720	
	Solar Power	\$ 185	Panel	6	\$ 1,110	-	
				Initial Cost:	\$ 63,810	\$ 110,580	
Ongoing	Bridge Service	Unit Cost	Yearly Occurrence		Wireless	Wired	
	Data Analysis	\$ 2,000	1		\$ 2,000	\$ 2,000	
	Maintenance	\$ 5,000	2		\$ 10,000	\$ 10,000	
				Ongoing Cost / Year:		\$ 12,000	\$ 12,000

1.3.2 Case Study 2

The Michigan Tech Transportation Institute (MTTI) and Michigan Tech Research Institute (MTRI), in cooperation with the Center for Automotive Research (CAR) and the Michigan Department of Transportation (MDOT), have investigated the use of remote sensing technologies to assess and monitor the condition of bridge infrastructures (CAR et al., 2012).

Visual inspection costs

This study collected data through literature review and face-to-face interviews with MDOT partners to establish realistic agency cost estimates of current bridge inspections. Costs of labor are the primary component of bridge inspection costs. For most routine inspections, a team of two can complete four to five bridges per day. Large or complex bridges may take longer. Thus, the cost of a routine inspection can be highly variable, based on factors such as size, location, traffic volume, and construction type. Non-routine inspections (e.g., in-depth, fracture critical) may also cost more. Historical bridge inspection cost data from CAR research are summarized in Table 1.3.

Table 1.3: case study 2, sample of bridge inspection costs for selected transportation agencies.

State/County/City	Bridge Inspection Cost	# of Bridges Inspected Annually	Period	Annual Inspection Cost Per Bridge	Type of Inspection Services
Michigan (1)	\$2.0 million	2,000	FY2010	\$900	In-house plus contract
Wisconsin (2)	\$2.32 million	2,542	FY 2006-07	\$917	\$1.01M for contractors; \$1.31M for in-house
Armstrong County, Pennsylvania (3)	\$482,172	34	2010 to 2015	\$2,398	Contract service with PennDOT
Tulsa County, Oklahoma (4)	\$70,000	195	2007 - 2008	\$359	Contract service
Coal County, Oklahoma (4)	\$18,300	52	2009 - 2010	\$352	Contract service
Logan County, Oklahoma (4)	\$88,000	231	2007 - 2008	\$381	Contract service
Oklahoma Turnpike (4)	\$150,000	399	Since 1998	\$376	Contract service
Tulsa District (4)	\$84,100	12 Spillway bridges	2003	\$7,008	Contract service
Tulsa District (4)	\$109,700	7 Spillway bridges	2004	\$15,671	Contract service

Sources: (1) Interviews with MDOT Bridge Inspection Team; (2) Wisconsin Legislative Audit Bureau (2008); (3) TribLive News (2010); (4) Oklahoma Department of Transportation (2010).

SHM costs

Three MDOT bridges were selected for field demonstration of using the following commercially available remote sensing technologies: 3D Optical Bridge-evaluation System (3DOBS), Bridge Viewer Remote Camera System (BVRCS), GigaPan Photography, Thermal Infrared Imagery (ThIR), Digital Image Correlation (DIC), Mobile Light Detection and Ranging (M-LiDAR), Synthetic Aperture Radar , Ultra Wide Band Imaging RADAR System (UWBIRS). The costs are Some cost elements, presented in Table 1.4, can be measured based on available market data and the field demonstration cost data collection efforts. Others, those with greater uncertainty, are not easily measured, such as final labor costs associated with inspection and data processing time.

Table.1.4: case study 2, summary of costs per individual technology.

Cost Description	3DOBS	BVRCS	GigaPan	ThIR	DIC	M-LiDAR	UWBIRS
Data Collection Systems	\$34,000	\$7,000	\$5,000	\$30,000	\$5,500	\$500,000	\$200,000
Data Collection Vehicle (Cost for one year)	\$9,600	\$9,600	\$9,600	\$9,600	\$9,600	\$9,600	\$9,600
Size of Data Storage File (GB per bridge)	0.1	2.0	10.0	1.0	32.0	7.0	0.2
Data Storage Rate (per GB per year)	\$10	\$10	\$10	\$10	\$10	\$10	\$10
Data Storage Cost (per year)*	\$1.2	\$20.0	\$0.0	\$10	\$320	\$70	\$2
Data Collection Time (hours per bridge)	0.5	0.5	4	0.5	2.5	0.5	0.5
Personnel Needed for Data Collection	2	2	2	2	2	2	2
Total Data Collection Staff Hours	1	1	8	1	5	1	1
Data Processing Time (hours per bridge)	2	1	4	8	4	16	8
Total hours per bridge for all data collection and processing	3	2	12	9	9	17	9
Labor Rate (dollars per hour)	\$50	\$50	\$50	\$50	\$50	\$50	\$50
Total Labor Cost (per bridge)	\$150	\$100	\$600	\$450	\$450	\$850	\$450
Total per bridge Operational Costs (data storage and labor)	\$151	\$120	\$600	\$460	\$770	\$920	\$452
Lane/Shoulder Closure Cost (dollars)	\$0	\$0	\$600	\$0	\$1,563	\$0	\$0
	3DOBS	BVRCS	GigPan	ThIR	DIC	M-LiDAR	UWBIRS
Contractor's Charge per Bridge (Service Fee) (dollars)	\$260	\$260	\$1,500	\$1,300	\$1,500	\$1,800	\$1,300

* GigaPan currently hosts the composite panorama images for its customers at no cost.

The study also provides three different packages: Basic, Enhanced, Premium. They respectively include only ThIR; ThIR and 3DOBS developed by the research team; ThIR, 3DOBS, and UWBIRS developed by the project team. For each package, a period of analysis of 5, 10, or 15 years can be selected.

1.3.3 The South Carolina DOT Saving Strategy

The South Carolina Department of Transportation (SCDOT) adopted a bridge monitoring system to cut costs and increase maintenance efficiency by reducing the number of in-person inspections and collecting real-time information on infrastructure stability and maintenance needs. The bridge monitoring system uses Advanced Condition Assessment Technology (ACAT) to help extend the life of transit infrastructure using girder sensors on the bridges to measure the carrying capacity of the infrastructure. The information collected from the sensors is relayed to the state's computers for 24/7 monitoring and assessment. While the monitoring system does not take the place of all in-person inspections, SCDOT found that the sensors are more accurate in determining the condition of the bridge than visual inspections. This helps to better prioritize projects and, in the long run, it will save money. Indeed, since deploying the monitoring system, SCDOT has reported substantial cost savings. The department was able to retrofit one bridge for \$100,000 rather than replacing it for \$800,000. In the future, an estimated \$5 million in savings is expected as a result of more accurate readings on just eight bridges currently being monitored. If just 1% of the 8,000 bridges in the state were equipped with the monitoring system, a predicted \$56 million could be saved (Greenfield, 2014).

1.3.4 Summary

The two studies presented two ways to analyze the cost of visual inspection and SHM. The first study provides a very detailed analysis of the cost to inspect and monitor one specific bridge with known geometry and known problems of corrosion and scour. The second showed a broader cost analysis, resulted from averaging all bridges in the states considered in the survey. The average includes all kind of bridges, long-span as well short-span, structurally deficient, newly renovated and new bridges, located in different areas with different challenges. It is important to do these considerations when consulting those tables of costs, otherwise, the costs might seem not comparable. In terms of choosing the best method, it can be concluded that both visual inspection and SHM present benefits which complement each other to provide a complete and detailed assessment of bridges.

1.4 SUBJECT BRIDGE DESCRIPTION AND DATA ACQUISITION SYSTEMS

The Hurley bridge is the test bed structure for many techniques developed and tested in this thesis (Fig.1.5). It is a five-girder, three-span continuous steel bridge over two piers with a composite concrete deck and carries the two westbound lanes of US Highway 2 over the Montreal River (the eastbound lanes cross on a separate but identical structure) between Ironwood, Michigan, and Hurley, Wisconsin (Fig.1.6).



Figure 1.5: Hurley Bridge (photo credits David Kosnik).

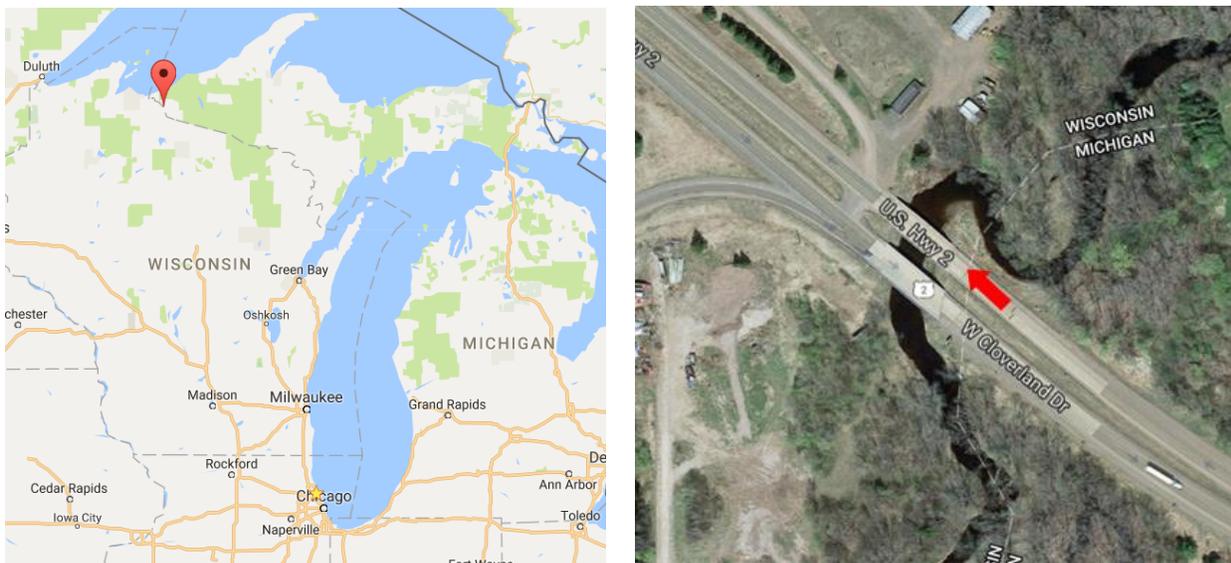


Figure 1.6: a) bridge map location; b) aerial view of the bridge, Montreal River and US Highway 2 (Google Maps).

The overall length of the structure is 37 m (120 ft) (Fig.1.7). The two lanes are 3.65 m (12 ft) wide and the two asymmetric shoulders are 3 m and 1.8 m (10 ft and 6 ft) (Fig.1.8). The structure is skewed by 30°.

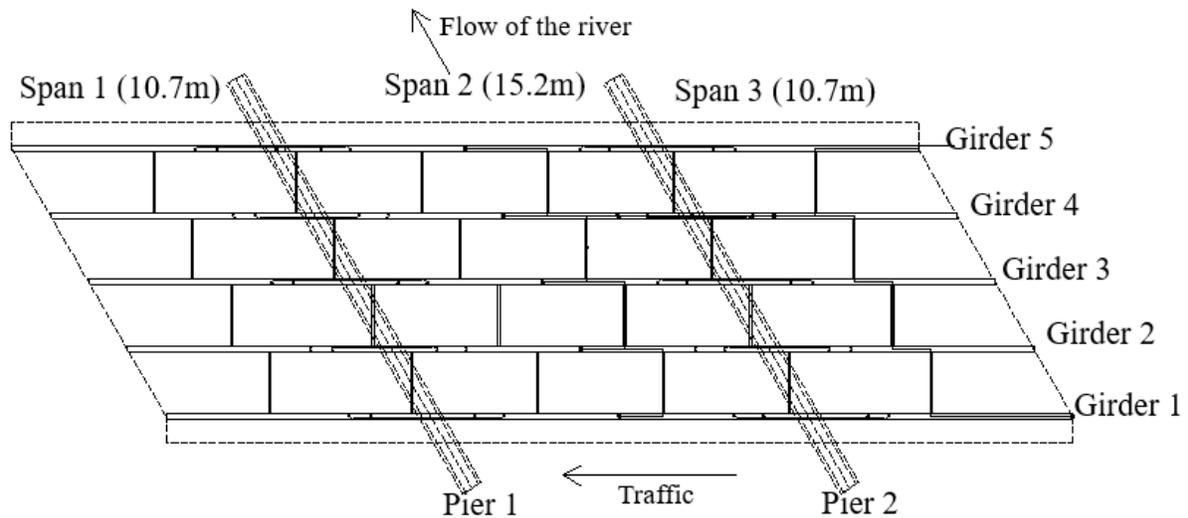


Figure 1.7: plan view.

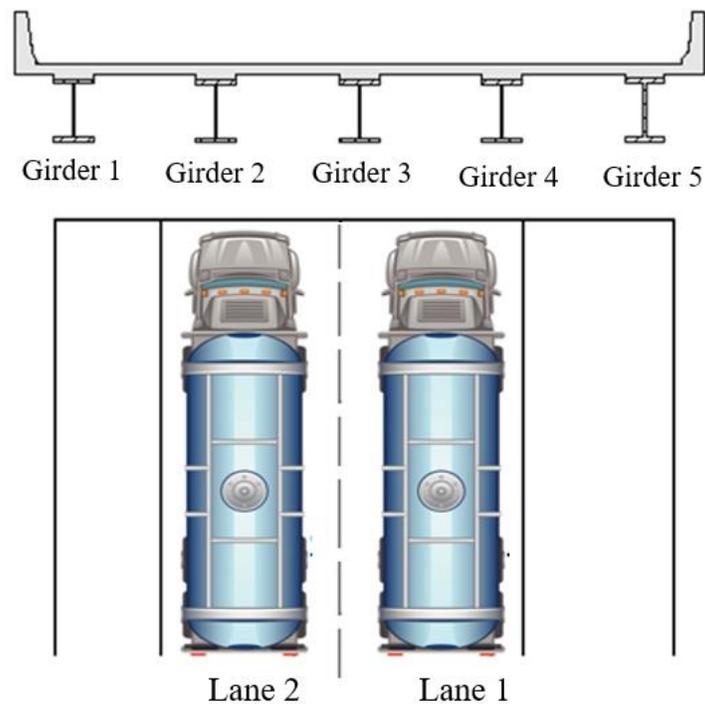


Figure 1.8: cross section with truck position visualization.

Per Wisconsin Department of Transportation (WisDOT) records, the average annual daily traffic is 5200 vehicles, 13% of which are trucks. Concerns over possible effects of increased load limits on the US-2 corridor, including the subject bridge, led to an opportunity to measure both vehicle parameters (system inputs) and structural responses (system outputs) on a multi-year basis by installing a weigh-in-motion (WIM) system and an SHM system during the Summer of 2009. Logging is an important economic activity in the area (Fig.1.9-1.10), and load limits for logging trucks were increased in the subject corridor to promote transportation of timber from Michigan to a sawmill in Wisconsin, approximately 64 km (40 miles) west of the subject bridge. Specifically, the regulation change allows logging trucks loaded per Michigan regulations, which include greater gross vehicle weight (GVW) than Wisconsin, to travel into Wisconsin along the US-2 corridor. For non-logging trucks, the GVW limit on the route is approximately 36.3 metric tons (80 kips); logging trucks are allowed to exceed this gross weight, and routinely do so. In addition, special permit loads such as heavy machinery routinely also cross the bridge. As these increased GVWs and different axle configurations are beyond the design basis for the subject bridge, concerns of overload damage have arisen. For these reasons and for its age and design, typical of many river crossings throughout the country, this bridge is representative of many similar structures nationwide.



Figure 1.9: logging truck (photo credits David Kosnik).



Figure 1.10: sawmills and lumber store close to the bridge.

1.4.1 WIM System

A commercial in-pavement WIM system, which allows measurements of moving vehicles without interrupting traffic, was installed approximately 60 m (200 ft) ahead of the bridge (Fig.1.11) Each lane includes two load cells, enabling the system to weigh left and right wheels

(or wheelsets) on each axle individually. Two inductive loops (L1, L2, L3, L4; Fig.1.12), one before and one after from the scales, and a piezo sensor are also deployed (P1, P2, Fig.1.12). Two cameras, one on either side of the roadway, are available to photograph each vehicle as is passing the WIM site (Fig.1.13). The photo record is especially useful for clarification in situations with multiple vehicles on the bridge, unusual axle or tandem trailer configurations, and snow removal or other specialty vehicles on the bridge. In particular, snow removal equipment was found to produce unusual strain signatures due to a tendency to drive along the shoulders.

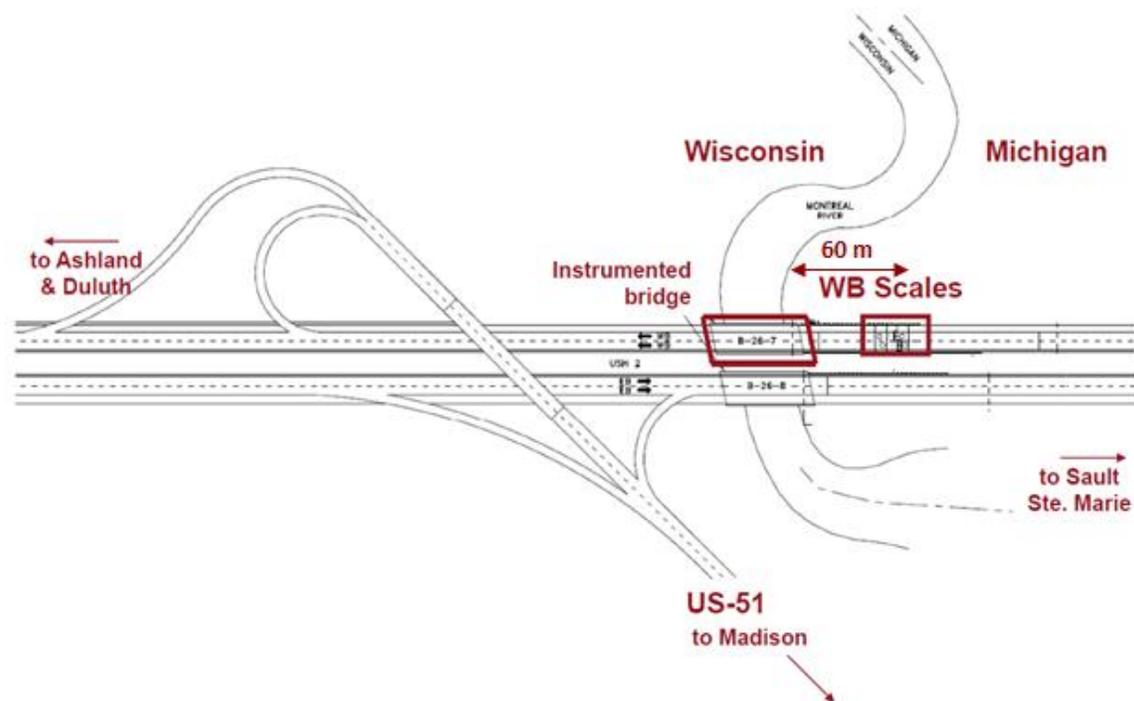


Figure 1.11: WIM site overall layout (courtesy of David Kosnik).

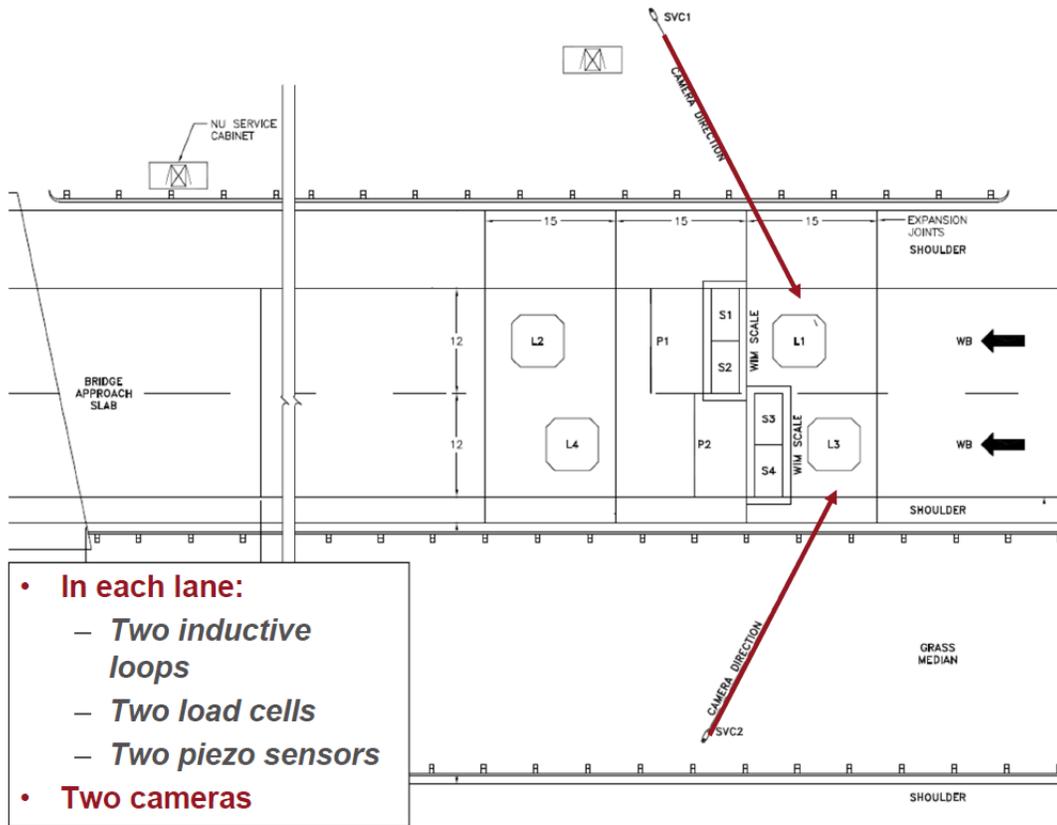


Figure 1.12: WIM details westbound lanes, dimensions in ft (courtesy of David Kosnik).



Figure 1.13: photo of the WIM site (Google Maps).

The commercial system processes signals from the in-pavement sensors to determine vehicle properties such as GVW, overall length, axle configuration (axle weight and spacing), speed, lane of travel, and Federal Highway Administration (FHWA) vehicle class (Fig.1.14). When the WIM system detects a truck classified as FHWA Class 9 or above, it signals the SHM system installed on the bridge. The Class 9 threshold for triggered burst recording was determined empirically to capture trucks of interest while minimizing spurious triggers from passenger or light commercial vehicles (e.g. pickup trucks with trailers). This approach was necessary because FHWA vehicle classes are related to axle configuration, not weight.

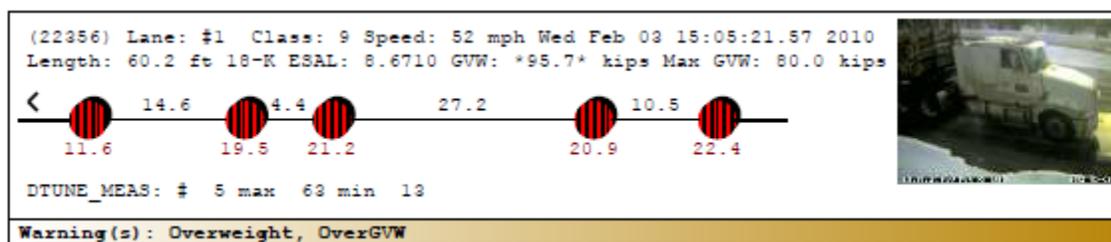


Figure 1.14: two types of WIM's output with axle configuration information.

1.4.2 SHM System

The Infrastructure Technology Institute at Northwestern University (NU-ITI) designed, constructed, and installed an autonomous long-term SHM system on the bridge (Kosnik, 2012). The sensors employed in this study are five strain gauges on the bottom flange of each of the five girders at mid-span, one of the high-stress regions on a girder (Fig.1.15-1.16). When the SHM system is triggered by the WIM system it starts recording dynamic waveforms (sampled at 100 Hz) of the strain, for five seconds, including a half-second pre-trigger buffer. Real-time data are automatically transmitted back to NU-ITI servers, where they are processed and made available in near real-time on a password-protected project Web site (<http://data.iti.northwestern.edu>). Fig.1.17-1.18 show the SHM output for a class 12 truck crossing the bridge on lane 1 (Fig.1.17) and a similar class 12 truck crossing the bridge on lane 2 (Fig.1.18). It is clear from the strain waveform that the load of the former is mostly supported by girder 3 and 4 while the latter is mostly supported by girder 2 and 3. The fascia girders (1 and 5) show a minimal response in both cases. It is also noteworthy that the measured response is dominated by quasi-static strains induced by the weight of the truck; that is the free vibration response is minimal in magnitude and is most evident after the truck leaves the bridge.



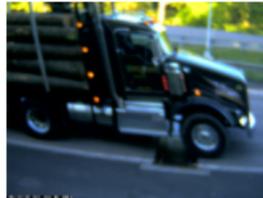
Figure 1.15: location of SHM sensors on the bridge (photo credits David Kosnik).



Figure 1.16: bridge configuration and location of SHM sensors on the bridge.

Dynamic Event Friday, June 1, 2012 at 5:37:19

[Previous Truck \(Friday, June 1, 2012 at 5:32:52\)](#)
 [Next Truck \(Friday, June 1, 2012 at 5:49:12\)](#)



WIM Vehicle #4873 Class 12 **GVW 88.7 kip** length 65 ft 55 mph Lane 1 ([expand](#))

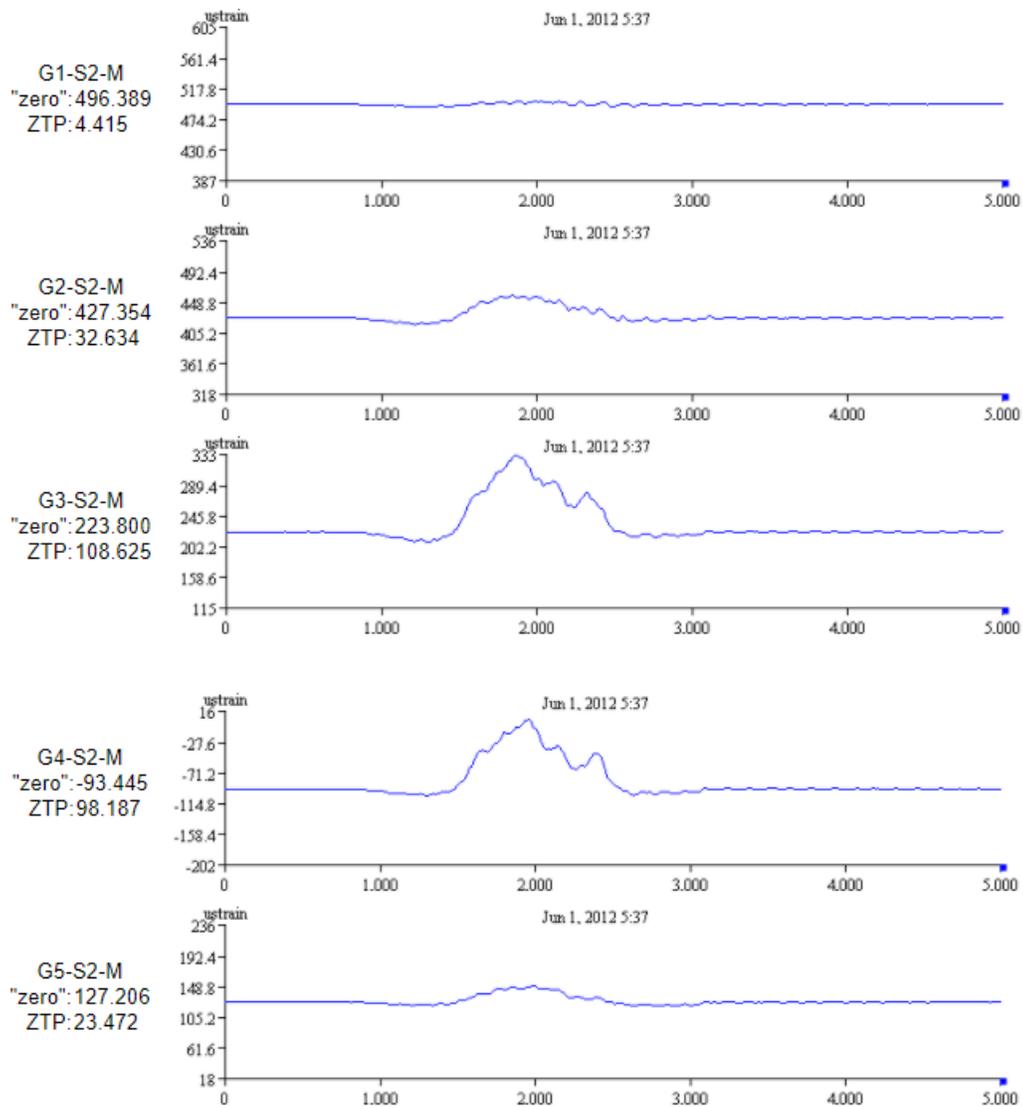


Figure 1.17: dynamic responses of the bridge caused by truck crossing in lane 1.

Dynamic Event Friday, March 30, 2012 at 11:17:07

[Previous Truck \(Friday, March 30, 2012 at 11:06:17\)](#)

 [Next Truck \(Friday, March 30, 2012 at 11:19:24\)](#)



WIM Vehicle #9846 Class 12 **GVW 98.7 kip** length 57 ft 58 mph Lane 2 [\(expand\)](#)

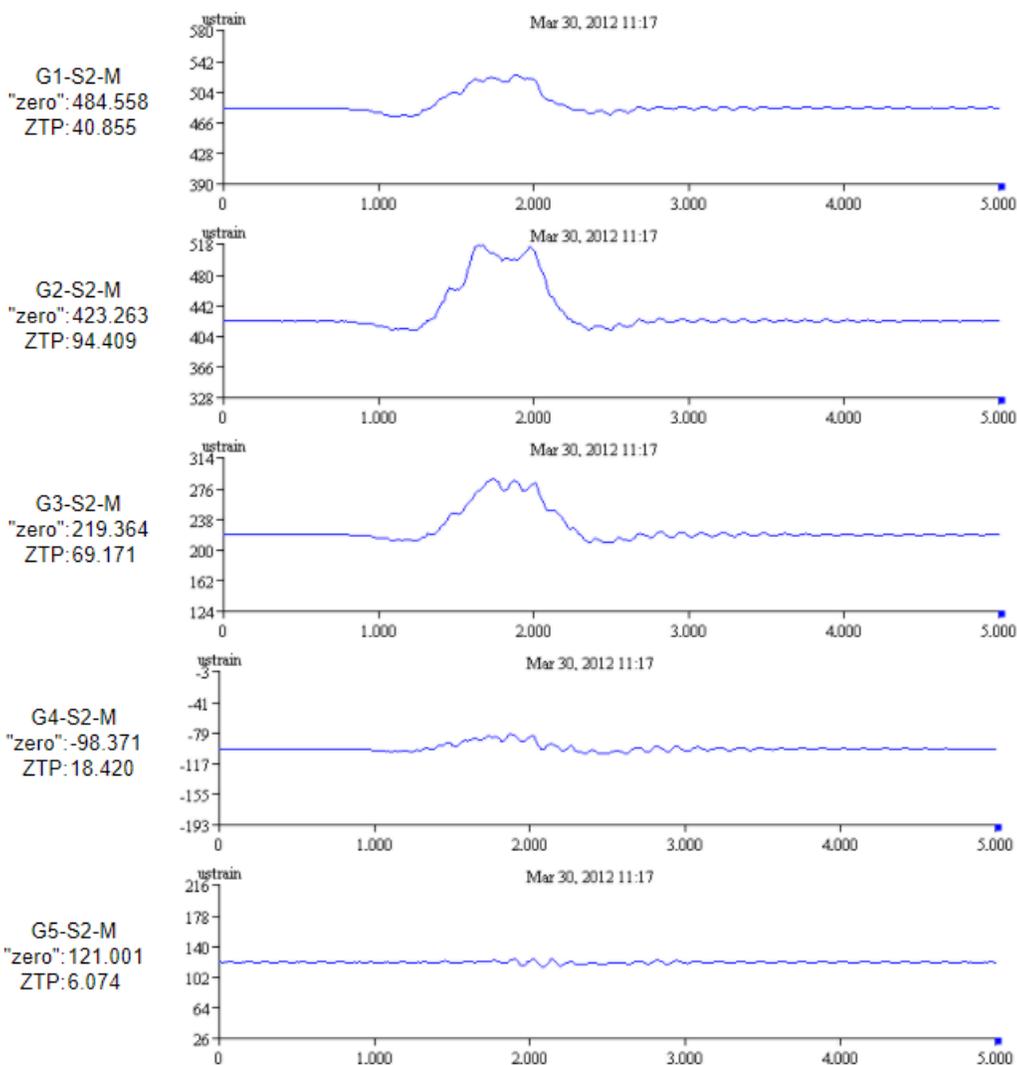


Figure 1.18: dynamic responses of the bridge caused by truck crossing in lane 2.

A combination of scheduled and randomly-triggered recording allowed the SHM system to capture both the long-term and dynamic responses of the bridge. The data acquisition system buffered measurements from all sensors at 100 Hz at all times, with various subsets of the data (or statistics describing those subsets) permanently recorded according to one of the following schemes:

Long-term hourly averages: once an hour, the average of the preceding 60 minutes' worth of data for each instrument was written to the data file. These averages characterize changes in strain and displacements over long periods of time, supporting detection of long-term trends, drifts, or variations in the bridge response (Fig.1.19). This dataset has been used for applications of statistical process control methods in SHM (Chen et al., 2014; Chen and Durango-Cohen, 2015).

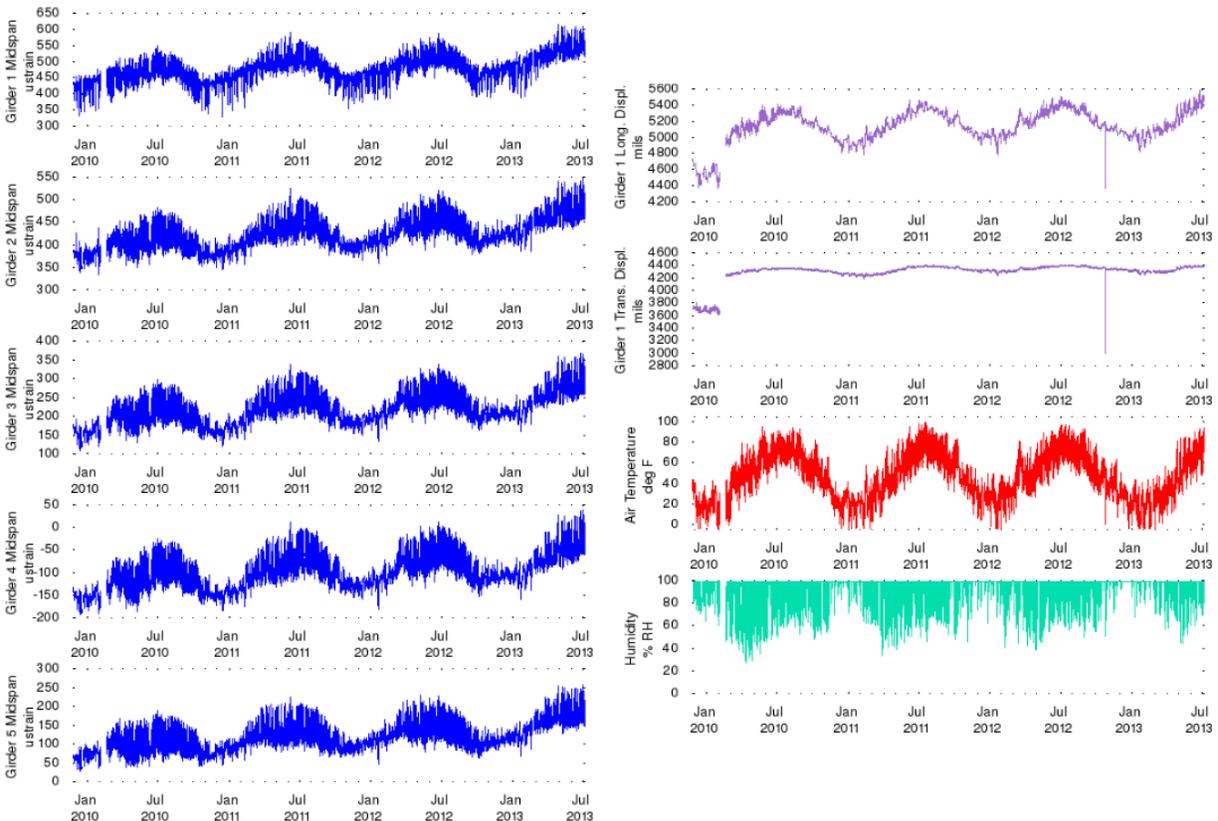


Figure 1.19: hourly data since November 2009, (left) strains, (right) displacements, air temperature, and relative humidity.

Triggered burst events: when the WIM system detects a truck classified as FHWA Class 9 or above, it signals the SHM system to record dynamic waveforms sampling 501 strain data-points in a five-second time frame. Triggered burst events are deployed in this thesis.

1.4.3 Dataset

Data were collected continuously, under live traffic, for over three years (2010-2013) generating a unique long-term WIM-SHM combined dataset. Approximately 171,411 trucks were recorded by the SHM system and 904,604 by the WIM system, however, for only 31,000 trucks the WIM time and the SHM time match. For the preliminary analyses developed in the first part of chapter 2, the SHM dataset is used in its entirety, while for the remaining part of the thesis the combined WIM-SHM dataset is used because both traffic characteristics and bridge response to it are needed (Fig.1.20).

There are some assumptions, described following, which are important when handling real data where several uncontrollable factors have the potential to influence the data. The presence of multiple vehicles on the bridge is not the focus of this study (Fig.1.21). Since the type of traffic examined is heavy trucks which tend to keep right, most of the analyses, not all, consider only lane 1 traffic where girders 3 and 4 carries play the most important role in carrying the load. For some analyses, at each crossing, the average of the strain waveform of girder 3 and 4 is computed for the following two reasons. First, to limit the delay due to the fact that the bridge is skewed of 30° and the signal reaches the mid-span of girder 3 and 4 with a delay of a few milliseconds. Second, to limit the uncertainty due to the lateral position of the truck on the bridge, (i.e. driving more towards the centreline will produce a bigger response on girder 3 while driving more toward the shoulder will produce a larger response on girder 4). It is also noteworthy that, unlike controlled

experiments with pre-weighted trucks with constant speed, these are all measurements from real live traffic, where changes of speed during the crossing can happen and can cause variation in the shape of the strain waveform.

The combination of truck characteristics from the WIM system and bridge response from the strain sensor system provides a comprehensive dataset of bridge loading and subsequent response. The parameters measured were GVW [kg or kN], length of the truck [m], speed [km/h], weight of each axle (ax1, ax2, ax3, ax4, ax5, ax6, ax7, ax8, ax9) [kg or kN], spacing between axles (s12, s23, s34, s45, s56, s67, s78, s89) [m] from the WIM system and strain measurements [$\mu\epsilon$] from the SHM system.



Figure 1.20: combined output of the WIM-SHM system.

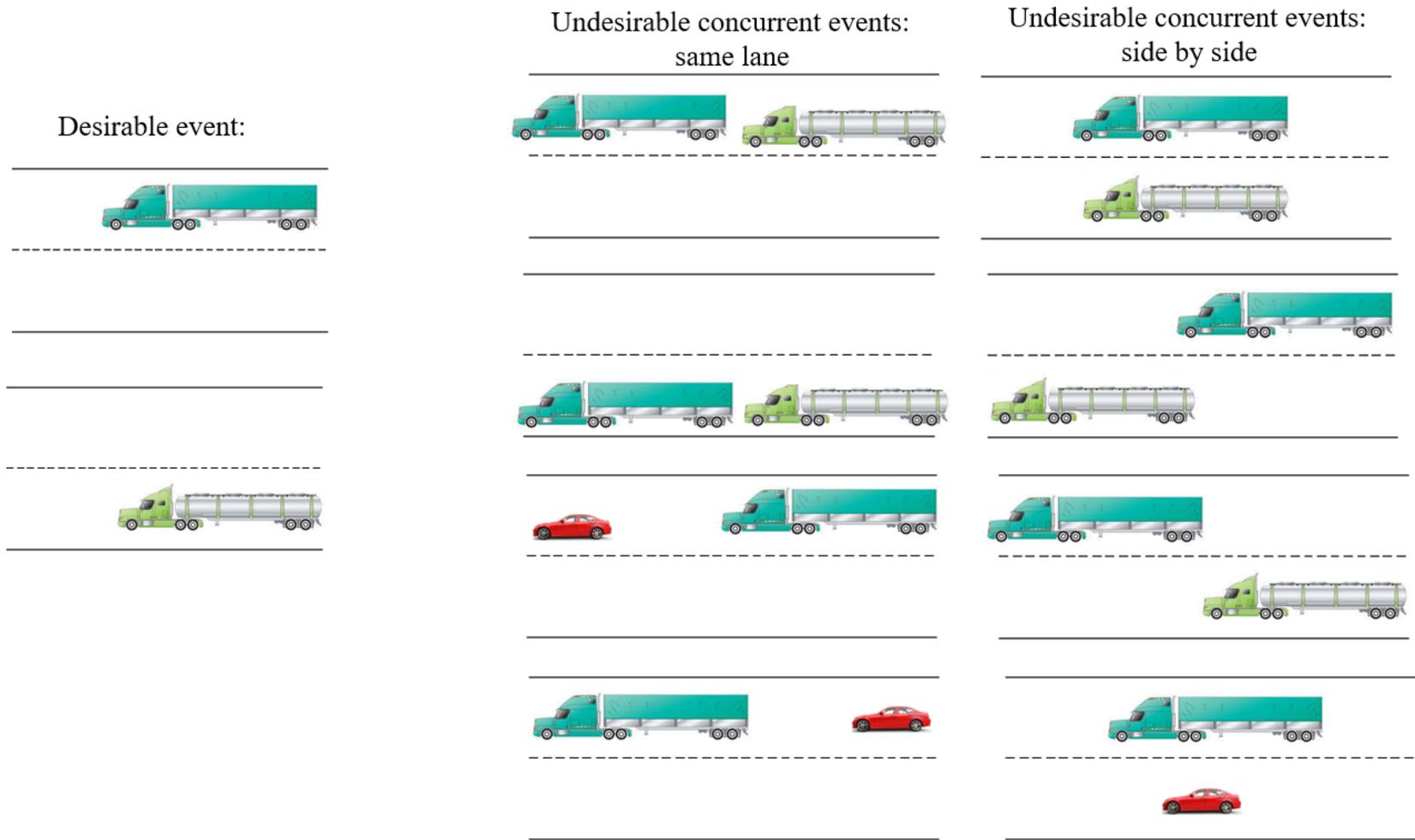


Figure 1.21: scheme of the concurrent events not included in the study.

1.5 FUTURE OF BRIDGE MONITORING

The combination of bridge inspection and health monitoring is probably the most valuable strategy which can help identify structural problems before they become critical and endanger public safety. However, a typical major urban center may possess several hundreds of bridges, which makes it difficult to upgrade all these bridges with surface-mounted sensors to monitor their structural performance due to practical and economic reasons. In-situ sensors are appropriate for targeted monitoring of selected structures, but cannot be readily deployed on a large scale, due to the already limited budgets for bridge maintenance and rehabilitation. Satellite-based monitoring data may offer a viable source of independent information products that may be used to remotely monitor the structural health of bridges, confirm conclusions drawn from in-situ sensor data, and feed decision-support models and tools for pre-emptive bridge rehabilitation. A two-step approach may be used, in which potentially critical bridges are first identified through a screening process by remote satellite-based monitoring, and then further investigated with in-situ monitoring and detailed inspection.

1.5.1 Monitoring Bridge from Space: the R2SHM Project

Canada is investigating satellite technologies for use in prioritizing in-situ monitoring and maintenance of critical bridges. A project entitled RADARSAT-2 Structural Health Monitoring (R2SHM) was initiated to supply bridge products, combined with in-situ data to identify safety-critical structures and quantify the risk they pose to their users. Preliminary results from the application of this technology to transportation infrastructure assets in selected major Canadian urban centers like Vancouver and Montreal are reported. Interferometric SAR (InSAR) is an advanced processing technique applied to radar images of the Earth's surface that can detect very

small movements from ground features such as infrastructure systems, including roadway and railway bridges and their major components. By applying InSAR processing techniques to a series of radar images over the same region, it is possible to detect vertical movements of infrastructure systems on the ground in the millimeter range, and therefore identify abnormal or excessive movement indicating potential problems requiring detailed ground investigation. A major advantage of this technology is that a single radar image, which can be obtained in darkness and in any weather, can cover a major urban area of up to 100 km by 100 km, and therefore all bridges in the area could be monitored cost effectively. The expected accuracy of elevation measurements being in the millimeter range makes this technology very attractive to displacement monitoring of bridges and other infrastructures. Space-borne InSAR, compared to other non-destructive evaluation techniques, offers the potential of rapid assessment of numerous bridges in a single scene from high standoff distance without requiring calibration or preparation of the structure and without interfering with traffic, which is a considerable benefit for busy highway bridges. This technology is best suited for the monitoring of bridge differential settlement, bridge deformed shape and, to a lesser extent, changes in bridge length and bridge deck transverse displacement by extracting the horizontal components of the measured satellite light-of-sight displacement.

Case study 1

The R2SHM project uses radar image data from Canada's RADARSAT-2 satellite launched in 2007, which is one of the world's most advanced commercially available Earth observation radar image providers, over five selected bridges in Vancouver (Granville, Burrard, Cambie, Lions Gate, and Ironworkers Memorial bridges). InSAR can generate wide-area displacement maps that identify and quantify displacement basins occurring in the entire imaged region. Fig. 1.22 illustrates the cumulative displacement map generated over the Greater Vancouver area using a Multi-Fine

RADARSAT-2 dataset (5-meter resolution) over a period of 2 years and many basins are proximal to bridge approaches and exits. The Lions Gate Bridge and the Skytrain (Canada Line) exhibit conspicuous thermal displacement, manifesting as cyclic displacement fringes along the length of the structure. Although some seasonal effects may be observed, the general trends show steadily increasing ground sinking by 20 mm over two years.

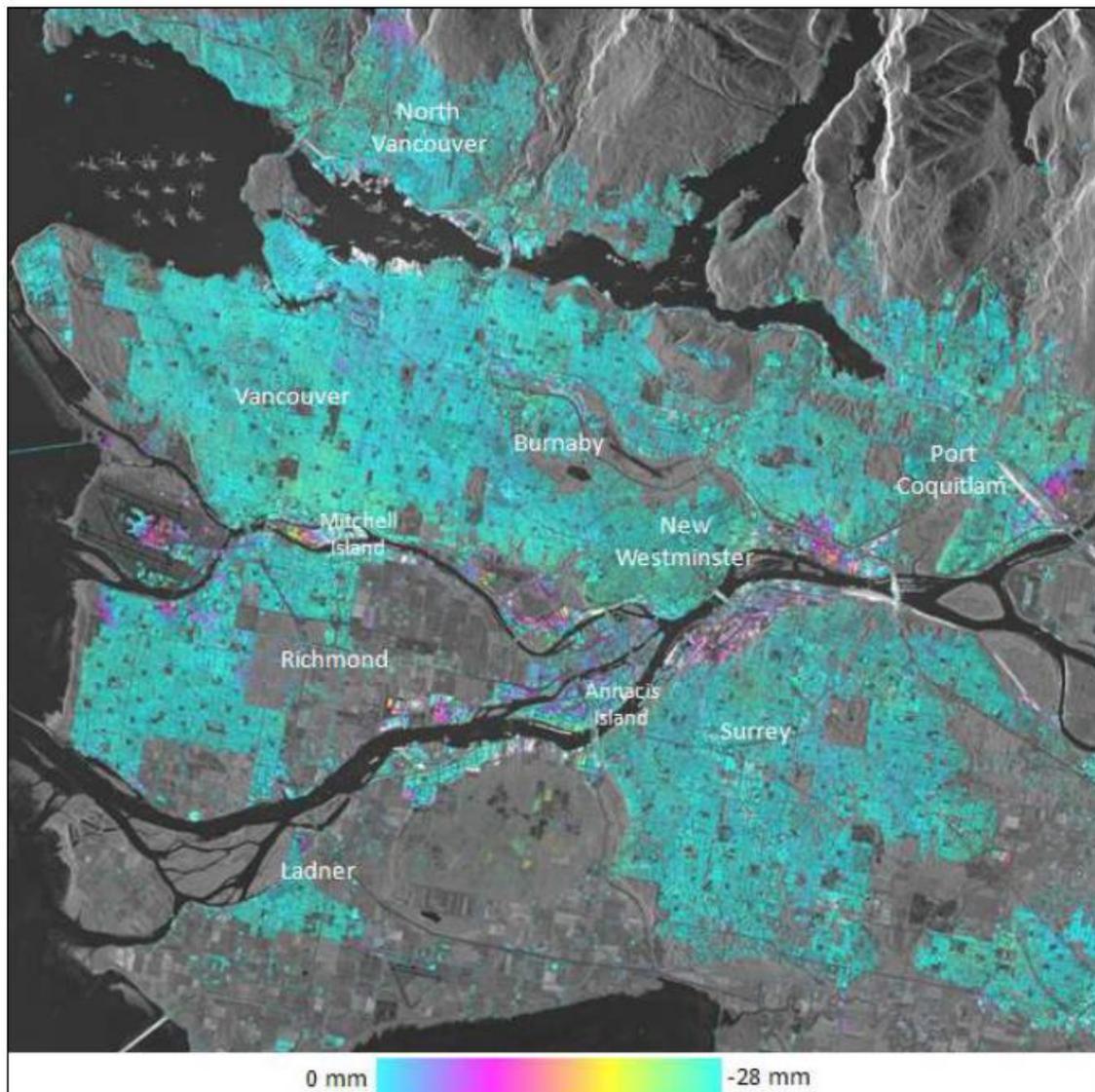


Figure 1.22: cumulative displacement map over greater Vancouver (Jan 2010- Feb 2012).

Case study 2

The Lions Gate suspension bridge in Vancouver, Canada crosses the First Narrows of Burrard Inlet and connects the City of Vancouver to other municipalities on the north shore. The bridge was built in 1938 with a main span of 473 m and two approach spans of 187 m each. Among the number of bridges being monitored in the R2SHM project, measuring vertical displacements of the Lions Gate Bridge is the most challenging, as the interferogram pixels over the bridge comprise two uncorrelated signals: elevation error and strong displacement due to environmental factors. Fig.1.23 illustrates an interferogram of the Lions Gate Bridge, in which the numerous colored fringes are indicative of large displacements occurring over the bridge.

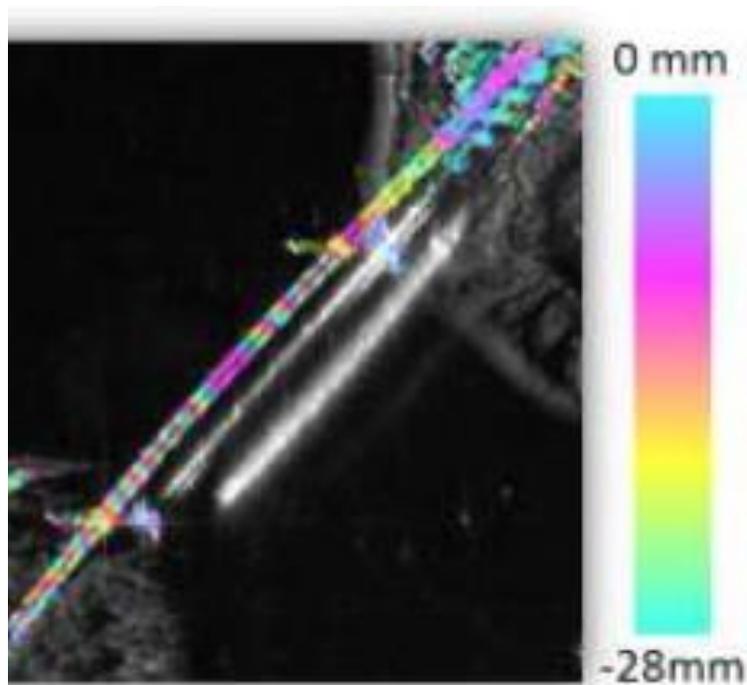


Figure 1.23: interferogram of the Lions Gate Bridge.

Deformation due to changes in ambient temperature between pairs of interferograms is of interest for satellite-based SHM of bridges in two regards. Firstly, it is an expected displacement component and thus must be modeled and removed in order to reveal suspect mechanical

displacements, if any. Secondly, it is used as a means of validating any vertical displacement measurements taken from space: two independent data sets are compared (i.e. temperature records and InSAR-derivate displacements) and the strength of the correlation indicates the degree of confidence in the measurements. The Lions Gate Bridge confirms the finding of another case study by Westgate et al., (2011), where a similar long-span suspension bridge (over the Tamar River in the UK) was monitored using in-situ sensors, and its structural response to environmental variables was analyzed using a calibrated finite element model of the bridge. The effects of temperature on vertical displacements of the bridge were found to be dominant among all other variables. They obtained an almost identical (absolute) correlation coefficient of 0.92 for the correlation of temperature changes and the in-situ vertical displacements measured on the Tamar River suspension bridge. The strong correlation of the regression analysis provides high confidence in the InSAR-derived displacement results at the Lions Gate Bridge and indicates that the thermal component of the displacements can be accurately modeled and removed.

Overall, space bridge monitoring can overcome some limitations of the current practice such as: eliminate lane closure and traffic disruption, as these technologies do not come in direct contact with the structure, access remote bridge locations, hazards in remote areas that often do not receive the required attention due to accessibility issues.

1.5.2 Monitoring Bridge from Space: the GeoSHM Project

The University of Nottingham with the operator of the Forth Road Bridge in Scotland developed a system, called Global Navigation Satellite System (GNSS) and Earth Observation (EO) for Structural Health Monitoring (GeoSHM), which uses space assets for monitoring purposes (ESA, 2016). GeoSHM is an SHM system integrating remote sensing (EO) and GPS-

based sensors (GNSS) monitoring the structural health system of long suspension bridges. While the GeoSHM solution has been initially designed mainly for monitoring long-span bridges, the service also has potential, however, for shorter bridges. Increased loads and extreme weather conditions such as strong winds, cause stressed structural members, unexpected deformations, and frequent bridge closures, respectively. To mitigate these problems, the GeoSHM consortium set out to determine whether it was technically feasible and economically viable to deploy GNSS and EO technologies to monitor more precisely the Forth Road Bridge. Specifically, they wanted to provide real-time indicators of bridge movement by means of highly sensitive sensors at key locations (Fig.1.24), as well as map potential long-term subsidence of the supporting structures.

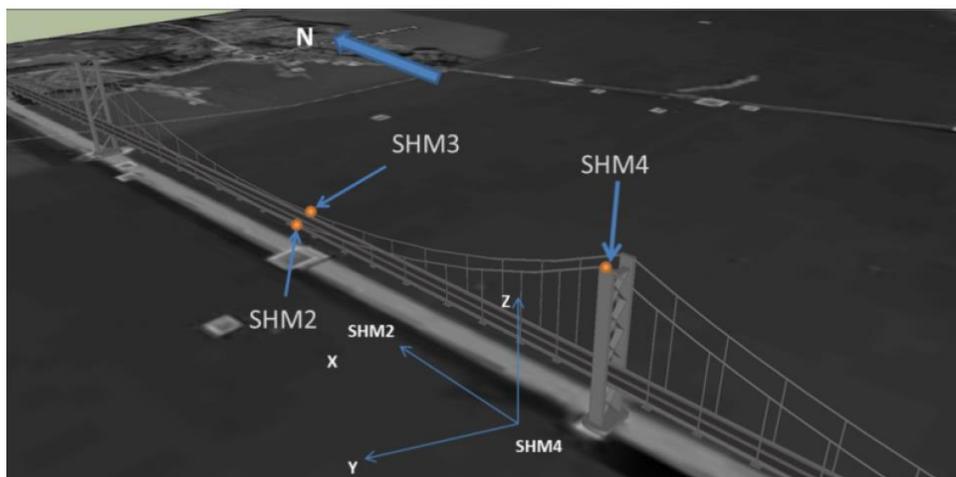


Figure 1.24: GeoSHM sensor locations on the Forth Road Bridge.

The EO is considered a powerful tool to monitor both local changes to the surrounding ground as well as displacements of key components of such structures. Highly sensitive GNSS receivers and anemometers were developed to measure short-term movement, and EO data to measure long-term ground movement. The study on Forth Road Bridge did not show significant changes because the bridge was very stable, however, it was helpful to understand how much the bridge can move under extreme weather conditions. This allows deciding to close the bridge based on precise

deformation information. On the other hand, the same studies were conducted in China were ground subsidence around bridge sites was founded in two cities, Shanghai and Wuhan, caused by underground engineering and groundwater extraction.

2 MONITORING BRIDGE PERFORMANCE

Bridges are structures designed to display ductile failure modes, and consequently, most of them will display early warnings when the structure is under extreme loads, providing enough time for remedial actions. For this reason, it is crucial to monitor bridge response for long-term. Changes in the material properties or in the integrity of the bridge components do not necessarily indicate a risk, however, over a decade or less these changes can be a signal of a trend which should rise a red flag and should be monitored.

The second chapter aims to show a preliminary analysis to monitor the bridge response over a 3-year period by using the strain data available and the concept of the girder distribution factor. Different levels of aggregate analysis are shown. This chapter also aims to validate the AASHTO load distribution factors for the Hurley bridge which, as many others in the country, is subjected to increasing loads due to the increasing demand.

2.1 INTRODUCTION

Interpretation of real data is always a challenging task because of the noise present in the data, environmental and human. Temperature, humidity, driver's behavior, truck's features (suspension, frame stiffness, etc.), cross-section of the bridge all generate uncertainty. One of the biggest complications while passively monitoring is the still unknown interaction between truck's geometry, speed, location, and weight. Although the bridge does not show any evidence of anomalous behavior, not knowing what is the healthy behavior under normal operating conditions is certainly a challenge. It is always good practice to get an overall idea of the bridge health by using field strain data to calculate how the loads crossing the bridge are distributed over the five steel girders using girder distribution factors (GDF). The goal of this chapter is to provide preliminary analyses to monitor the performance of the structure using SHM data for over 171,000 trucks. Long-term aggregate analyses are also developed and shown.

2.1.1 Literature Review

The need for a more efficient use of steel in highway bridge design led the American Iron and Steel Institute (AISI) to the publication of "Tentative Criteria for Load Factor Design of Steel Highway Bridges" in 1969 (Vincent, 1969). The first formulation of the GDF (AASHTO, 1996) intended as the percentage of the design load that each girder will carry, was defined as the ratio between the girder spacing S and a constant based on the bridge type. This formula, also called "S-over", was applicable to non-skewed bridge only. Since that time, the use of AASHTO Load Factor Design (LFD) for steel bridges has continued to arouse interest, mainly due to the fact that there were many important parameters that were not included in it. In order to achieve better accuracy and flexibility across a wider range of bridges, in 1993, the AASHTO LRFD formulas

were developed under the National Cooperative Highway Research Program (NCHRP) project 12- 26 (Zokaie et al., 1991). The new specifications presented major changes. In addition to girders spacing, bridge length and slab thickness, and bridge skewness were finally considered, as well as stiffness of girders and deck. Further corrections were introduced to take into account the difference between interior and exterior girders. In 2001, Tabsh et al. developed a gauge modification factor that multiplies the AASHTO GDFs to consider oversized trucks. This factor is a function of the effect of the flexural or shear load effect, truck configuration and bridge geometry. Also, Bae and Oliva (2009), developed their modification to the code for overloaded trucks. Having more accurate load distribution may allow designing a bridge that can support higher loads, this may result in less money spent on freight transportation since companies can have more loaded trucks (Eom and Nowak, 2001). At this reference, Barr et al. (2001) stated that if the bridge had been designed using the distribution factors calculated with the finite-element model rather than the code values, the required concrete strength can be reduced by 6.9 MPa (1,000 psi) or the live load can be increased by 39%.

Ghosn et al., in 1986, using experimental data, defined the GDF as the ratio of the static strain at the girder to the sum of all the static strains, valid for identical girders only. Stallings and Yoo in 1993 refined this method to account for bridges with different interior and exterior girder sizes using weighted strains. When all girders have the same section modulus, their formula is equivalent to that of Ghosn et al. A similar approach to Stallings and Yoo was used by Neely (2001) to determine the distribution factors from both strain and deflection data for the Tom's Creek Bridge field test. Reiff et al. (2015) analyzed four damaged bridges subjected to crack, corrosion, diaphragm crack, and deck delamination and compared their GDFs to the undamaged structure. While Shenton and Hu (2006) showed how the dead load is redistributed when damage

occurs, Kennedy and Grace (1983), who developed two 118-scale bridge models to verify their theory, stated that transverse cracking of the concrete deck, at the intermediate support(s), does not appear to influence significantly the transverse distribution of the design parameters. Cardini and DeWolf (2008) suggested a continuous long-term SHM system to determine if there are significant changes to the deck or the girders due to major damage.

2.2 PRELIMINARY ANALYSIS

When a load, like a passing vehicle, is applied to the bridge, a reaction is immediately produced. The strain gages installed on the bridge record that reaction by producing five strain waveforms, one for each girder. Preliminary analysis of the SHM strain is necessary to understand the bridge response to heavy trucks. As mentioned in chapter 1 and shown in Fig.1.17-1.18, lane 1 is mostly supported by girders 3 and 4 while lane 2 by girders 2 and 3. It is also noteworthy that the measured response is dominated by quasi-static strains induced by the weight of the truck; that is the free vibration response is minimal in magnitude and is mostly evident after the truck leaves the bridge. The strain waveforms also provide a first hint on the axle configuration of the truck where the peaks represent either a single axle or a tandem axle.

To facilitate the analysis of hundreds of thousands of vehicle crossings, some parameters are immediately computed by the data post-processing script and stored for reference. One important parameter, computed during every vehicle crossing, is the maximum absolute departure from the zero point (ϵ_0) which is the initial strain value prior to loading (Fig.2.1). This parameter is called zero-to-peak (ZTP) and it is defined as follows:

$$ZTP_{i,j} = \max(\epsilon_{t=1}; \epsilon_{t=2}; \dots; \epsilon_{t=500}) - \epsilon_0 \quad (1)$$

where, i is the truck crossing and j is the girder, $j=1, \dots, 5$. During a five-second truck crossing record, 500 data points ($\varepsilon_{t=1}; \varepsilon_{t=2}; \dots; \varepsilon_{t=500}$) are stored for each girder. The zero-point ε_0 is the average of the first ten data points (100 ms); this is within the pre-trigger buffer recorded before the vehicle reaches the bridge. ZTP is an attractive parameter because it captures the effect of each truck crossing while removing effects of temperature and other quasi-static phenomena.

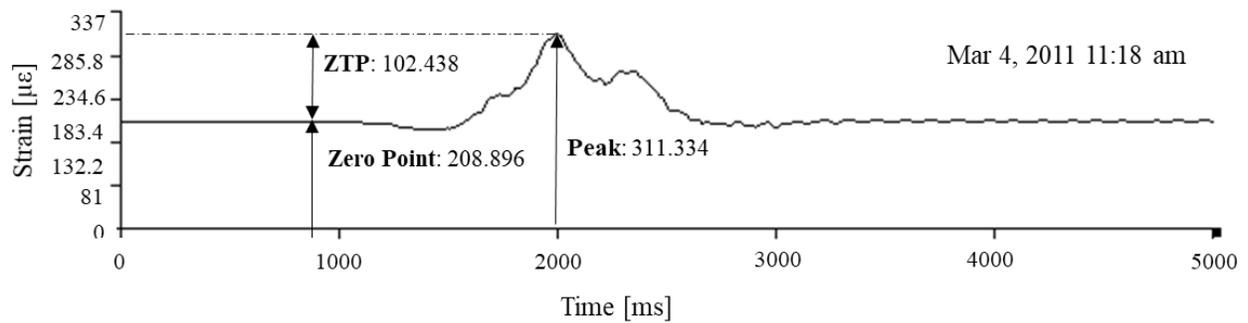
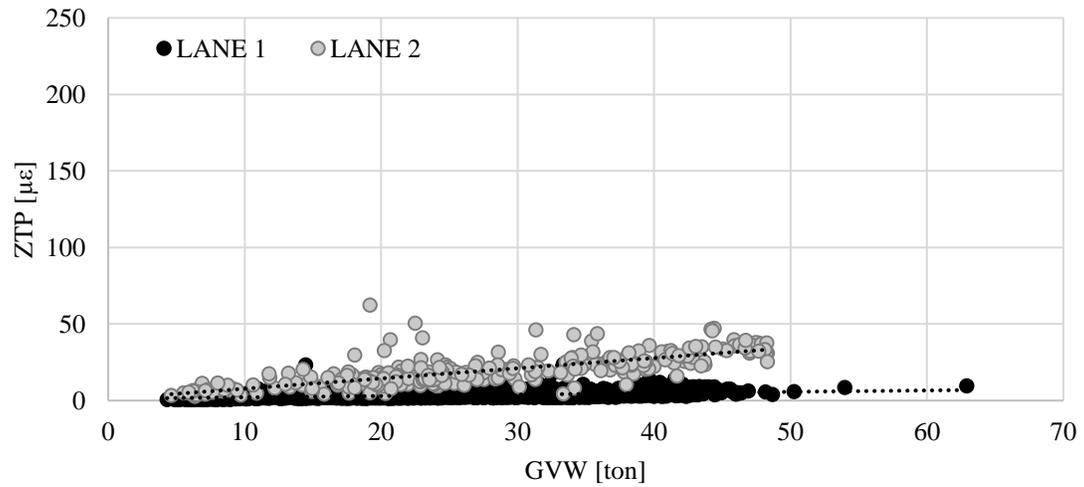


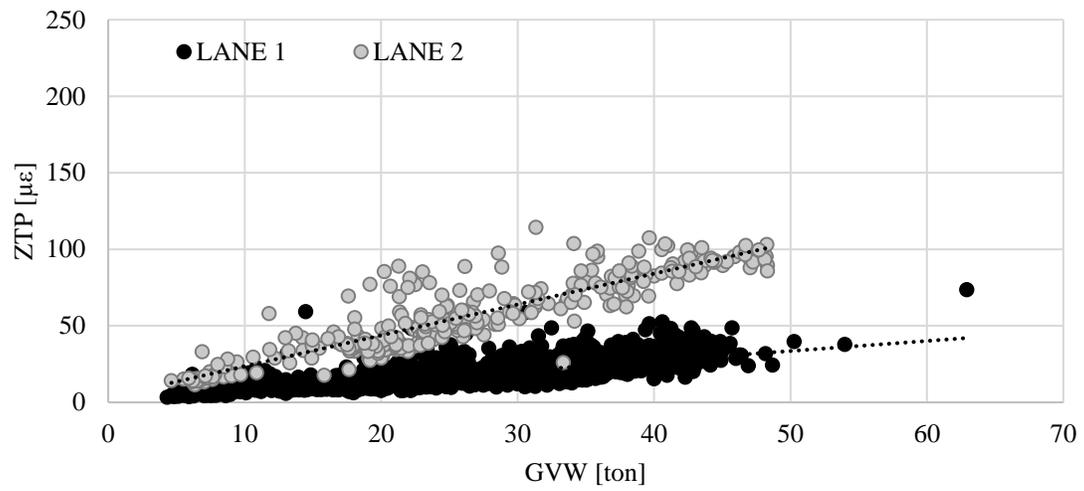
Figure 2.1. waveform generated by the truck crossing and ZTP identification.

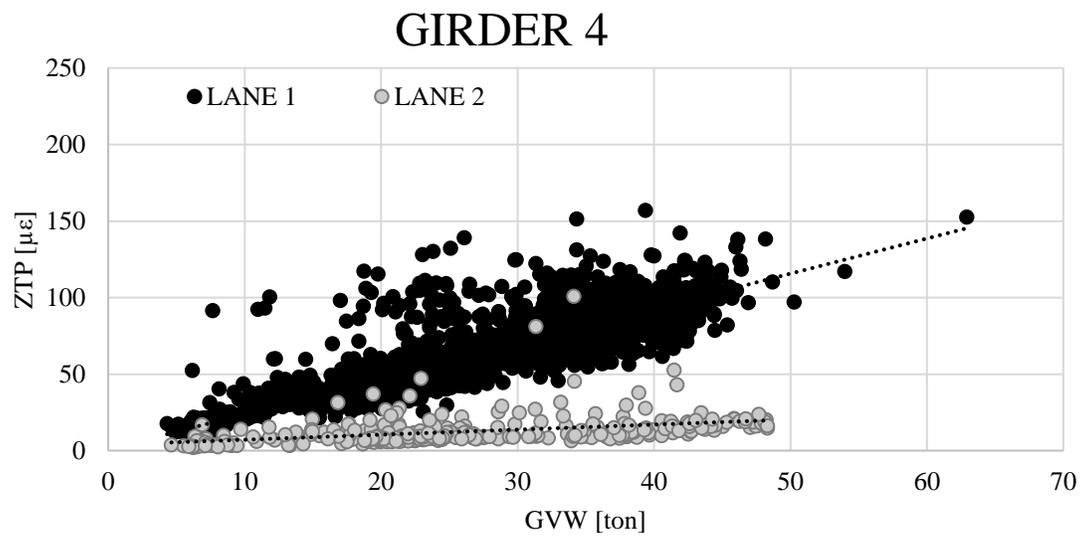
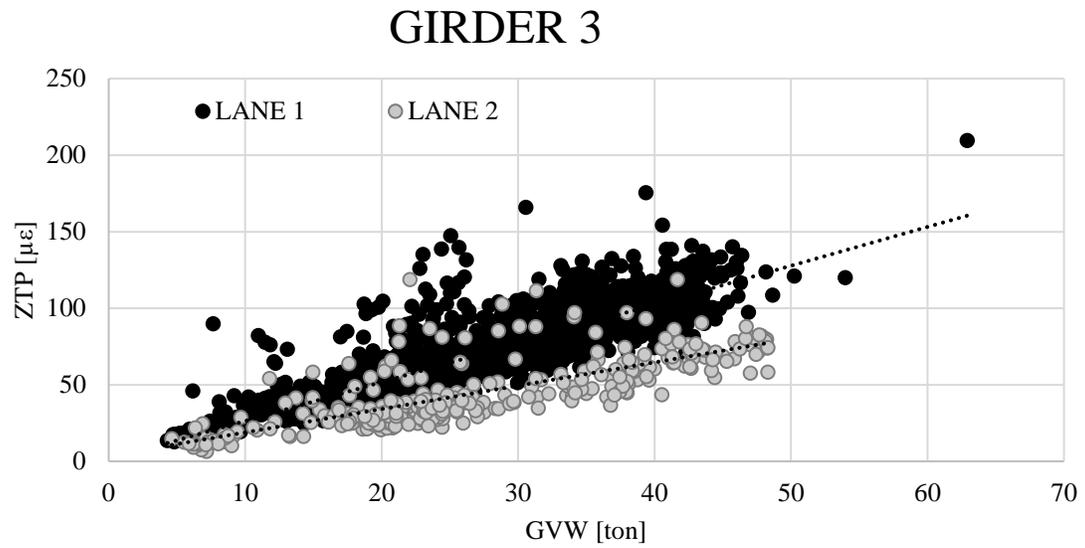
The theoretical linear relationship between strain and GVW are confirmed by the scatterplots in Fig.2.2. For Girders 2 and 4, the lobes in the scatter plot representing data points for the two different lanes are more distant and well-separated. This effect is reduced on Girder 3 (the middle girder), where the distinction between Lanes 1 and 2 is minimal; as such, the resulting data cloud is more compact, though two distinct lobes remain visible, presumably due to the asymmetry of the road and shoulders with respect to the centerline of the bridge structure. The responses of the fascia girders (1 and 5) are significantly lower.

GIRDER 1



GIRDER 2





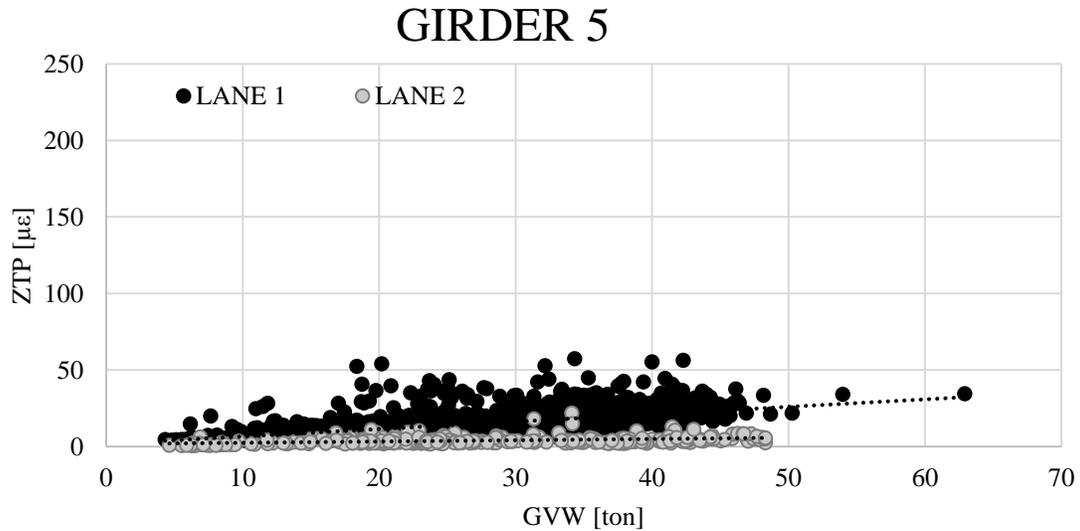
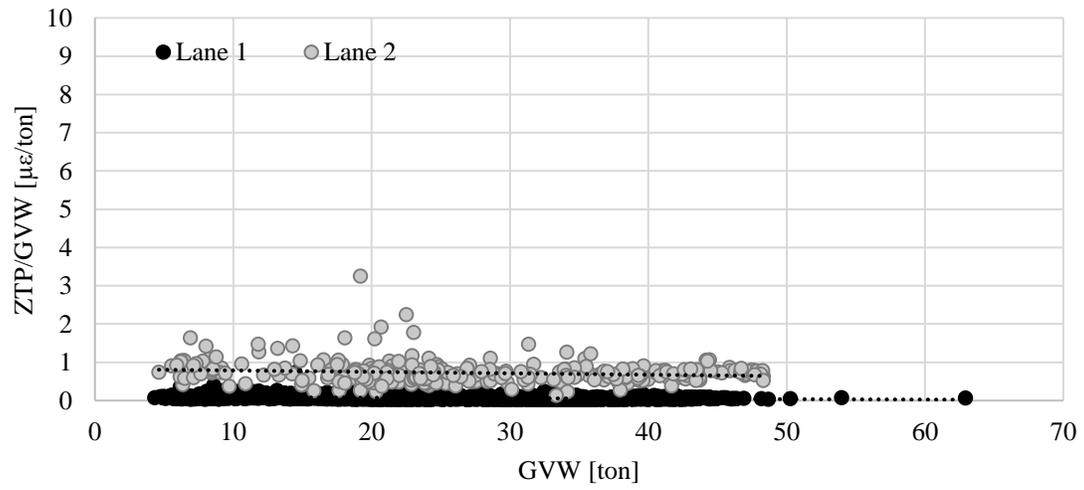


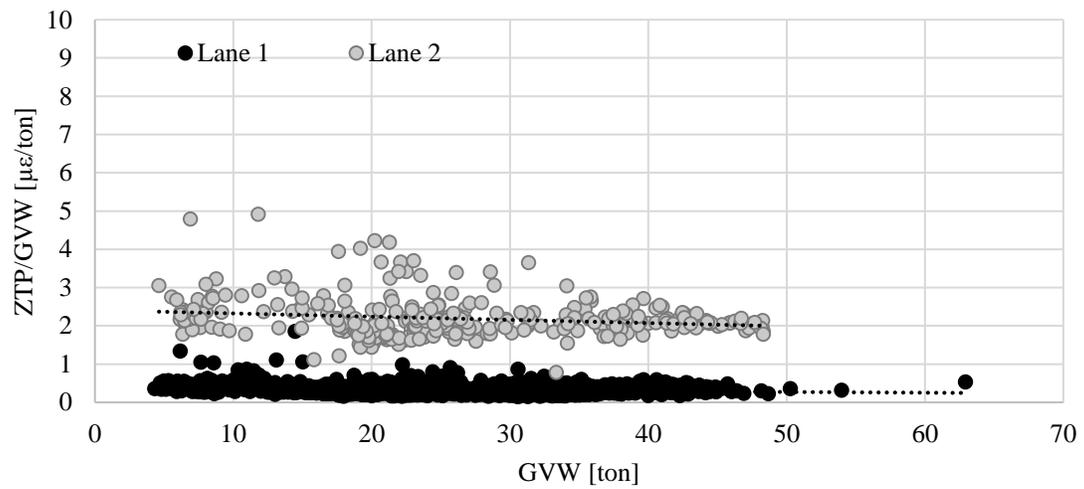
Figure 2.2: scatterplots of ZTP vs GVW, by girder (March 2011).

Measured ZTPs were normalized by the GVW to remove the linear relationship described in the previous section. Fig.2.3 shows how, for each girder, the normalized response of the bridge is approximately constant. The scatter is attributable to the large number of trucks with different axle configurations, as well as external factors, including the position of the truck within the lane. It is noteworthy that the scatter is wider on the more heavily loaded girders (Girders 2, 3 and 4) but it is slightly reduced when the trucks are significantly heavy.

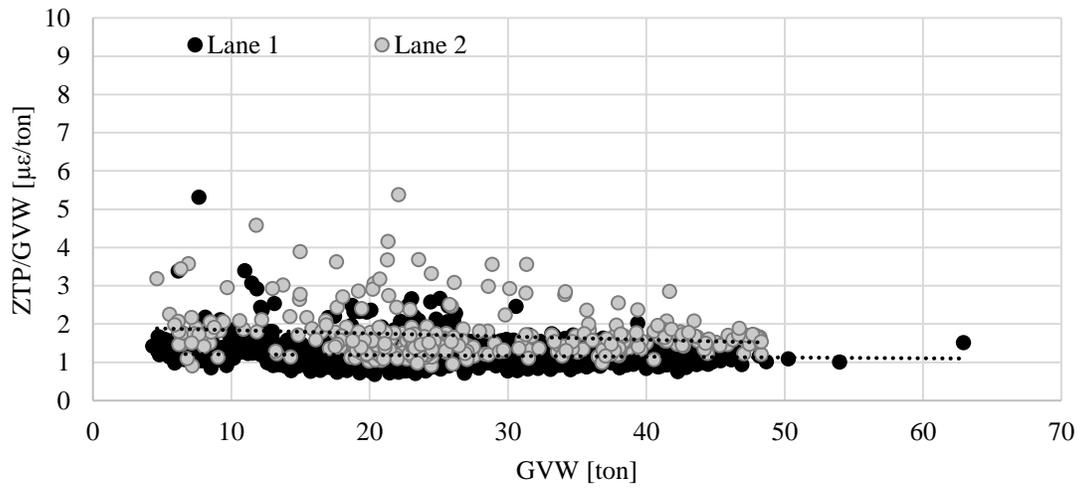
GIRDER 1



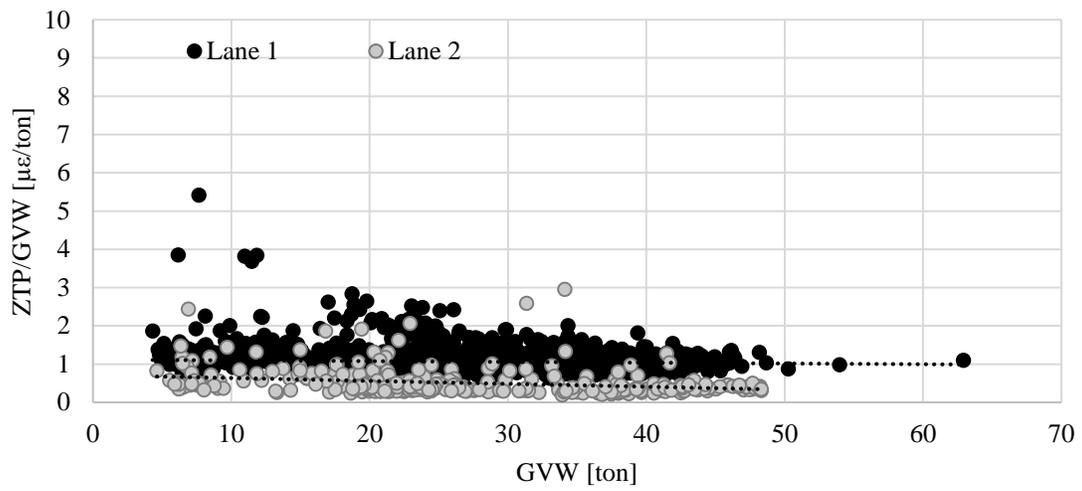
GIRDER 2



GIRDER 3



GIRDER 4



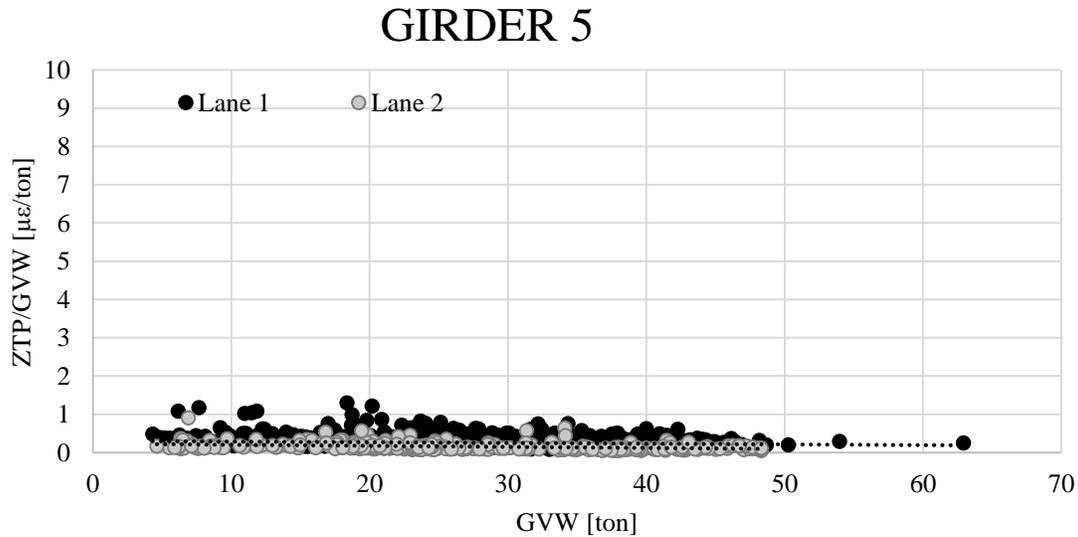


Figure 2.3: normalized ZTP vs GVW, by girder (March 2011).

2.3 GIRDER DISTRIBUTION FACTOR (GDF)

2.3.1 AASHTO Design GDF

The calculated GDFs, which describe the structural behavior of the bridge, have been compared with the GDF formulation for design provided by AASHTO LRFD specification. AASHTO GDFs identify the percentage of the design maximum live load that each girder will carry (AASHTO 2007). The following formulation is valid for the concrete deck, supported by steel beams, for moment in interior girders.

One design lane loaded:

$$\text{AASHTO}_{\text{GDF}_1} = \left[0.06 + \left(\frac{S}{14} \right)^{0.4} \cdot \left(\frac{S}{L} \right)^{0.3} \cdot \left(\frac{K_g}{12 \cdot L \cdot t_s^3} \right)^{0.1} \right] = 0.44 \quad (2)$$

Two or more design lanes loaded:

$$\text{AASHTO}_{\text{GDF}_I} = \left[0.075 + \left(\frac{S}{9.5}\right)^{0.6} \cdot \left(\frac{S}{L}\right)^{0.2} \cdot \left(\frac{K_g}{12 \cdot L \cdot t_s^3}\right)^{0.1} \right] = 0.6 \quad (3)$$

Eq.3 is chosen due to its more conservative value. Since the bridge has skewed supports the code suggests a reduction of the load distribution:

$$\text{AASHTO}_{\text{GDF}_{I_s}} = \text{AASHTO}_{\text{GDF}_I} \cdot [1 - c_1(\tan\theta)^{1.5}] = 0.54 \quad (4)$$

The variables in the equation are explained below:

S = Girder Spacing [ft];

L = Span [ft];

t = Deck Depth [in];

$K_g = n \cdot (I + A \cdot e_g)$ = Longitudinal Stiffness parameter [in⁴];

$n = \frac{E_B}{E_D}$ = Modular Ratio;

E_B = Modulus of Elasticity of Beam Material [ksi];

E_D = Modulus of Elasticity of Deck Material [ksi];

I = Moment of Inertia [in⁴];

A = Cross-Sectional Area of the Girder [in²];

e_g = Distance between the centers of gravity of the basic girder and the deck [in];

θ = Angle of Skewness, for $30^\circ \leq \theta \leq 60^\circ$;

$c_1 = 0.25 \cdot \left(\frac{K_g}{12 \cdot L \cdot t_s^3}\right)^{0.25} \cdot \left(\frac{S}{L}\right)^{0.5}$, for $\theta < 30^\circ$ then $c_1 = 0$.

2.3.2 Experimental GDF

As mentioned in the literature review, the formula developed by Ghosn et al. (Eq.5), (which, in this case, study, coincides with the formulation of Stallings and Yoo because all girders have the same section modulus, $w_i=1$), is used here and it is rewritten in Eq.6.

$$GDF_i = \frac{\varepsilon_i \cdot w_i}{\sum_{j=1}^{\#girders} \varepsilon_j \cdot w_j} \quad (5)$$

Given the ($\mu\varepsilon$ - t) waveform, obtained by recording the strain ($\mu\varepsilon$) for five seconds (t), and the calculated ZTP for each girder, the ZTP_{GDFs} are calculated, for each truck crossing (i), as the ratio between the ZTP at girder j and the sum of all girder ZTPs.

$$ZTP_{GDF_{i,j}} = \frac{ZTP_{i,j}}{\sum_{j=1}^5 ZTP_{i,j}} \quad (6)$$

The ZTP_{GDFs} can be used as a reliable method for lane classification. It is important to notice that the WIM system is installed 60 m prior to entering the bridge, meaning that the driver can potentially change lanes resulting in an incorrect lane record. In order to avoid lane mismatch, the GDF of the two most representative girders has been used. More load on girder 4 indicates that the truck is in lane 1, otherwise, the truck is clearly in lane 2, as explained by Eq.7-8.

$$ZTP_{GDF_{girder\ 4}}(i) > ZTP_{GDF_{girder\ 2}}(i) \rightarrow \text{Lane 1} \quad (7)$$

$$ZTP_{GDF_{girder\ 4}}(i) < ZTP_{GDF_{girder\ 2}}(i) \rightarrow \text{Lane 2} \quad (8)$$

This method has been validated by using the photos taken by the WIM system cameras and has consequently been used to confirm WIM lane classifications. The rate of WIM misclassification is not very concerning because out of 31,107 trucks crossings only 462 (1.5%) are misclassified, among which, 285 (62%) trucks crossing on lane 2 were detected as lane 1 and 177 (38%) trucks crossing on lane 1 were identified as lane 2. WIM misclassification is not necessarily due to a malfunction of the system, but it can be due to a change of lane after passing the WIM system, due to roadworks which signal to use the other lane or due to the need to overcome a vehicle. Numbers show that the WIM misclassification rate is low, however, a consistent incorrect lane classification

can raise some concerns due to a potential disruption of the pavement which force the drivers to change lane.

2.4 RESULTS

The overall average of ZTP_{GDF} was calculated for all trucks over the three-year period for each girder, for each lane (Fig.2.4). As expected Lane 1 and Lane 2 distributions are symmetric and slightly skewed due to the asymmetric shoulders. Since Girder 3 is in the center, it always plays an important role in carrying both lane's live loads. GDF averages never exceed the AASHTO design values (Fig.2.4), therefore, it may be concluded that the AASHTO specifications are appropriate for this bridge. This outcome is also significant in relation to the age of the bridge. It was built in 1960, and it was subsequently subjected to deck replacement in 1987, i.e. much earlier than when the code was released, meaning that the AASHTO parameters are also valid for aging bridges. For completeness, the AASHTO GDF for non-skewed bridges (dotted line; Fig.2.4) is also presented in the graph, in addition to the value for skewed bridges (solid line; Fig.2.4), because 30° is the border line limit to apply the formula for skewed bridges. To validate the distribution obtained, the results of a similar study conducted by Tennyson et al. (2001) is shown in Fig.2.5. The values of the graphs shown in Fig.2.4 are presented in percentages in Fig.2.6.

The long-term data recorded at Hurley were used to check the AASHTO specifications also against extraordinary events (crossing of heavy machinery, snow removal, house moving, etc.). For this purpose, the maximum values of GDF were calculated and plotted for both lanes (Fig.2.7-2.8). As expected, Girders 2, 3 and 4 present the highest values that almost reach the design value suggested by AASHTO, however, they never exceed them. This indicates that the

AASHTO_{GDFs} are valid for both the average truck and the exceptional event, and not excessively conservative as other studies suggest (section 2.4.1).

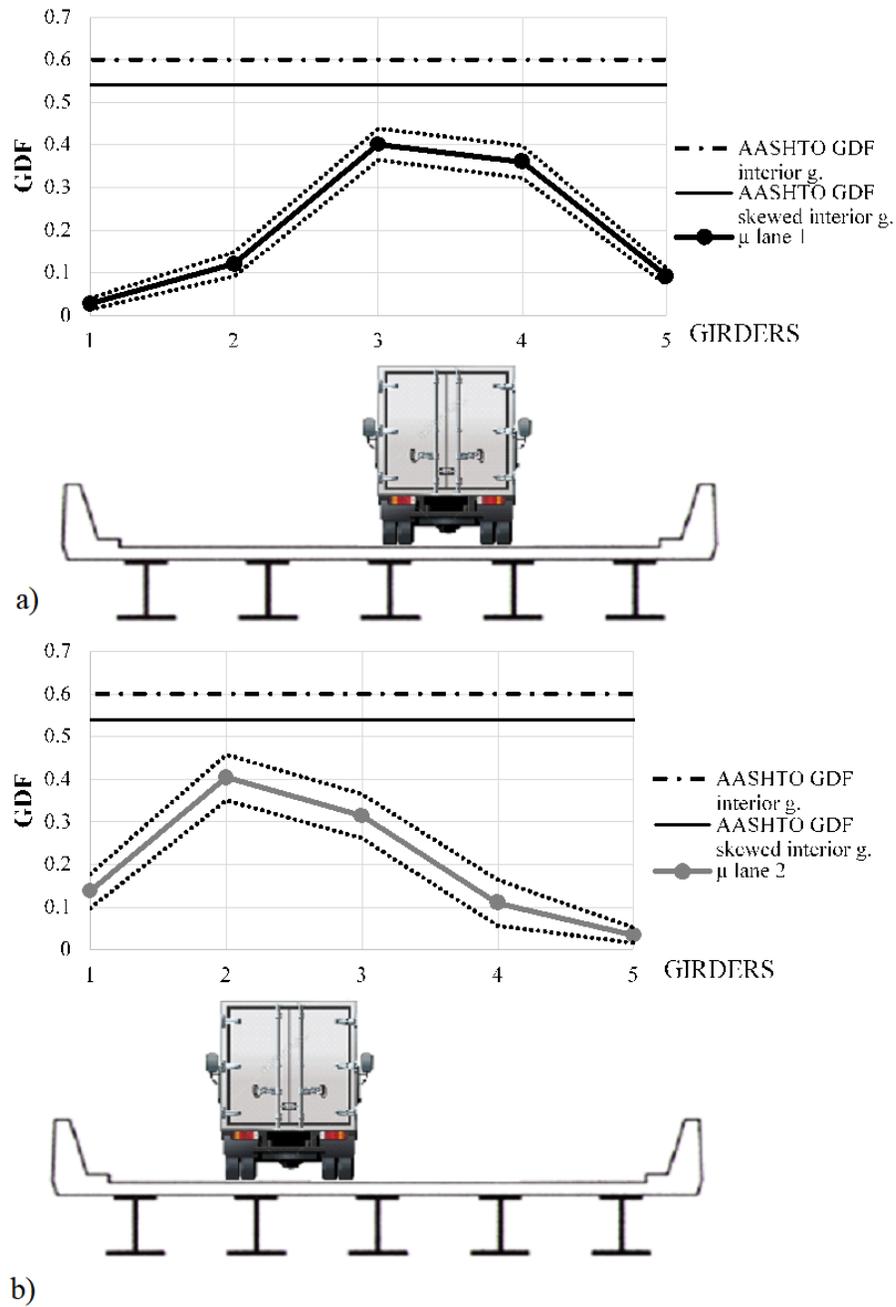


Figure 2.4: AASHTO_{GDF}, 3-year averages (μ) of ZTP_{GDFs}, \pm std. dev. a) lane 1; b) lane 2.

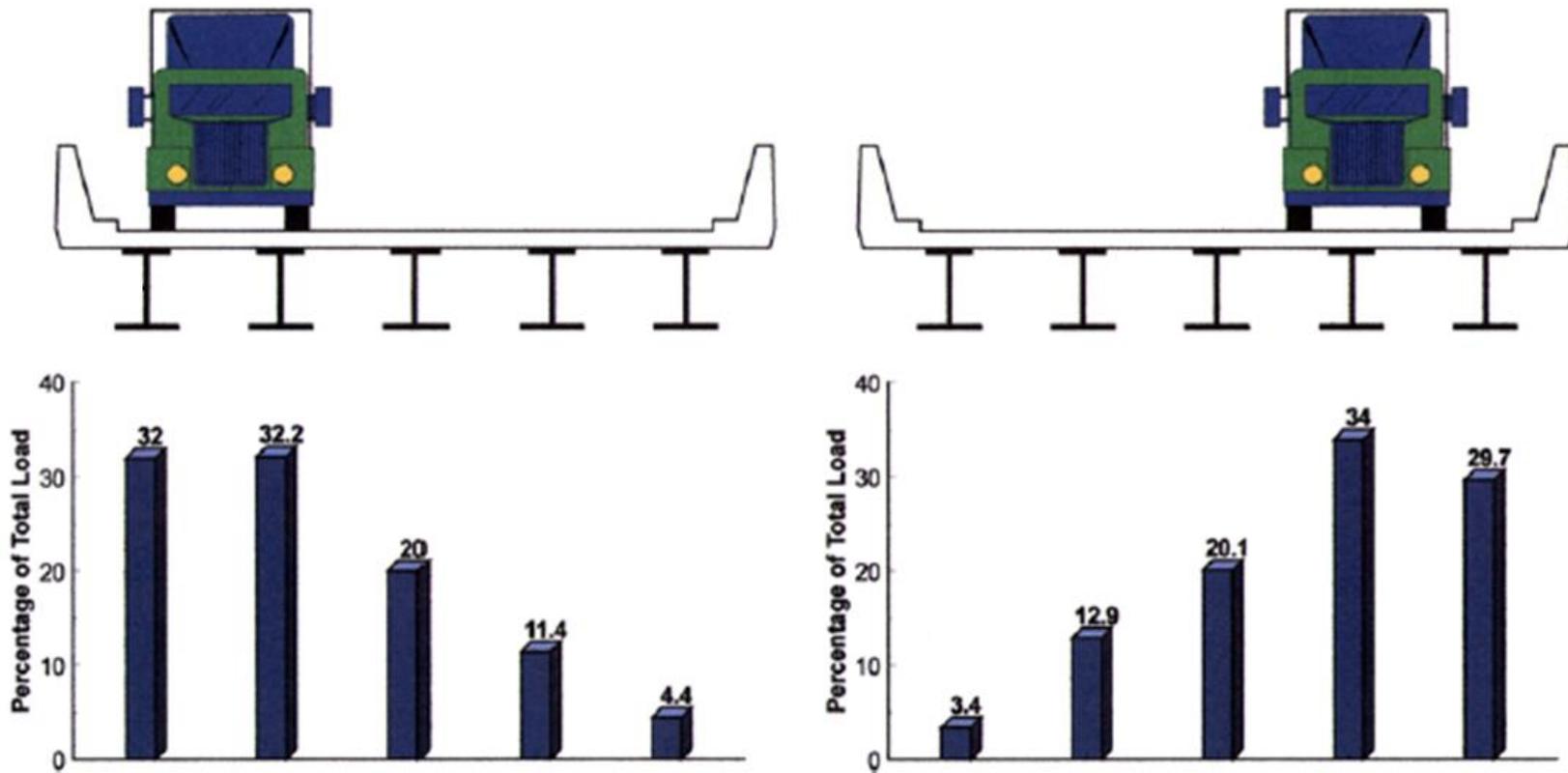


Figure 2.5: “Load Sharing among girders based on dynamic strain measurements 1997” (Tennyson et al., 2001).

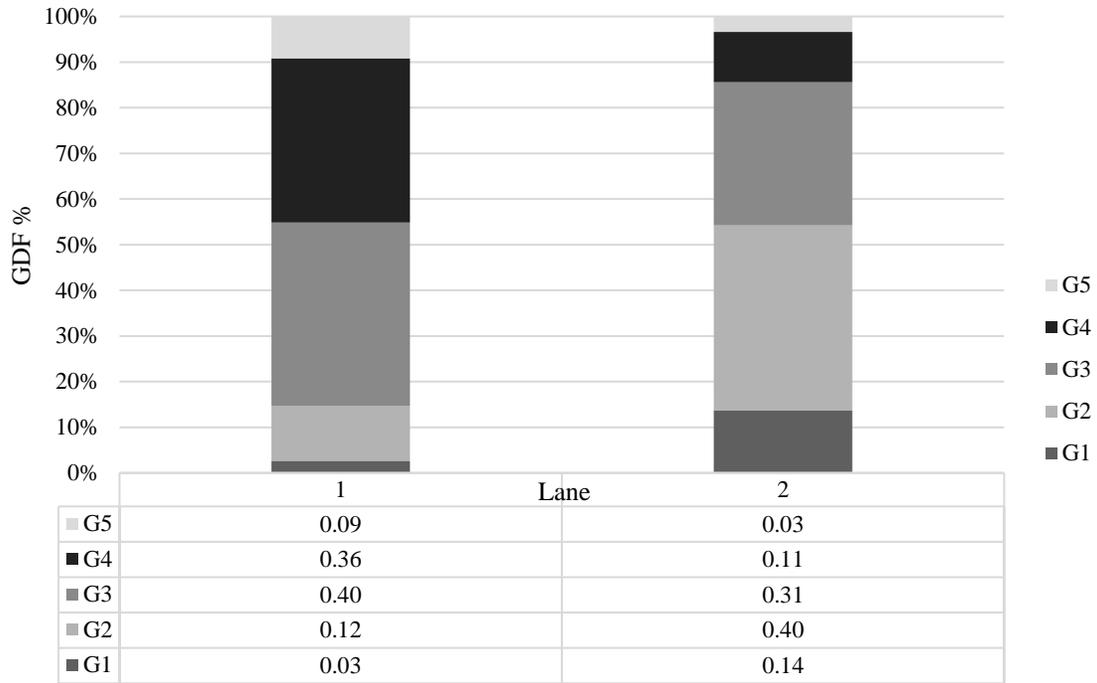


Figure 2.6: load percentages carried by each girder, by lane.

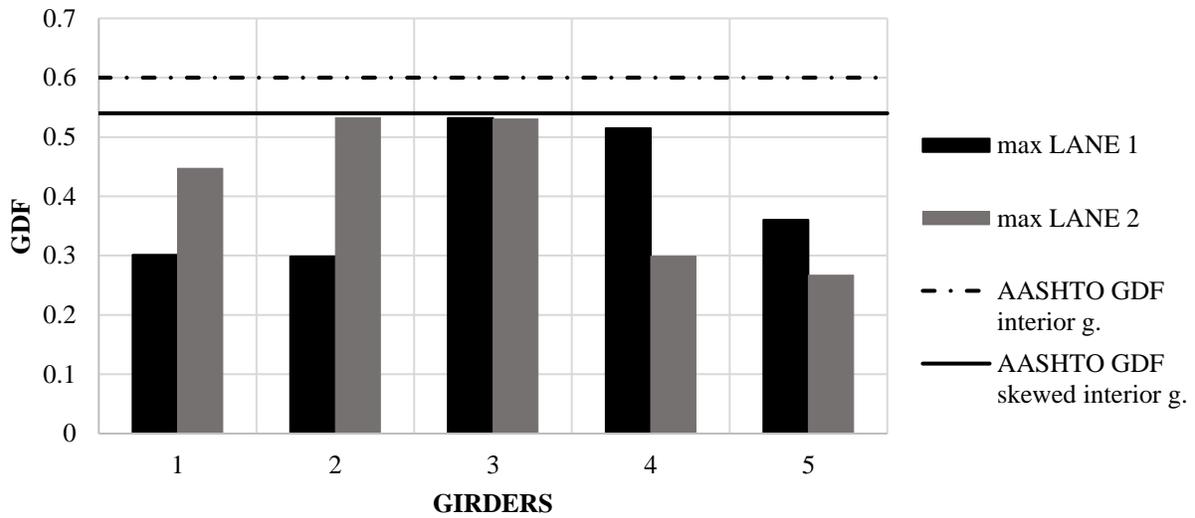
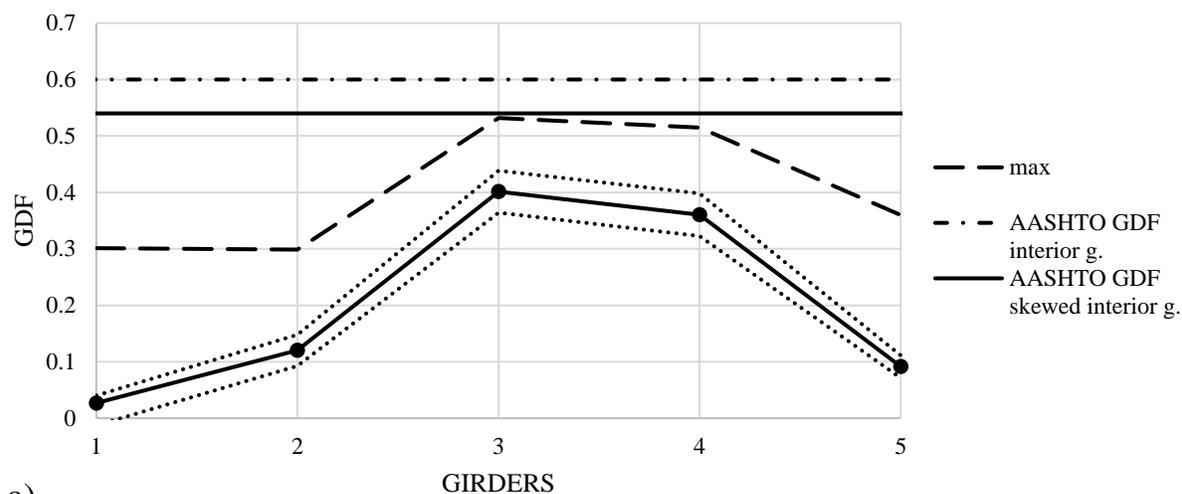
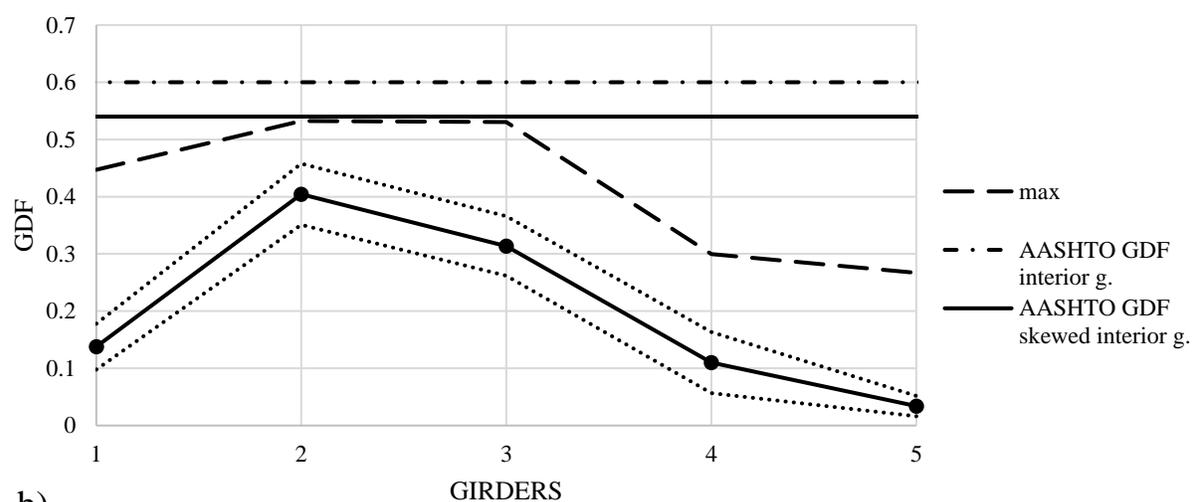


Figure 2.7: AASHTO_{GDF} and calculated ZTP_{GDF}s maxima.



a)



b)

Figure 2.8: 3-year averages, maximum values and AASHTO specifications. a) lane 1; b) lane 2.

2.4.1 AASHTO Values in Other Studies

The results of some related studies concerning the AASHTO values are presented here. Some studies show that the AASHTO specifications may be over-conservative. Kim and Novak (1997) stated that GDFs from specifications are much larger than measured GDFs. In the cases of two bridges, they found that the $AASHTO_{GDFs}$ are respectively 28% and 19% larger than the measured GDFs. Novak et al. (2003) showed that for simply-supported, single short-span bridges,

with two traffic lanes the AASHTO specifications are appropriate. Olund and DeWolf (2007) stated that their long-term bridge monitoring program showed that the live load distribution, calculated for each of the eight girders during crossings of two trucks with known GVW, was approximately 42% lower than the AASHTO specifications. Brendler and Yasser (2015) stated that the AASHTO specifications were reasonable for one loaded lane but overly conservative for multiple loaded lanes. They also stated that the code is more conservative for simply supported bridges than equivalent integral abutment bridges.

2.4.2 Influence of Truck's Characteristics

Gross vehicle weight, overall length, and speed of the truck are analyzed to investigate their influence on the calculated ZTP_{GDFS} . To ensure that the recorded data do not contain operational errors, the following checks were applied, (Dai, 2013): $GVW > 0$, $15 \text{ km/h} < \text{Speed} < 160 \text{ km/h}$, $0 < \text{Length} < 60 \text{ m}$. For trucks in Lane 1, Girder 3 GDFs were plotted against the truck's characteristics. Fig.2.10-2.11-2.12 suggests that heavier trucks, including overweight ones, show a smaller range of GDF values because larger GVW trucks produce larger strain values while fluctuations due to bridge vibration or other factors will not increase with truck GVW. Trucks with a length between 18 m and 24 m (60-80 ft) are closer to the design limit. Speed does not seem to be influential. Ultimately, the position of the truck on the bridge, in terms of the lane and the driver's tendency to drive toward the centerline or towards the shoulders, are the most important parameters that influence the distribution of the load on girders.

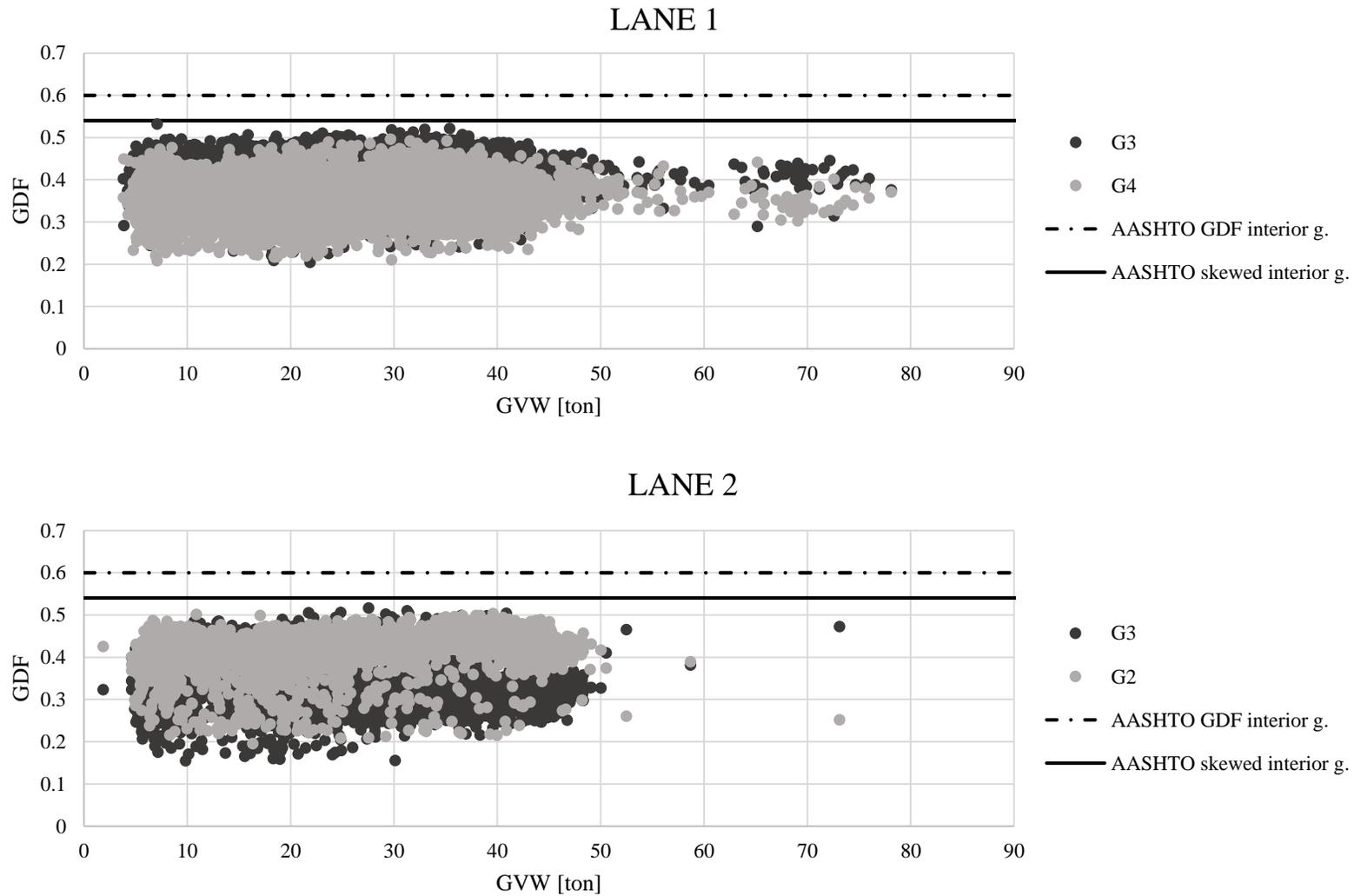


Figure 2.9: ZTP_{GDF} vs GVW, girder 3 and 4, lane 1; girder 2 and 3, lane 2.

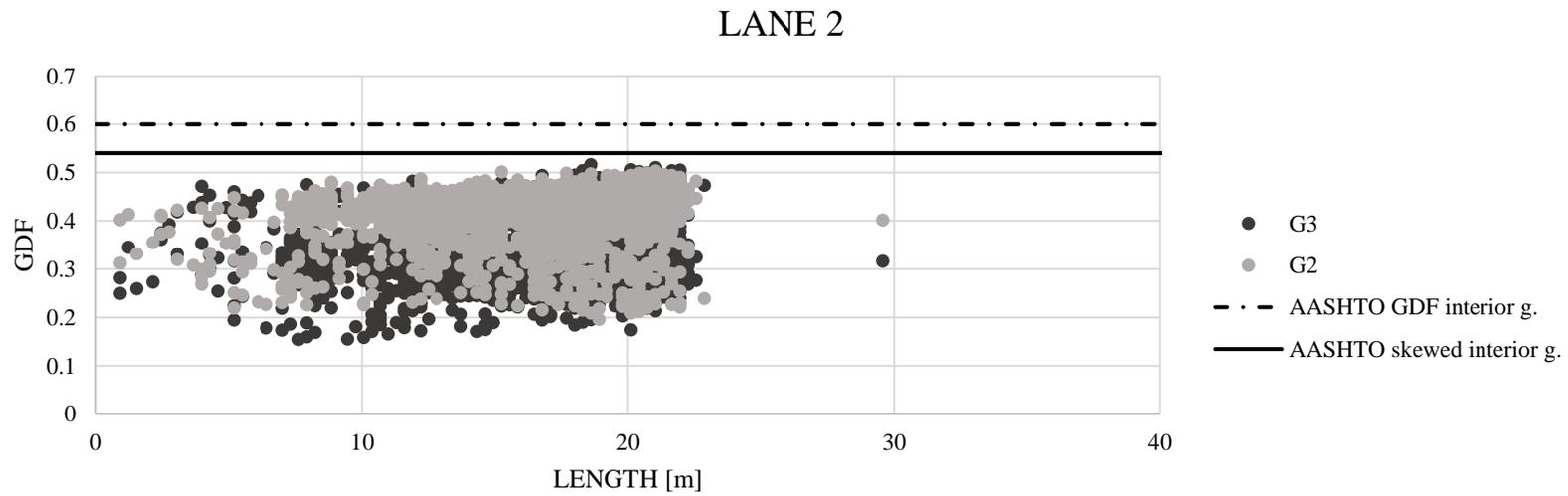
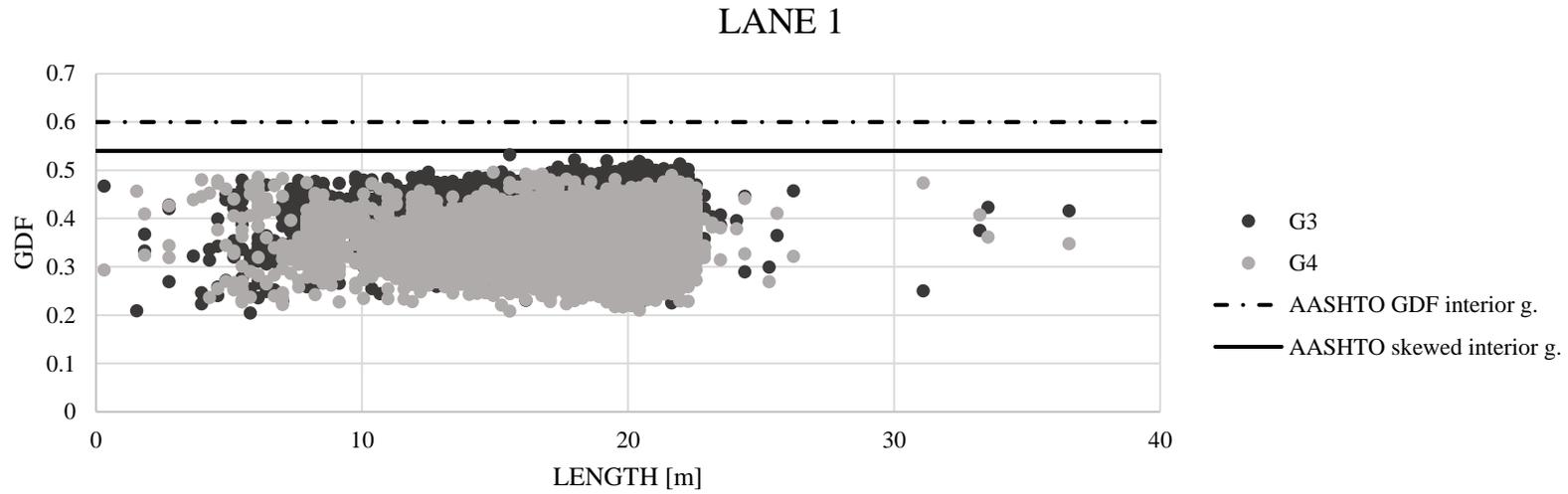


Figure 2.10: ZTP_{GDF} vs length, girder 3 and 4, lane 1; girder 2 and 3, lane 2.

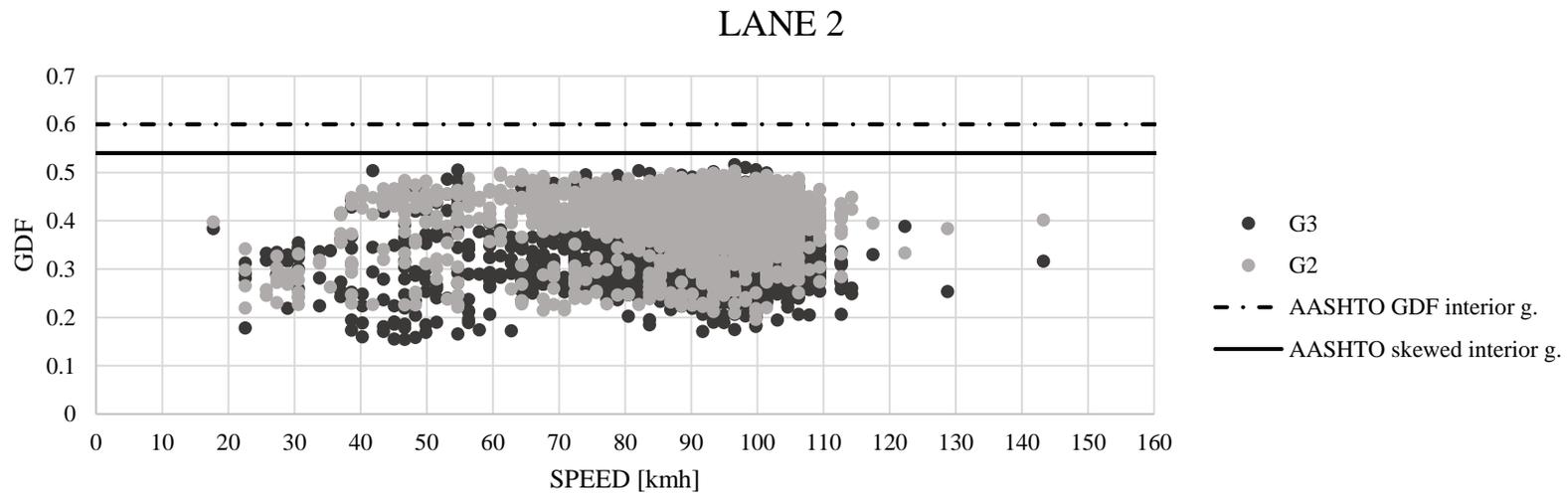
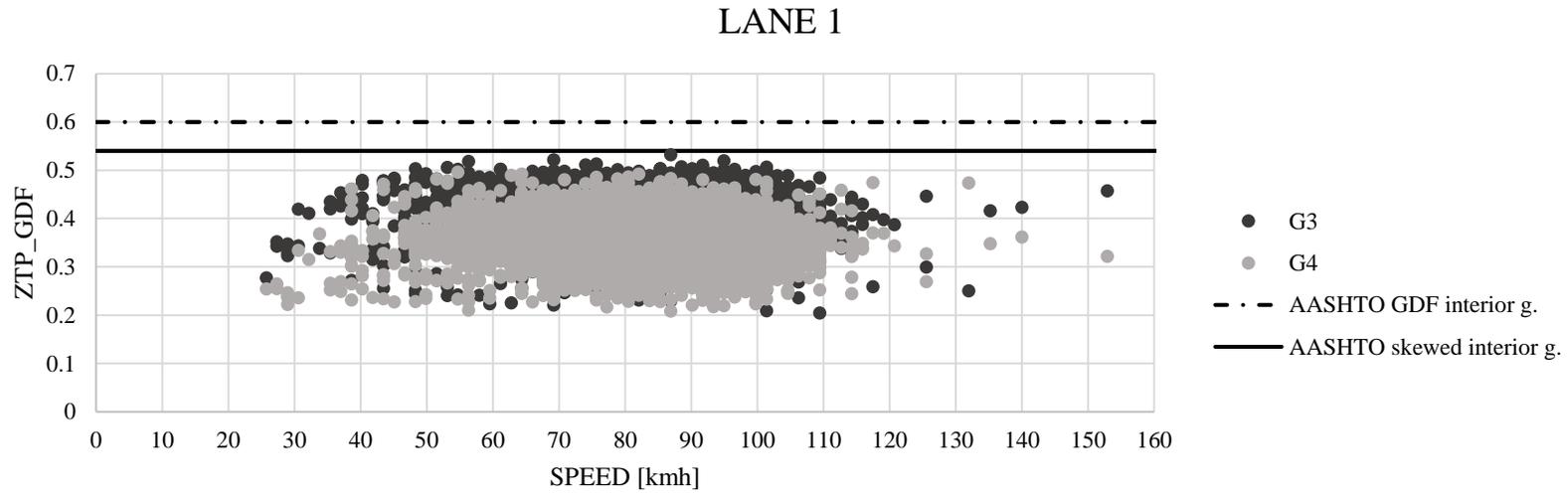


Figure 2.11: ZTP_{GDF} vs speed, girder 3 and 4, lane 1; girder 2 and 3, lane 2.

2.5 LONG-TERM AGGREGATE ANALYSES

A long-term aggregate analysis is developed to detect potential trends, changes, sudden anomalies and seasonal events, over time. Daily, weekly and monthly averages of ZTP_{GDF} for each girder and lane are calculated. Figure 2.12 shows one day of truck crossings (February 20, 2010), as expected, the frequency of trucks starts to increase after 7 am and it is much less during the night. A common feature of all following plots is that lane 2 presents much more scatter and a much lower number of trucks, presumably because heavy vehicles tend to drive in the slow lane (lane 1).

The first level of aggregation is daily average meaning that Fig.2.12 a-b are represented in Fig.2.13 and 2.14 as five data-points, one for each girder. In Fig.2.13 girder 3 presents upward trends towards the end of each year and downward trends with the beginning of the new year, while girder 2 and 5 reasonably show the opposite behavior. More scatter is also noticeable during the same period during the transition of all three years. The increase of scatter between years is even more noticeable in lane 2 (Fig.2.14). Fig.2.15 shows a typical week of truck crossings (February 20-26, 2010). The plots show a very clear separation between days, the presence of fewer trucks identifies the nights. The sharp gap on Thursday represents a malfunction of the system which was promptly repaired. In lane 2 (Fig. 2.15b), the separation between loaded girders (2 and 3) and less loaded (4 and 5) is less clear suggesting that trucks crossing the bridge on lane 2 tend to drive more toward the center, creating more compact clouds in the plots.

The second level of aggregation is the weekly average where changes are smoother but still appreciable. Fig.2.16 still shows the same temporary shift around the new year, already mentioned

for the daily average, of girder 3 and 4 between 2010 and 2011, and between 2011 and 2012 although less marked. Fig.2.17 shows the presence of scatter at the transition between years.

The third level of aggregation is the monthly average which is the final widest aggregation range that still allows potential changes to be captured. A further step of aggregation will cause significant loss of information. The trends in Fig.2.18 are much smoother, but it is still possible to detect a periodical behavior (or seasonality) with different intensity between the years, with a more accentuated trend in lane 2 (Fig. 2.19). Girder 3 shows a permanent shift in the GDF values, starting January 2011.

Some of the changes highlighted by the plots can be linked to some events or decisions. Three types of events were considered potential influencers of the bridge behavior over the 3-year period and they are logging season, weather conditions, and Wisconsin statutory change.

- The logging season generally begins after the first frost in October, when sap begins to run and ends in March when it becomes difficult to drive in thawing soil. Monday through Friday hauling is common, while a few mills operate 24/7, Saturday operations are generally mulch or chips from logging during the weekdays. There is a general increase in loaded logging truck traffic during the winter months with peaking occurring mostly in December and February, while there is a significant decline of loaded logging trucks in March, this can be the most plausible cause of the consistent trend between the end of the year and the beginning of the new one. Due to the periodic nature of the phenomenon visible in all plots, with different levels of detail, and its persistence in each year, it can also be described as seasonality.

- There is a specific event which might have impacted the truck traffic. The event is the Wisconsin statutory change, effective from 1 January 2011, which allowed vehicle combination up to 98 kips on six axles to transport loads of raw forest products during the spring thaw suspension period (Owusu-Ababio and Schmitt, 2014). For completeness, for the zone 1, where the Hurley Bridge is located, the frozen road declaration was 84 days in 2010-2011 from December 17, 2010 to March 11, 2011 and 50 days in 2012, from January 19 to March 9 while the spring weight restriction was 83 days in 2011, from March 14 to May 14, and 35 days in 2012 from March 14 to April 18. This decision might have had a role in the permanent shift in the GDF starting January 2011.
- Another important event which might have contributed to the permanent shift after January 2011 is the severe winter. The Hurley-Ironwood area in Iron County experienced the most snow of 167 inches in the winter of 2010-2011, including the exceptional Groundhog Day blizzard of February 2, 2011 (emergencymanagement.wi.gov). This might have influenced the accentuated behavior of the most loaded girders (upward, then downward) visible at all three levels of aggregation. The whole bridge was supporting the additional snow load, vehicles might have been more loaded for safety reasons and snow removal vehicles were certainly in use during that event. Another extreme weather event happened during the weekend of December 11-12, 2010 where a so-called “monster storm” which left a foot or more of snow falling in Wisconsin and Minnesota (accuweather.com).

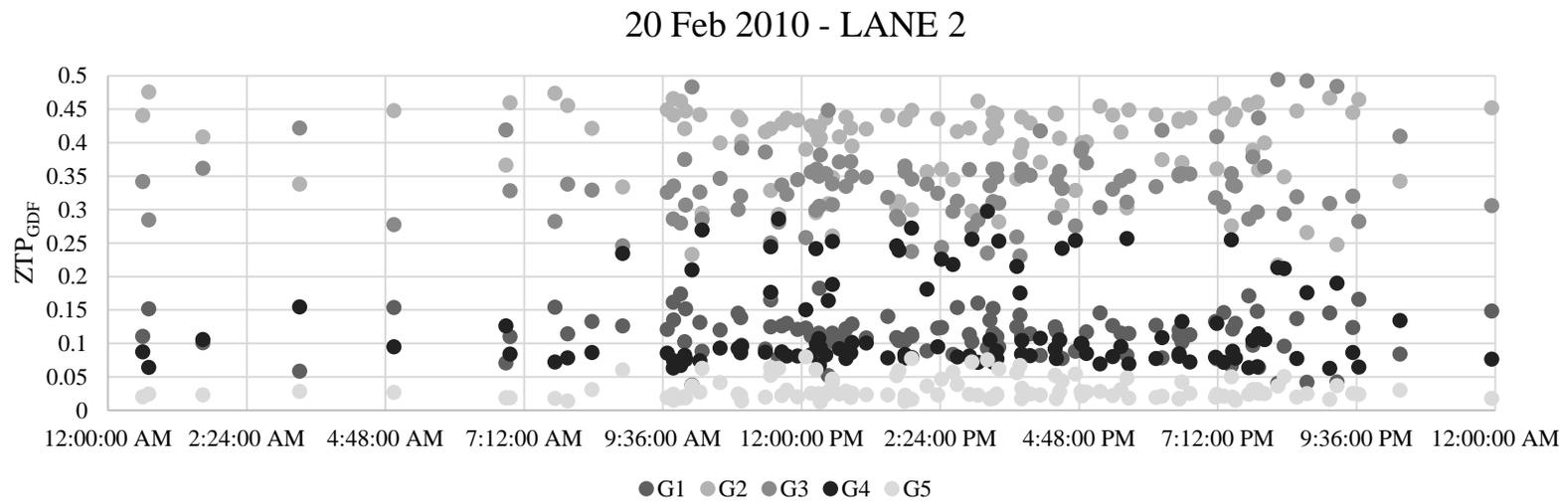
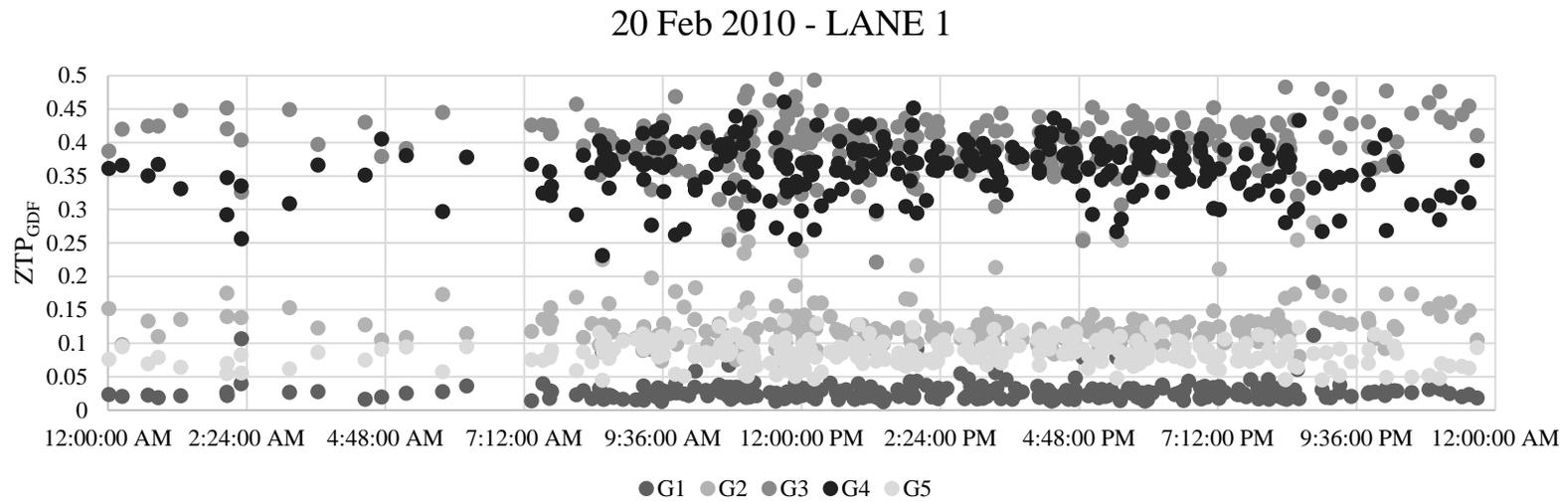


Figure 2.12: 1-day crossing, Saturday, 20 February 2010.

DAILY AVERAGE - LANE 1

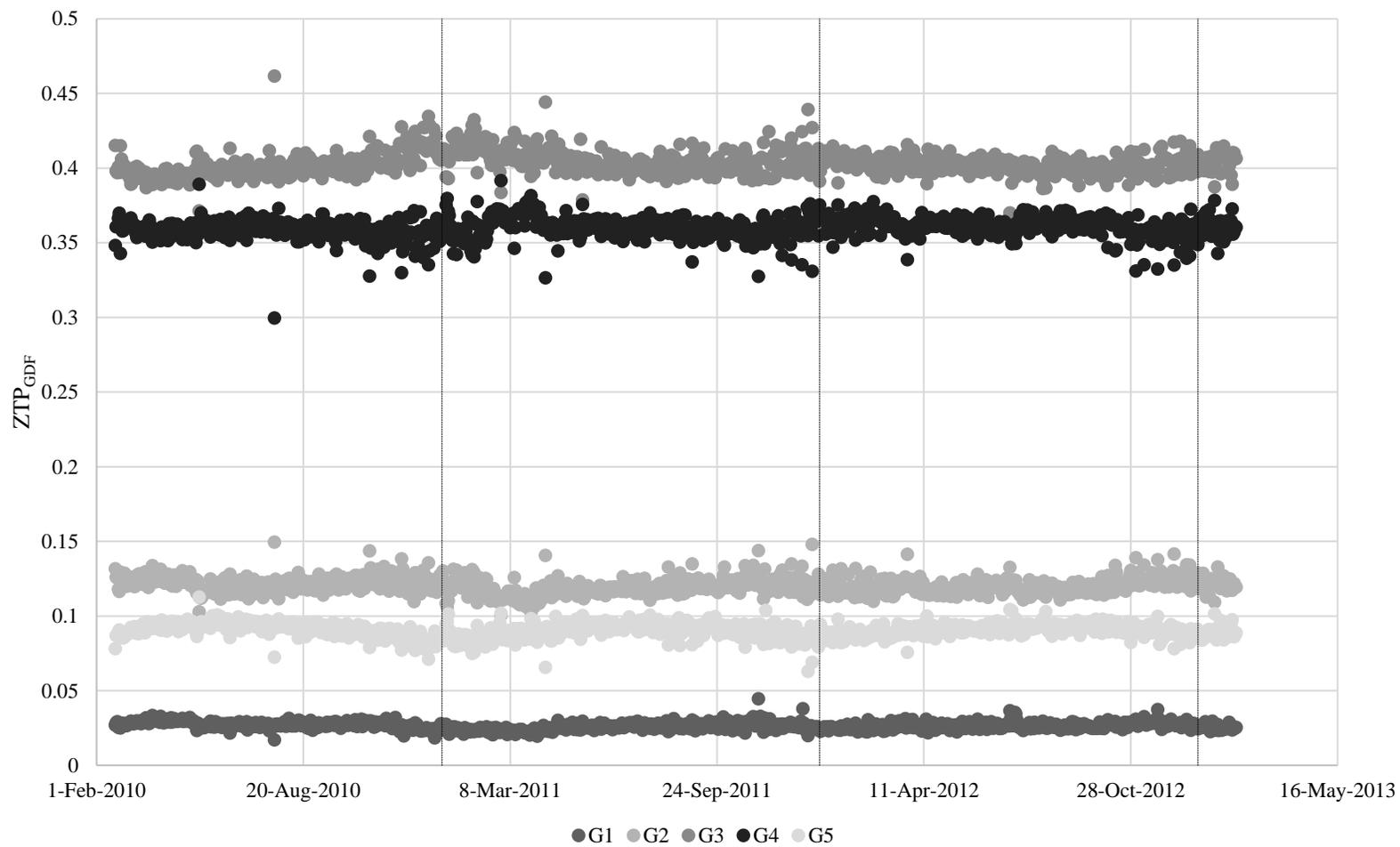


Figure 2.13: daily average, lane 1.

DAILY AVERAGE - LANE 2

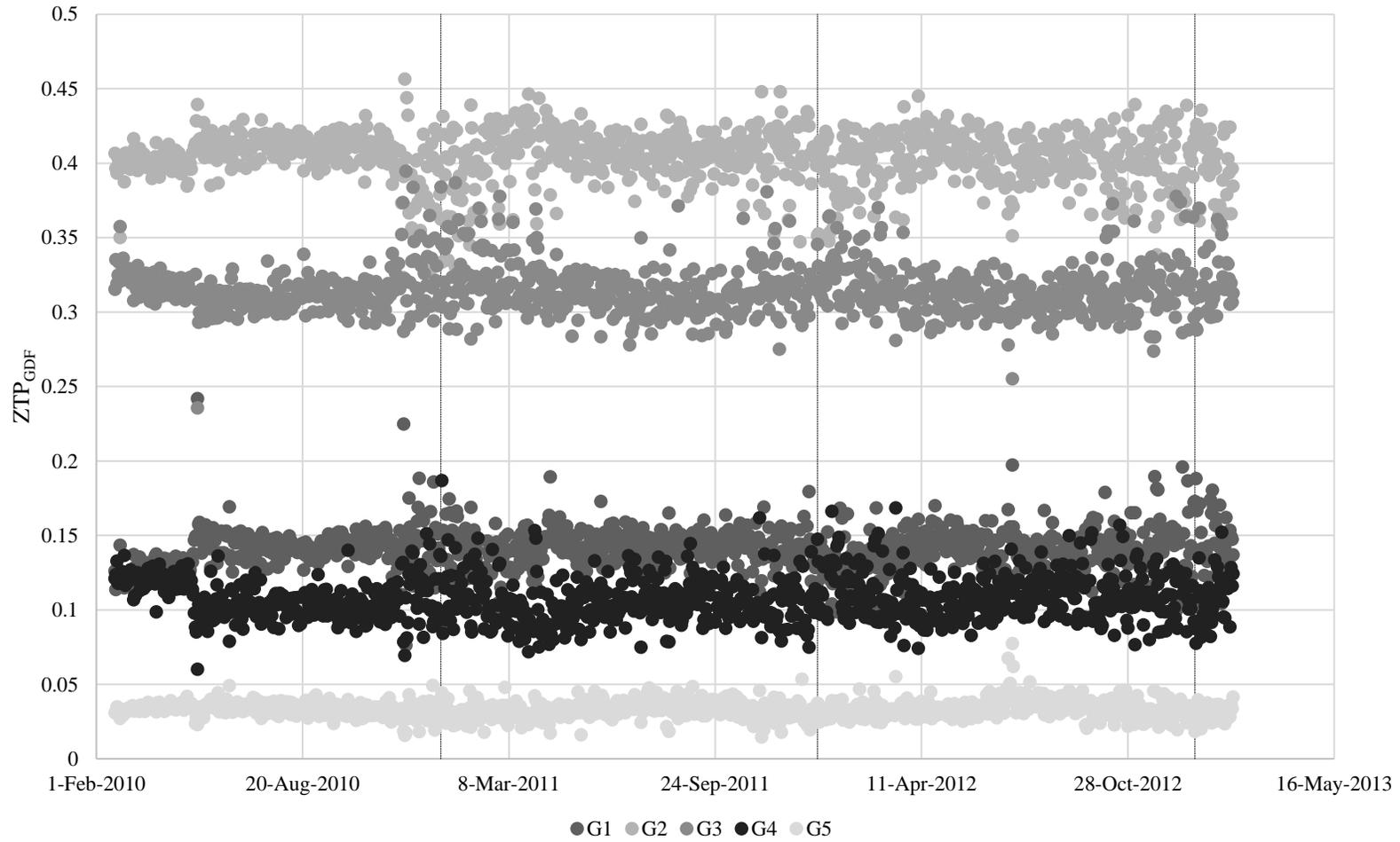


Figure 2.14: daily average, lane 2.

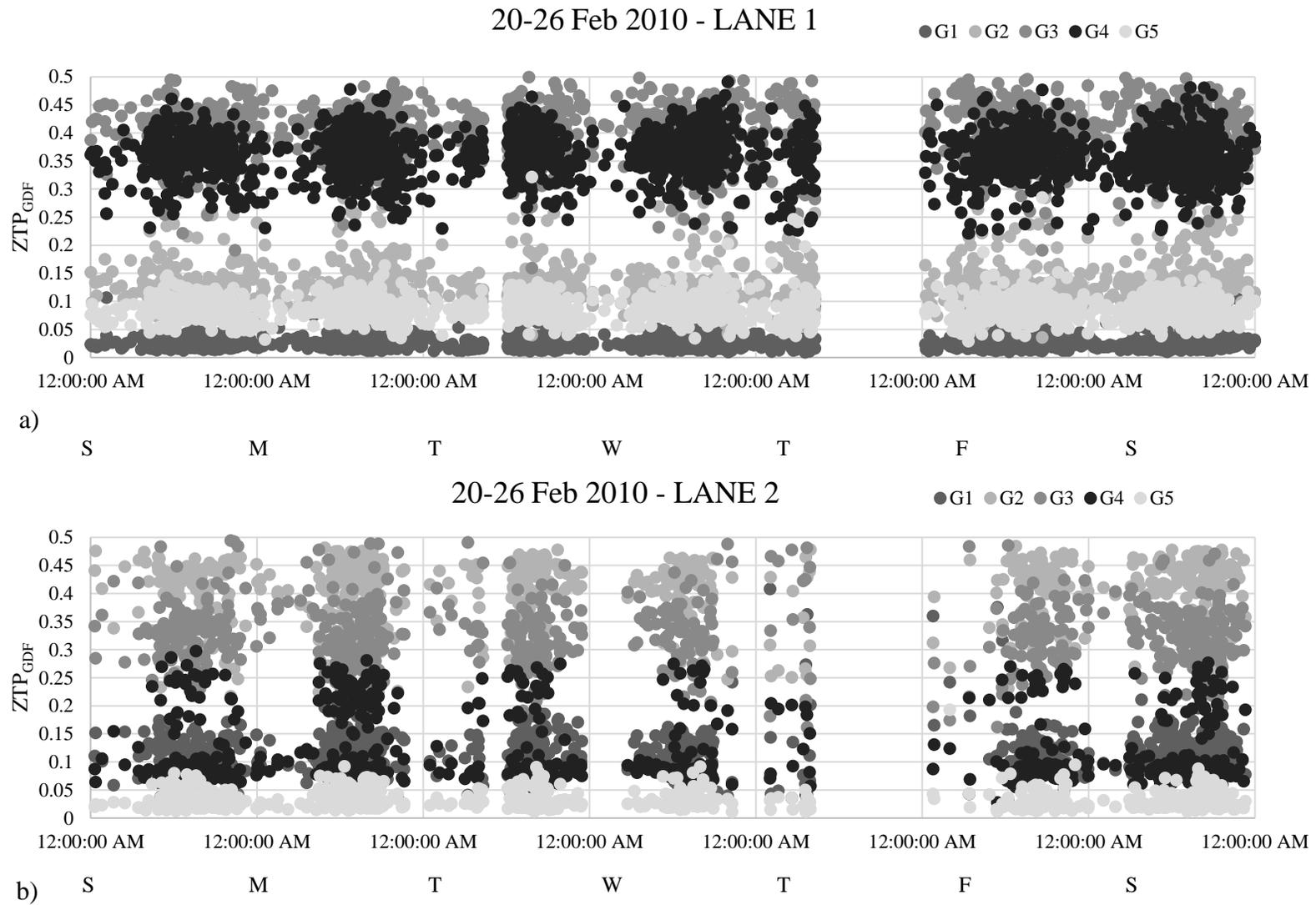


Figure 2.15: 1-week crossing, 20-26 February 2010.

WEEKLY AVERGAGE - LANE 1

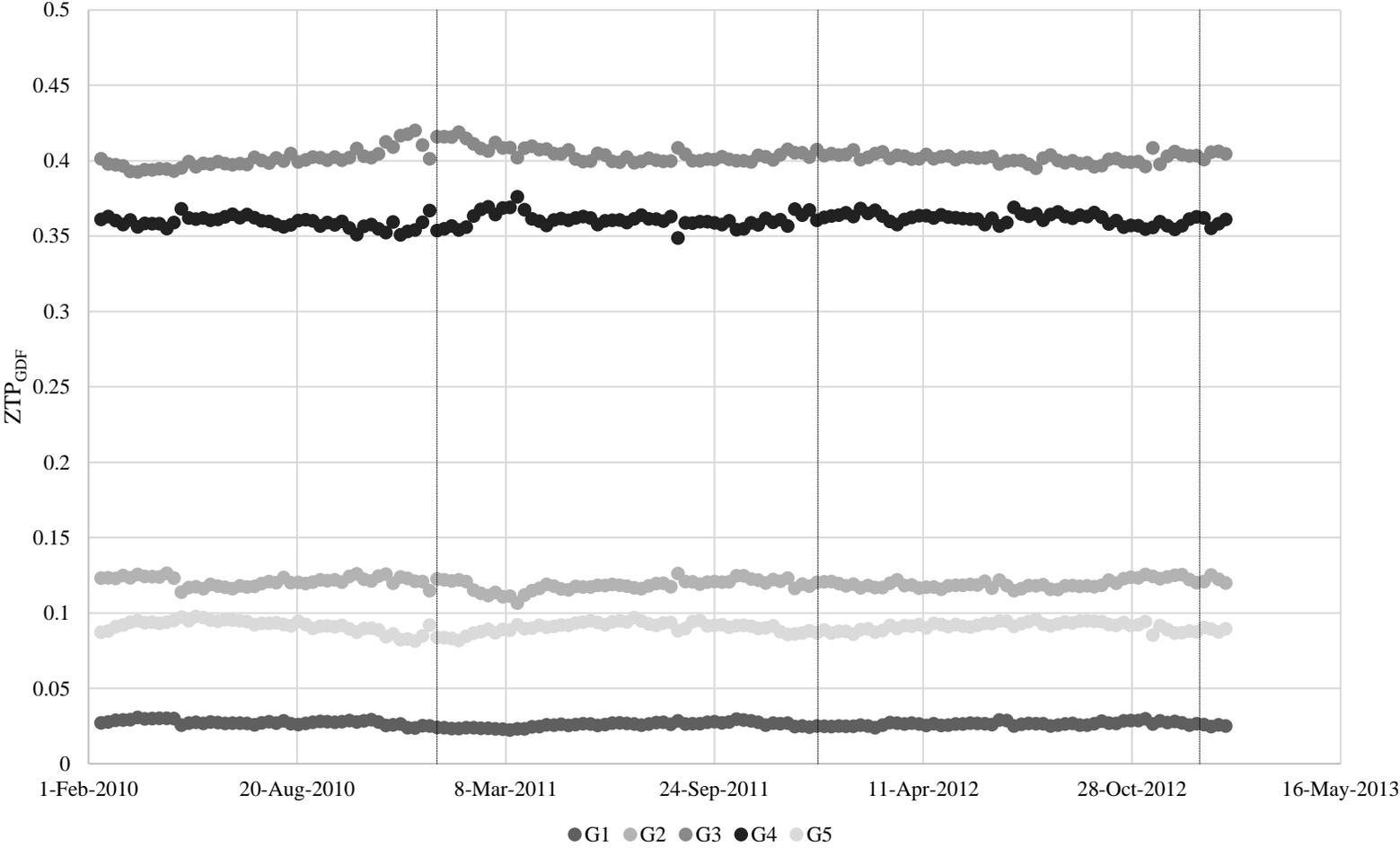


Figure 2.16: weekly average, lane 1.

WEEKLY AVERGAGE - LANE 2

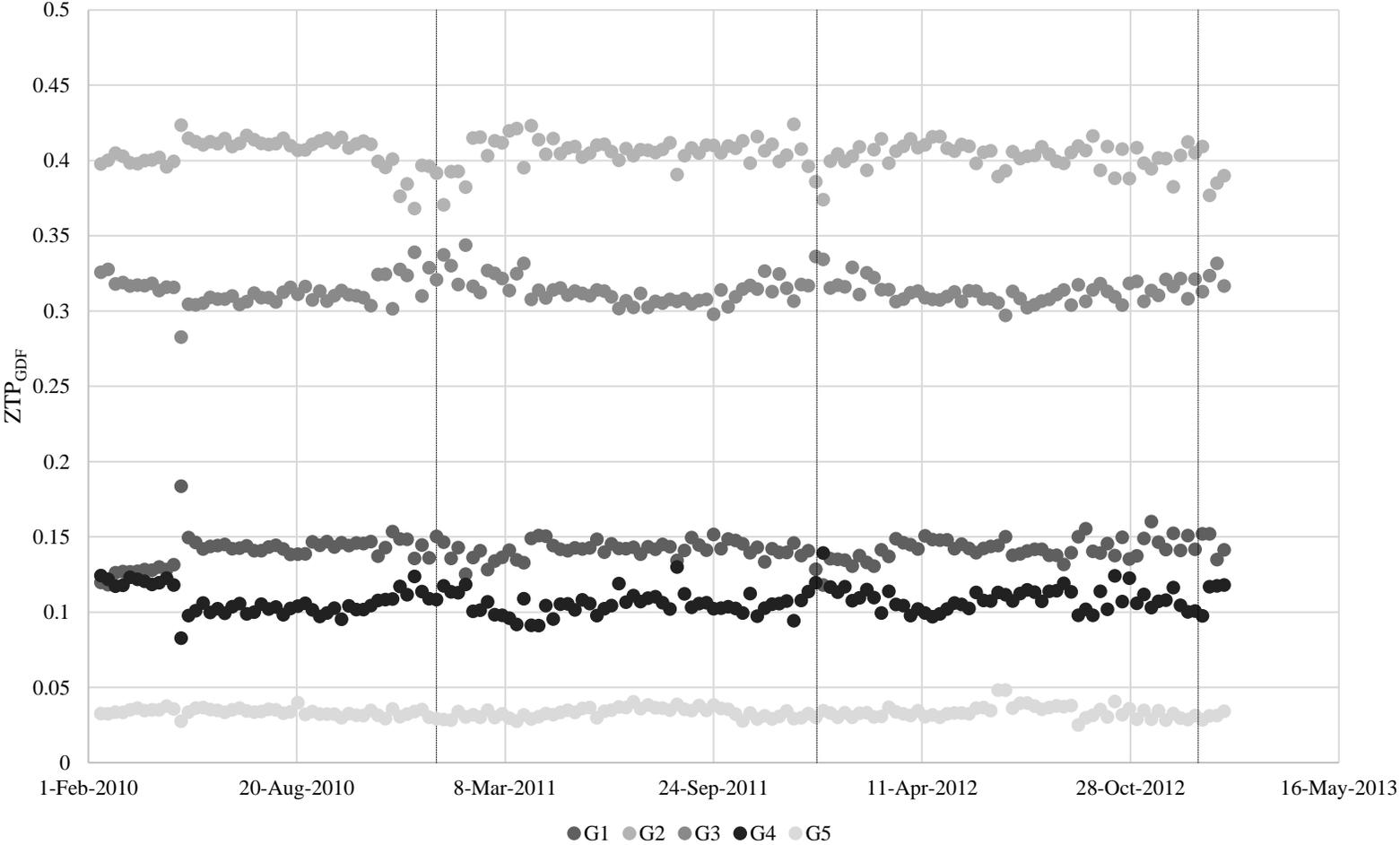


Figure 2.17: weekly average, lane 2.

LANE 1

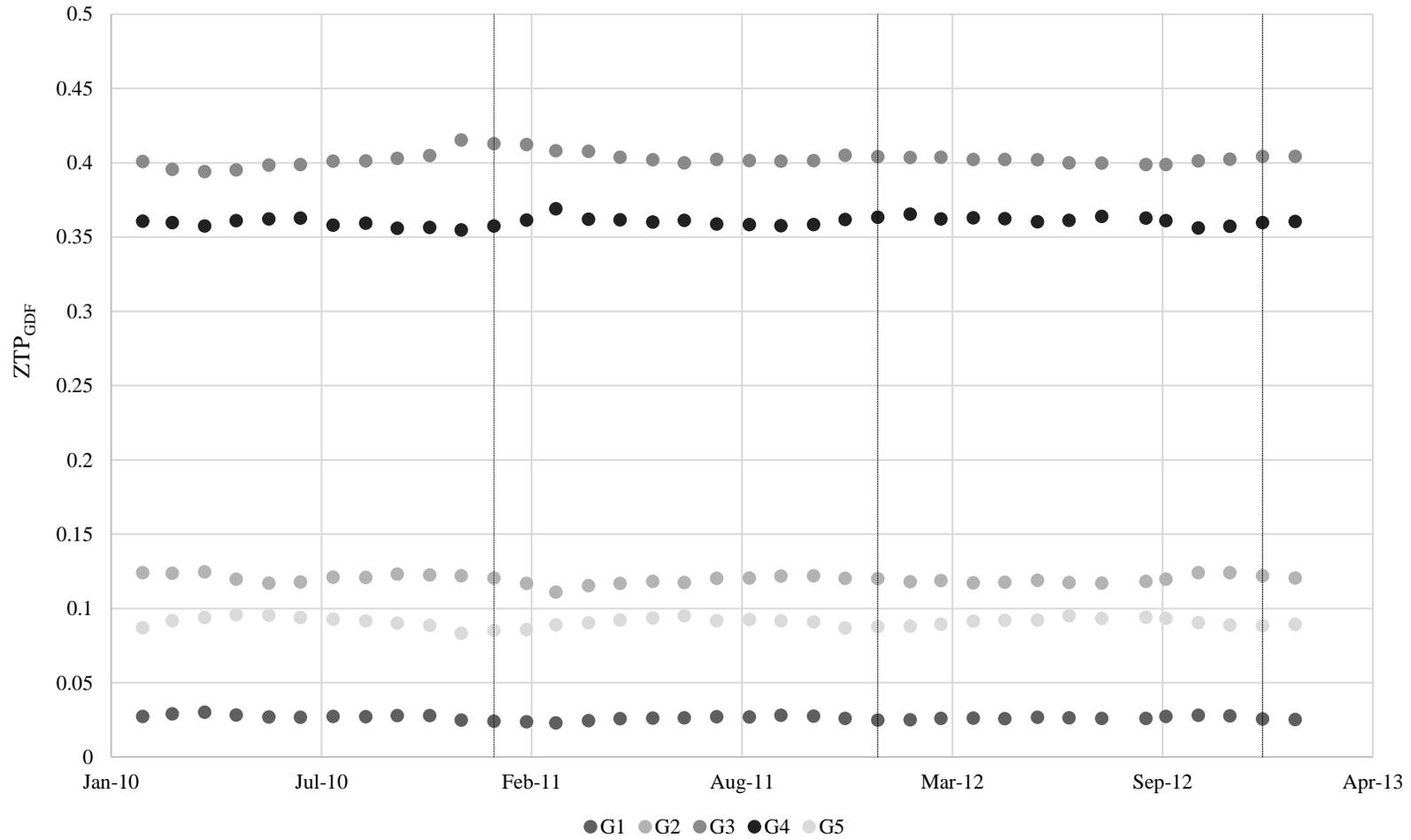


Figure 2.18: monthly average, lane 1.

LANE 2

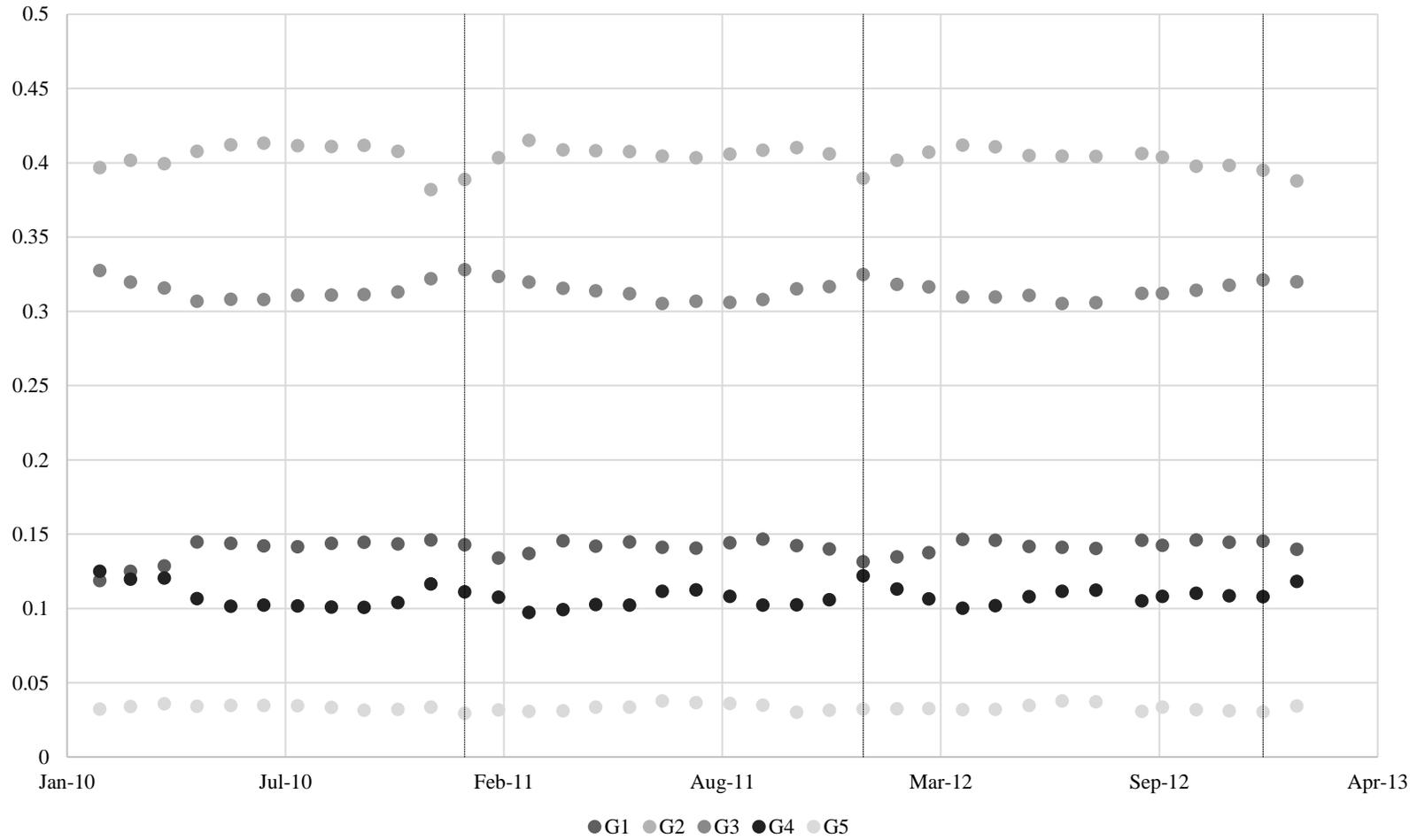


Figure 2.19: monthly average, lane 2.

2.6 CONCLUSIONS

In this chapter, a preliminary analysis is shown to monitor the performance of the structure using SHM data for over 171,000 trucks. GDFs are computed to study the structural behavior of the bridge under its current operational condition using ZTP data. The GDFs for over 171,000 trucks were compared to the AASHTO specifications and were found to be within specified limits. It is noteworthy that extreme events (with unusual high GDF) never exceed the AASHTO limits. This suggests that the AASHTO specifications are appropriate for this bridge. The GDF concept is also successfully used to provide a 100% error free lane classification method and to analyze the long-term behavior of the bridge at different levels of aggregation (daily, weekly and monthly average). A specific trend of GDFs is consistently noticeable at all levels of aggregation, towards the end of the year and the beginning of the new one, although with different levels of accuracy, this can be presumably explained with the beginning and end of the logging season (the most important economic activity of the area). The more accentuated GDF perturbations between 2010 and 2011 can be influenced by the extremely severe weather conditions of that specific winter with 168 inches of snow measured.

3 METRICS FOR BRIDGE PERFORMANCE

Prompted by the need for an efficient and effective use of monitoring data, the goal of this chapter is to propose novel metrics to monitor bridge performance. The first metric is a strain-distance-area measure which aims to monitor the bridge response to a wide range of truckloads. The second metric is a follow-up concept of the first metric and it is used to monitor the integrity of the bridge's component and materials over time. The overall goal is to provide decision-making support concerning the reduction in some aspect of performance and serviceability rather than warning for incipient failure.

3.1 INTRODUCTION

Bridges are essential elements of a complicated network of infrastructures (Liu and Frangopol, 2006) which is often underperforming, resulting in loss of efficiency in moving goods and people across the country. Long-term monitoring performed proficiently by using the best selection of instruments, thoughtfully placed, is the starting point of continuous monitoring field performance. Interpretation of the acquired data is equally important to obtain information concerning the structural integrity (Bergmeister and Santa, 2001; Feng et al., 2004). Both actions are fundamental to aid management decision-making (DeWolf et al., 2002; Hirachan and Chajes, 2005) and to plan for cost-effective allocation of (often limited) funds and resources. Assessing the integrity of existing bridges and planning effective maintenance (Lansdell et al., 2017) require accurate monitored bridge data, and therefore need for an efficient use of monitoring data (Frangopol et al., 2008). On this matter, there is large interest among researchers and practitioners in developing monitoring systems that efficiently incorporate performance assessment indicators (Strauss et al., 2012).

The overall objective of this work is to support asset management decisions to promote the ongoing serviceability of the structure and the transportation network in which it forms a critical link – not to provide warning of incipient failure. Such changes do not necessarily imply total loss of functionality of the structure, but rather reduction in some aspect of performance or serviceability. Judicious application of structural health monitoring instrumentation and analysis can provide timely decision-support information to extend service life and, as in the case of this bridge, evaluate effects of increased or unanticipated loads.

3.2 MEASURE OF BRIDGE RESPONSE

The ZTP, introduced in chapter 2, is an important measure of bridge response which does present some limitations, mostly due to the fact that it does not consider the axle configuration of the truck but only GVW and length. The new metrics proposed in this thesis are rooted in the influence line theory which is useful for conducting performance evaluation to understand the actual bridge behaviour under real traffic and environmental conditions, to monitor changes over time, to anticipate the response under various loads, and to identify damage (Moses, 1979; Hirachan and Chajes, 2005; O'Brien et al., 2006; Štimac-Grandić et al., 2011; Ieng, 2014; Zhao et al., 2014; Wang et al., 2017). Fitting the bridge response through the superimposition of influence lines allows the acquisition of data on axle weight of crossing vehicles (McNulty and O'Brien, 2003; Lydon et al., 2016; Ojio et al., 2016). Strain data and influence lines have been used over the years to develop methods to compute traffic characteristics such as axle weight, axle spacing and speed using only strain measurements (Cardini and DeWolf, 2008; Wall et al., 2009; Lansdell et al., 2017). In another study, Gagarin et al. (1994) estimated strain influence line in its application of neural network to determine trucks characteristics purely from strain-response measurements.

It is hoped that the metrics introduced in this dissertation will be widely applicable for long-term monitoring and performance evaluation to promote time effectively maintenance which, in turn, will lead to a strengthen asset management of similar bridges.

3.2.1 Strain-Distance-Area (SDA)

Both long-term monitored flexural strain data from SHM and traffic characteristics from WIM are used in the development of the strain-distance-area (SDA) measure, a novel metric to monitor bridge response to traffic. As already mentioned, this metric is based on the concept of

influence line (IL), which is the effect of a moving point load at a location other than the strain gages locations (Liu et al., 2009b; Buckley, 1998). In this study, the effect of interest is bending moment. Based on simple structural considerations, it is possible to relate strain (ϵ) and the bending moment calculated from an influence line (M_{IL}) (Fig.3.8). Stress (σ) can be expressed as the product between Young's modulus (E) and strain deformation (ϵ) (Eq.9) and also as the ratio between M_{IL} times the distance between the fibers subjected to the maximum tension stress and the neutral axes (y) and the moment of inertia (I) (Eq.10).

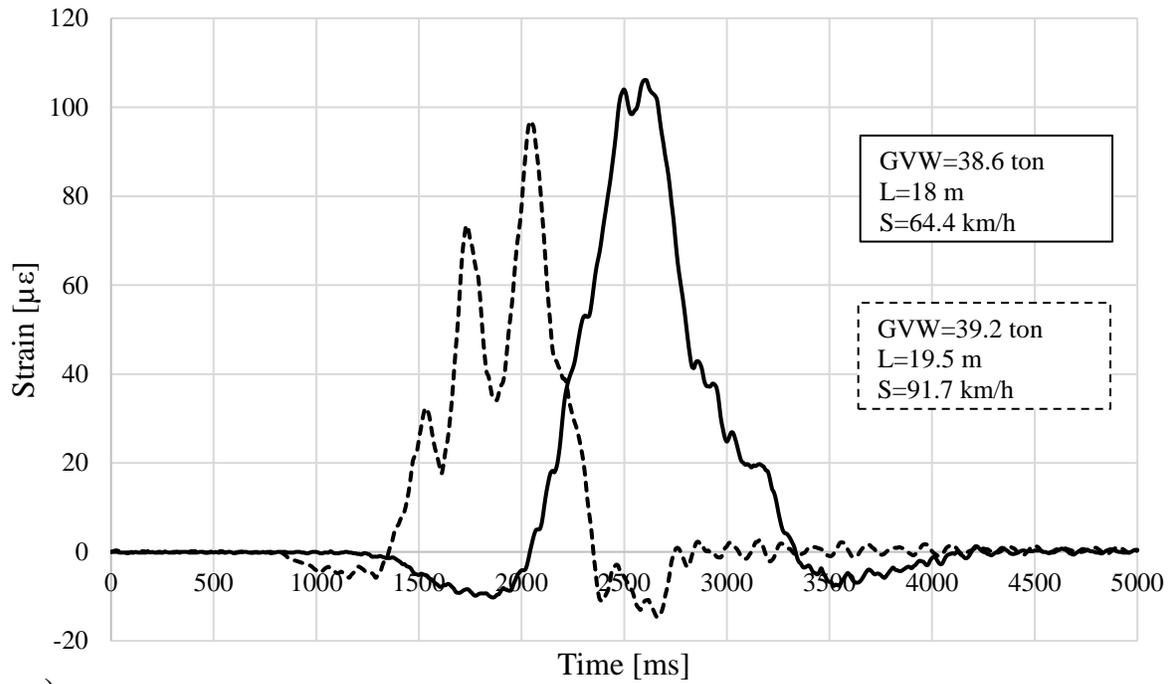
$$\sigma = E \cdot \epsilon \quad (9)$$

$$\sigma = \frac{M_{IL} \cdot y}{I} \quad (10)$$

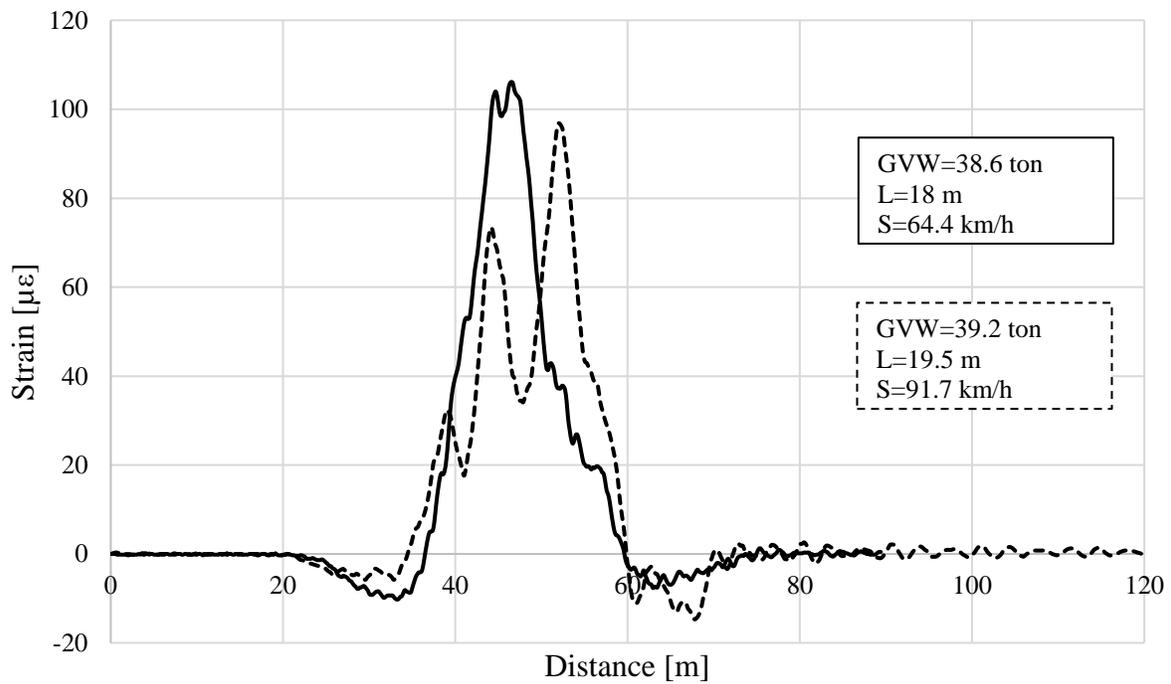
Combining Eq. (9) and Eq. (10) ϵ can be defined as follows:

$$\epsilon = \frac{M_{IL} \cdot y}{E \cdot I} = \alpha_d \cdot M_{IL} \quad (11)$$

All the constant quantities (y , E , I) are grouped into one single term called α_d which represents the bridge flexibility calculated from design (d) drawings (geometry and materials). The development of the novel metric starts with revising Eq.11. For this purpose, preliminary steps are required: first, the recorded strain-time waveform is transformed into a strain-distance waveform, where the distance traveled is calculated using the speed recorded by the WIM system (Fig.3.1a-b); second, after the waveform is transformed, the area under the strain-distance waveform represents the metric proposed in this chapter, SDA.



a)



b)

Figure 3.1: a) strain vs time; b) strain vs distance of two trucks with similar GVW, length, and speed, (girder 4).

M_{IL} (Fig.3.2) is constructed using the traffic data (axle weight and spacing) recorded from the WIM system and the moment-distance area (MDA) under the constructed M_{IL} curve is computed. SDA and MDA are shown in Fig.3.3. Based on previous considerations, Eq.11 is here revised as:

$$SDA = \alpha_e \cdot MDA \quad (12)$$

Where “e” stands for experimental. It will be shown later that, although α_e is a product of field data, it resulted to be a good candidate as a monitoring quantity due to its constant nature.

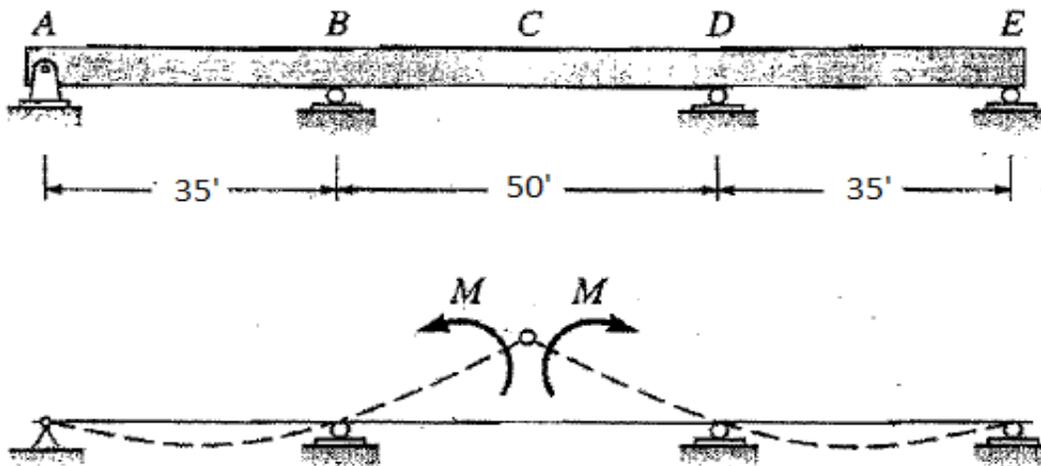


Figure 3.2: theoretical bending moment influence line (M_{IL}).

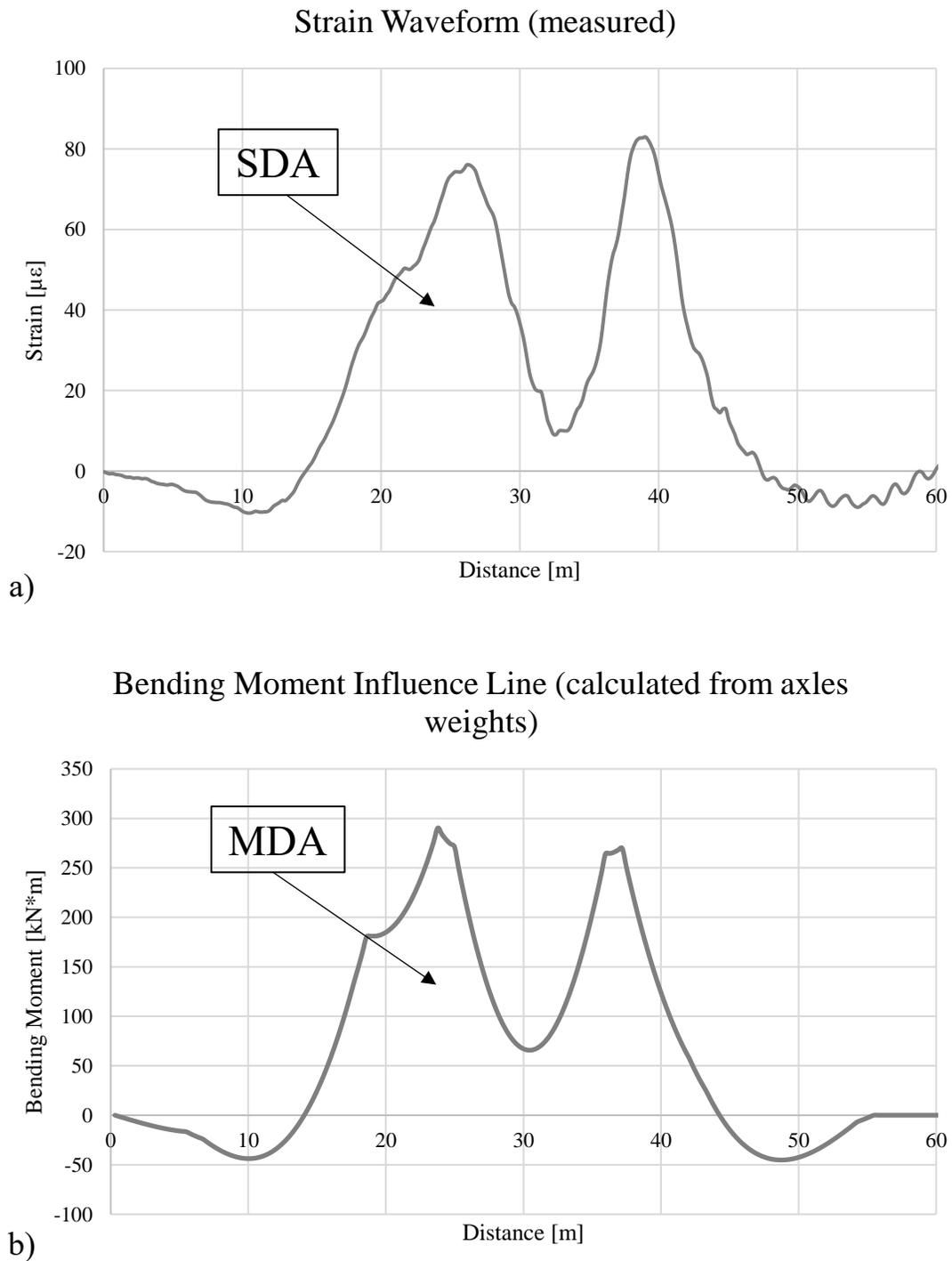


Figure 3.3: a) measured strain waveform (average girder 3 and 4); b) calculated M_{IL} at mid-span, generated by a 331.4 kN (74.5 kips), 22 m (72 ft) truck.

3.2.2 Data Pre-processing

Some pre-processing steps are necessary to guarantee that outliers do not affect the quality of the information in the data. Four conditions are empirically chosen to define an outlier:

- null GVW or length from WIM;
- negative SDA values;
- presence of another vehicle approaching or leaving the bridge while a truck is crossing;
- WIM mismatch between the GVW and the axles weights.

3.2.3 Benefits of SDA

The benefits of SDA over ZTP are discussed as follows. First, ZTP is a concentrated measure of the effect of the load on the bridge. The peak from which ZTP is computed occurs when the truck is in the proximity of the strain gage while SDA takes into account the effect of the truck at any point on the bridge. ZTP is a localized measure while SDA has the important benefit of being independent of the position of the truck, and it includes the effects of the truck throughout its crossing of the bridge.

Using SDA reduces dependency on the truck's length and therefore reduces the scatter in data. ZTP is highly dependent on the length, indeed, a load distributed over a smaller length generates a steeper waveform, while the same load distributed over a bigger length produces a shorter and wider waveform, this generates inevitable scatter in the data (Fig.3.4a). To visualize the reduction in scatter and compare same-unit quantities, ZTP and SDA values are normalized by the sum of all trucks crossing the bridge that month. Fig.3.4b shows, indeed, a significant reduction in the scatter.

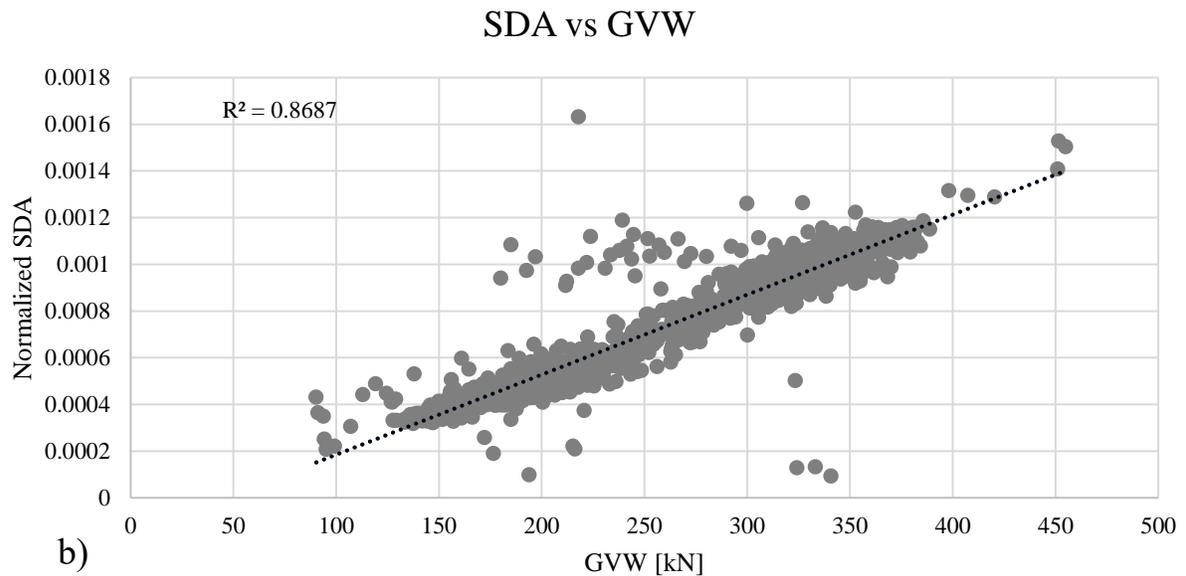
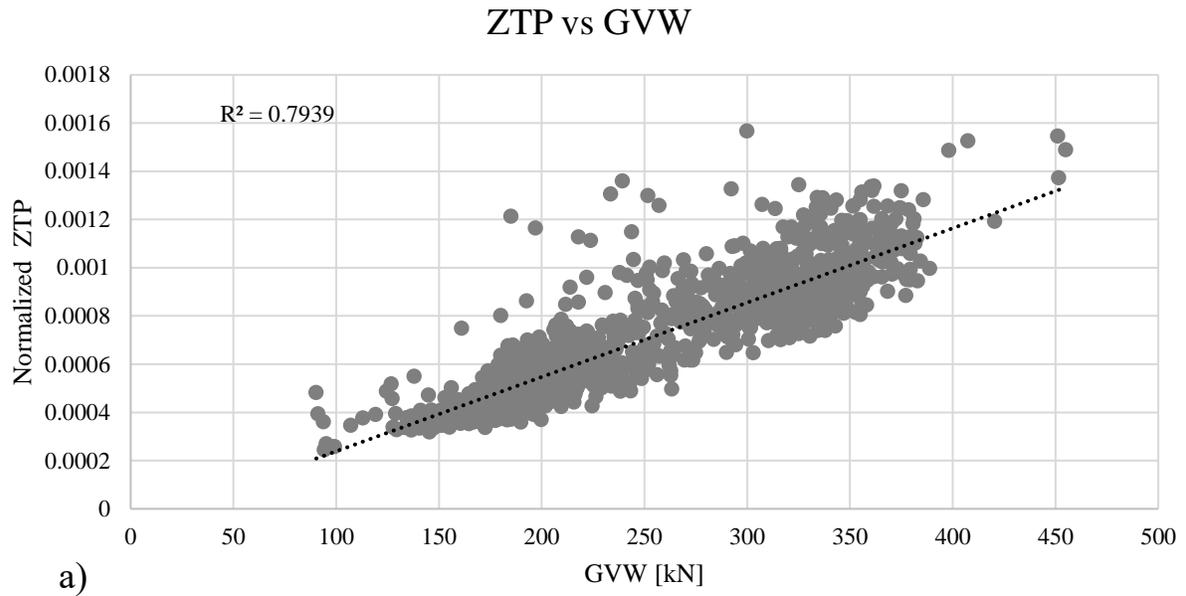


Figure 3.4: a) normalized ZTP vs GVW; b) normalized SDA vs GVW. Values are averages of girder 3 and 4, Lane 1, March 2011.

The third benefit is related to the ability to avoid response underestimation. For instance, the presence of another vehicle on the bridge produces a biased ZTP (Fig.3.5) This anomaly can be avoided or reduced by computing the SDA under the truck's waveform instead.

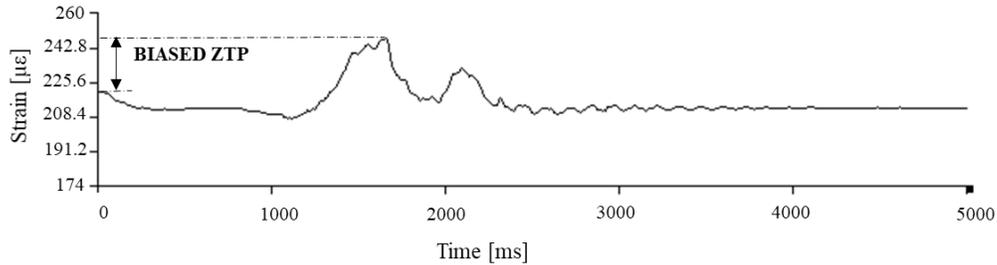


Figure 3.5: an example of underestimated ZTP due to the presence of another vehicle approaching the bridge.

SDA can also be used to calculate the GDF described in chapter 2.

$$SDA_{GDF_{i,j}} = \frac{SDA_{i,j}}{\sum_{j=1}^5 SDA_{i,j}} \quad (13)$$

Where i is the truck identifier and j is the girder ($j=1, \dots, 5$). The monthly average of ZTP_{GDF} and SDA_{GDF} , for both lanes, is calculated and compared in Fig.3.6. Not only it has been shown that SDA can be used to calculate GDF, but also, unlike traditional strain GDF, SDA_{GDF} has the potential to capture the presence of multiple vehicles by using strain data only. Fig.3.7 shows how the ZTP_{GDF} and SDA_{GDF} methods react to this. When ZTP_{GDF} is calculated, these episodes (black dots) are well hidden among the normal values (grey dots) (Fig.3.7a); when the SDA_{GDF} is computed the anomalies are easily recognizable (Fig.3.7b) so that they can be further studied to assess if they are just errors to discard or special events to be taken under special consideration (Chen et al., 2014). The black dots represent the case when two vehicles are crossing the bridge at

the same, typically a car and a truck (Fig.3.5). The black dots above the normal values are typically due to vehicles crossing the bridge on lane 2 while the truck is in lane 1, the ones below are vehicles crossing the bridge in the same lane as the truck. This analysis can potentially lead to the definition of safe thresholds to manage the traffic on the bridge.

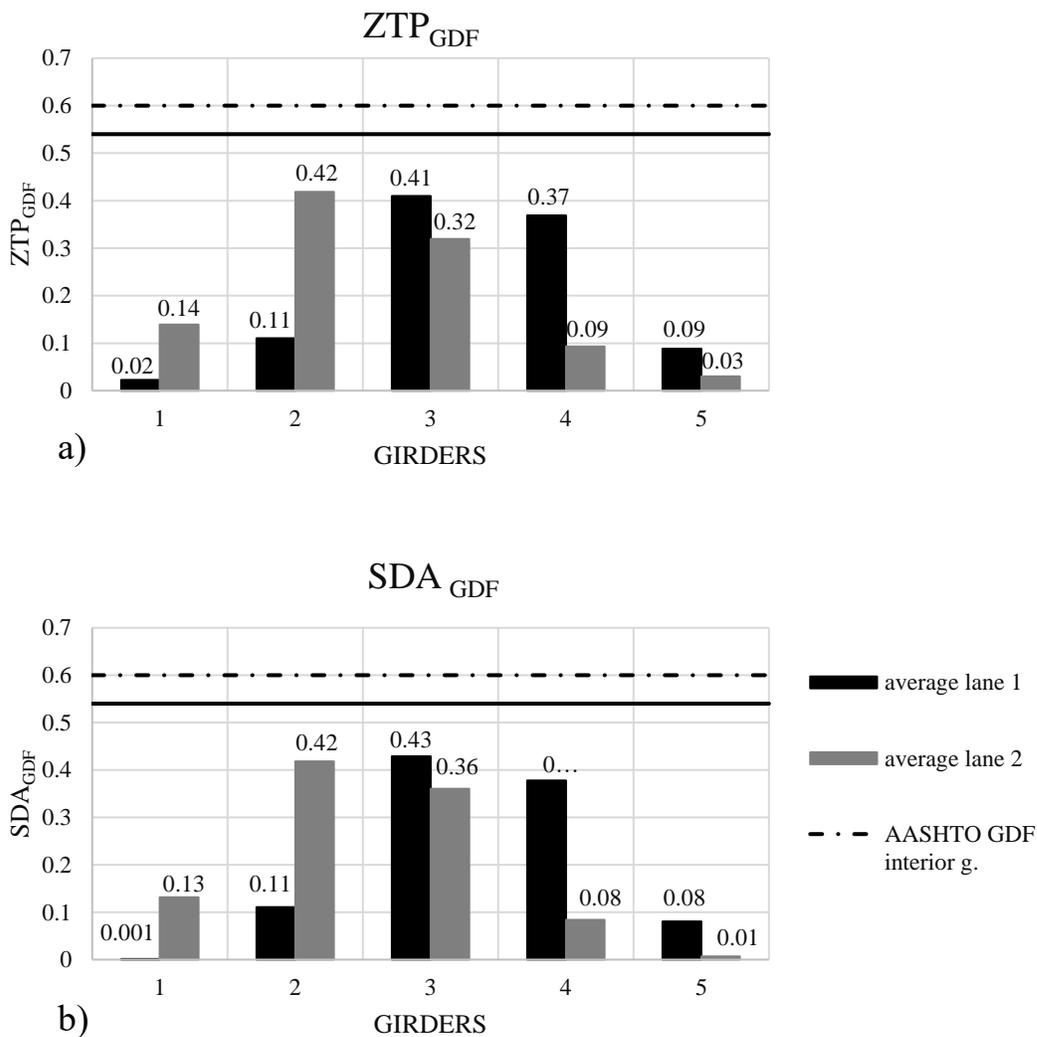


Figure 3.6: a) ZTP_{GDF}; b) SDA_{GDF}. Monthly average (March 2011).

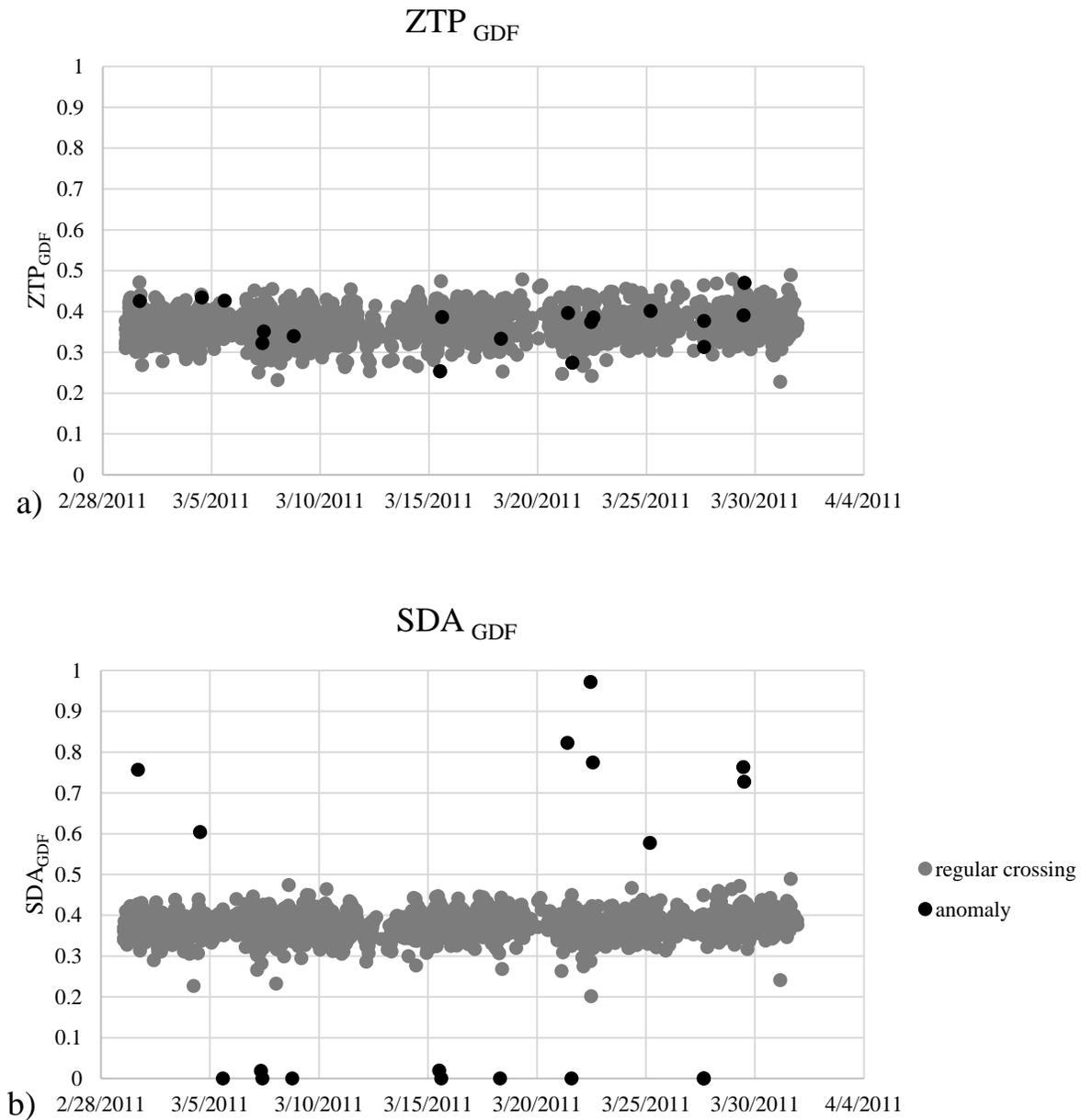


Figure 3.7: a) ZTP_{GDF} ; b) SDA_{GDF} , girder 4 (March 2011).

3.2.4 Bridge Flexibility Over Time

The development of SDA led to the opportunity to propose another metric to monitor changes in the bridge flexibility/rigidity over time. As mentioned before, α is a function of the

bridge, particularly its geometry (y , I) and the materials (E) of girders and deck. The design value of α_d can be calculated from the design drawings. The cross-section used in the calculations includes steel girder (W24x76) and concrete deck, steel bars in the concrete slab have a negligible effect, therefore, are not included. After homogenizing the cross-section in Fig.3.8 into one equivalent material, the design value of α resulted to be the following:

$$\alpha = \frac{y}{E \cdot I} = 1.08 \cdot 10^{-6} \left[\frac{1}{\text{kN} \cdot \text{m}} \right] \quad (14)$$

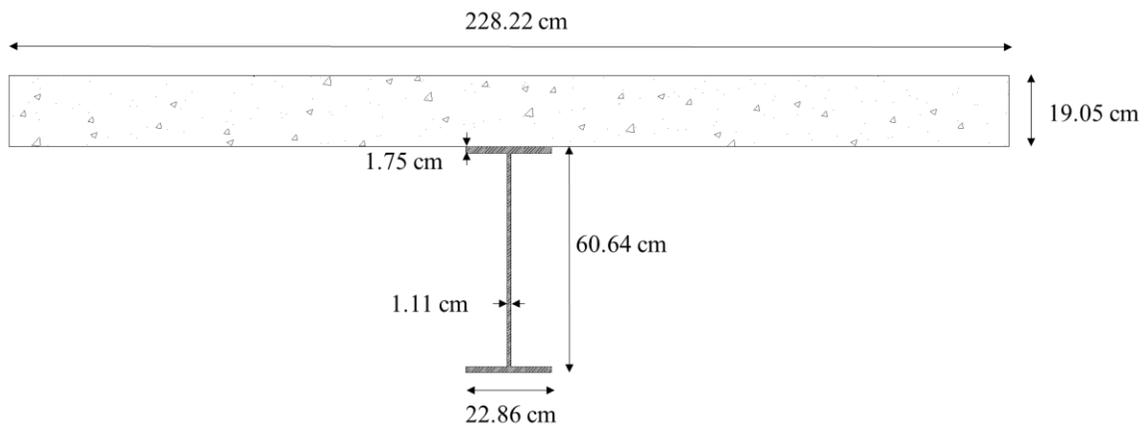


Figure 3.8: cross-section of steel internal girder and concrete deck.

Since only trucks on lane 1 are considered, it is important to get only the portion related to girder 3 and 4 by multiplying α for the appropriate GDF.

$$\alpha_d = \alpha \cdot \text{GDF}_{\text{AVE}(3,4)} = 4.2 \cdot 10^{-7} \left[\frac{1}{\text{kN} \cdot \text{m}} \right] \quad (15)$$

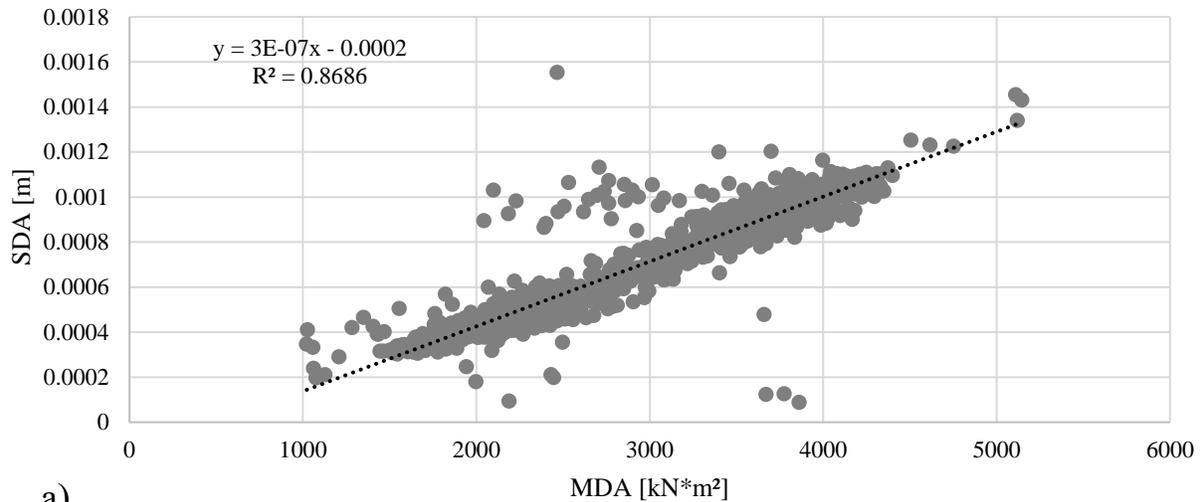
On the other hand, α can also be estimated from field data as the ratio between SDA and MDA (i.e. the slope of the scatterplots in Fig.3.9).

$$\alpha_e = \frac{\text{SDA}}{\text{MDA}} = 2.8 \cdot 10^{-7} \left[\frac{1}{\text{kN} \cdot \text{m}} \right] \quad (16)$$

In spite of the uncertainty and environmental factors which can affect the 3 years of recorded data, both design and experimental values of α are quite close to each other. The design value is larger, meaning more flexibility, likely due to factors such as variation in concrete stiffness, age, discrepancy between specified concrete compressive strength (f_c') and in-place concrete strength.

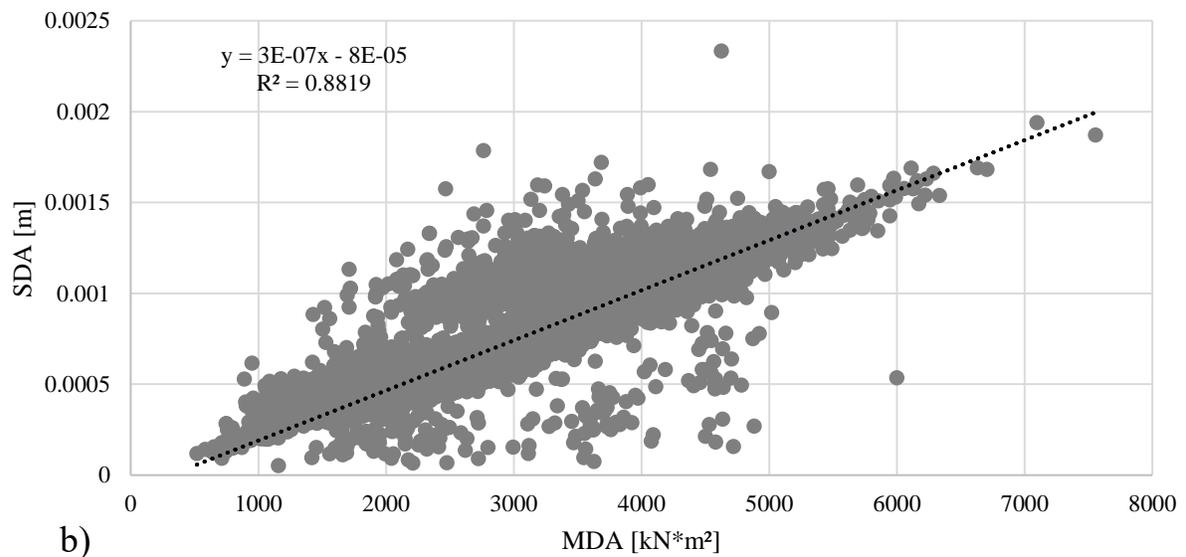
α_e is here proposed as a valid metric to monitor the bridge flexibility over time. During the first stage of life of the structure, any change from α_d , and later in time any change of α_e , can be considered as an alert of a change in the components of the bridge or in their materials. This monitoring can lead to targeted effective maintenance actions. Fig.3.9 shows two important pieces of evidence used to confirm the reliability of α_e as a metric of the bridge performance. First, the volume of traffic does not influence α_e , for this purpose, Fig.3.9a and Fig.3.9b shows that one month and three years of recorded data present the same value of α_e . Second, the type of traffic does not influence α_e , for this purpose, Fig.3.9a and Fig.3.9b show respectively only class 9 trucks (5-axle trucks) and all classes of trucks from class 9 to class 13 (from 5- axle to 9axle trucks) confirming the same α_e and therefore its validity.

SDA vs MDA (Influence Line calculated with axle weights)



a)

SDA vs MDA (Influence Line calculated with axle weights)



b)

Figure 3.9: a) one-month (March 2011), class 9 (5-axle) trucks only; b) three years, from class 9 to class 13 (up to 9-axle) trucks.

α_e can be estimated by plotting ZTP_ϵ vs ZTP_{IL} , however, there is one particular case where computing α_e from SDA-MDA and not from ZTP_ϵ - ZTP_{IL} is highly recommended. This happens when the axle configuration (axle weight and spacing) is not available but GVW is the only accessible piece of information. With this parameter is still possible to build M_{IL} and calculate MDA as shown in Fig.3.10. Fig.3.11a and Fig.3.11b show that, if the axle configuration is available, using ZTP_ϵ - ZTP_{IL} or SDA-MDA to estimate α_e is equivalent, indeed the slope shows the same value of α_e . However, when the axle configuration is not available, and M_{IL} is constructed based on GVW only, the SDA-MDA (Fig.3.11c) still produces the correct α_e value while ZTP_ϵ - ZTP_{IL} fails to do so (Fig.3.11d, Table 3.1).

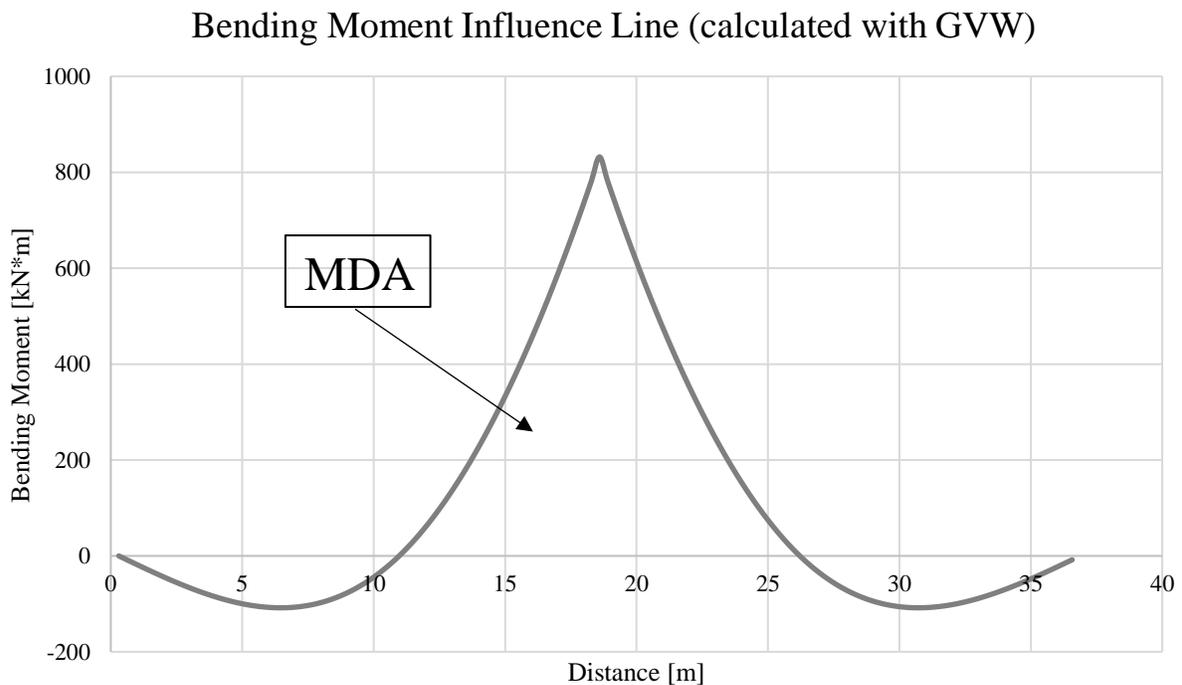
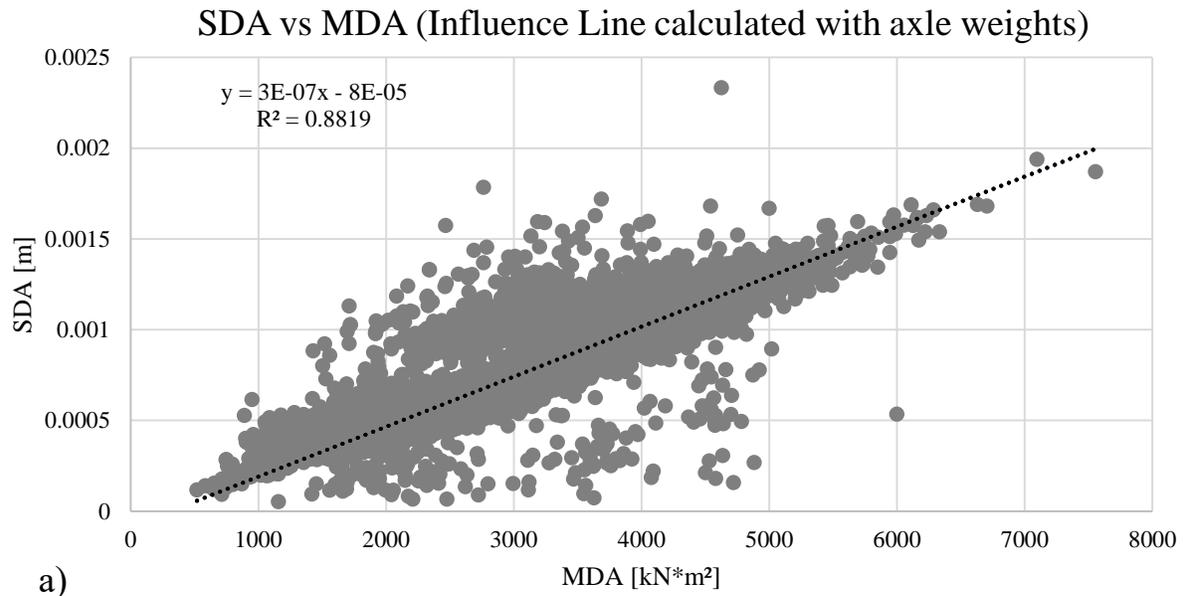
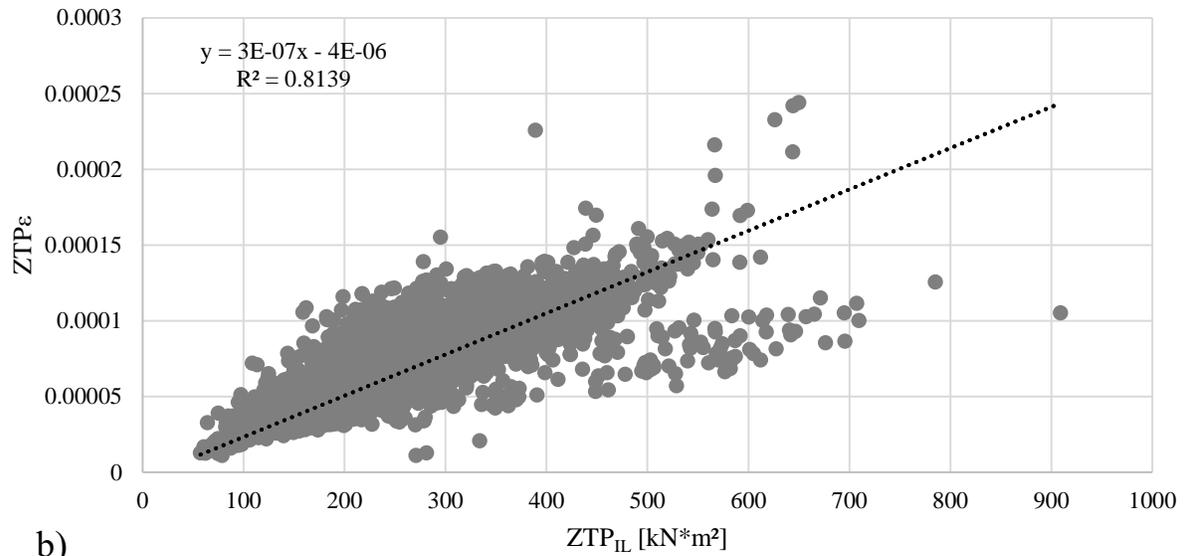
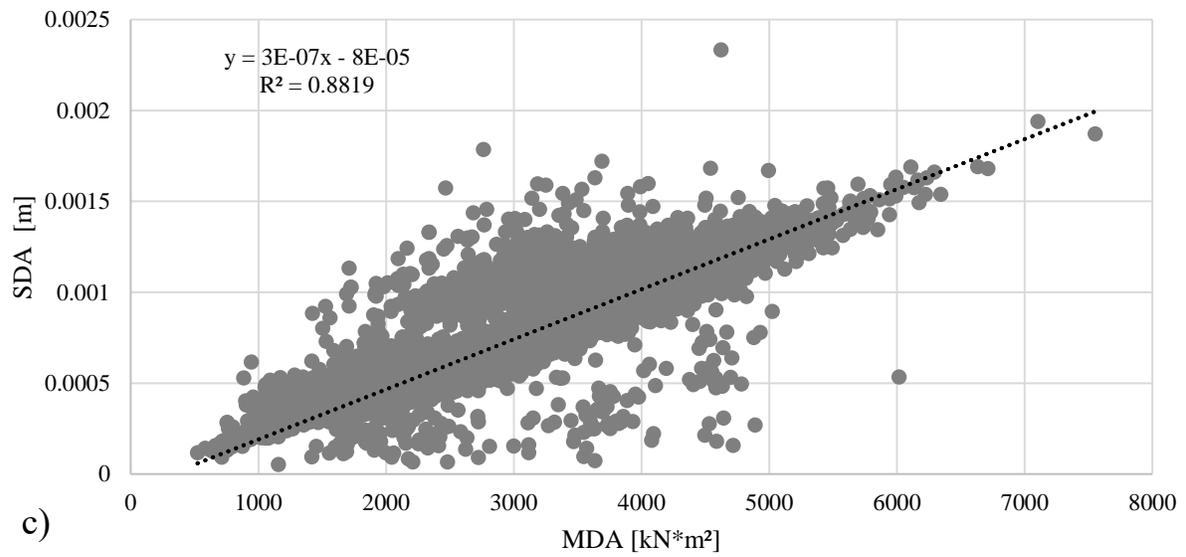


Figure 3.10: M_{IL} at mid-span, truck considered as a point load.

Table 3.1: α_e values from fitting line.

	α_e
SDA vs MDA (IL calculated with axle weights)	$2.75 \cdot 10^{-7}$
ZTP ϵ vs ZTP $_{IL}$ (IL calculated with axle weights)	$2.72 \cdot 10^{-7}$
SDA vs MDA (IL calculated with GVW)	$2.75 \cdot 10^{-7}$
ZTP ϵ vs ZTP $_{IL}$ (IL calculated with axle GVW)	$0.96 \cdot 10^{-7}$



ZTP ϵ vs ZTP $_{IL}$ (Influence Line calculated with axle weights)**SDA vs MDA (Influence Line calculated with GVW)**

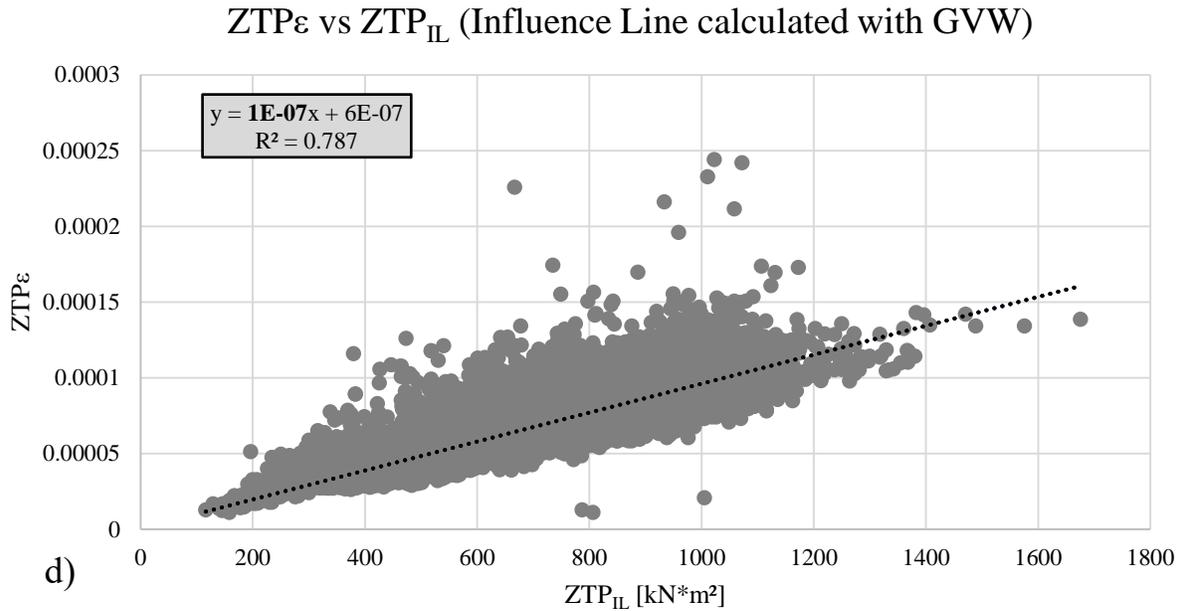


Figure 3.11: a) SDA-MDA and b) ZTP_{ϵ} - ZTP_{IL} IL calculated with axle weights; c) SDA-MDA and d) ZTP_{ϵ} - ZTP_{IL} IL calculated with GVW.

3.3 CONCLUSIONS

This chapter presents novel metrics to monitor the influence of heavy traffic on the bridge response and potential changes in rigidity/flexibility over time. Nominally, two metrics, rooted in the influence line theory, are proposed.

The development of the first metric for bridge response consists of two steps. First, the strain-time waveform recorded from SHM is converted into a strain-distance waveform, then, the area under that curve is computed. The computed area is here proposed as an alternative strain measure for bridge response and it is called strain-distance area (SDA). SDA presents several benefits over more traditional strain metrics:

- it provides a generalized measure of the response of the bridge as opposed to a localized measure;

- it shows significant reduction of the dependency on length, resulting in significant reduction of scatter in data;
- it identifies hidden anomalies which can lead to response underestimation;
- it provides an alternative way to calculate GDF and it allows to efficiently detect anomalies of from the monitoring systems.

SDA is also used in the implementation of a second metric to monitor the bridge rigidity/flexibility over time. The ratio between the area under the recorded strain waveform (SDA) and the area under the calculated bending moment influence line (MDA) provides α_e , which, if constantly compared to α_d (bridge flexibility/rigidity calculated from design drawings) is able to detect changes in the bridge rigidity/flexibility. It is shown that in the case of lack of detailed information on the truck's axle configuration, α_e obtained by plotting SDA vs MDA assure the achievement of reliable values of flexibility, unlike more traditional strain methods (ZTP_{ϵ} - ZTP_{IL}). This work aims to provide practitioners improved metrics with immediate applicability.

4 PREDICTING BRIDGE RESPONSE

After defining valid metrics to study the bridge response, the next step is to identify the best way to model and predict bridge performance to ensure resource allocation in a cost-effective manner. Bridge management can benefit from machine learning techniques to prolong the life of the structure and improve safety. The first goal of this work is to present a comparison between Multilinear Regression, Artificial Neural Network, and Regression Tree, which are used to model the performance of a bridge subjected to a wide variety of trucks. The second goal is to propose an alternative strategy for practitioners to avoid trading off predictive power for explanatory capabilities and interpretability. The models use both WIM and SHM data.

4.1 INTRODUCTION

To be able to monitor and assess bridge conditions, three key components are needed: (i) data availability, here facilitated by long-term monitoring using both WIM and SHM data (section 1.4); (ii) appropriate metrics of infrastructure performance (Chapter 2); (iii) efficient predictive models, achieved by means of analytical tools such as machine learning techniques to identify conditions that undermine the safety of a bridge, as well as to prevent critical events and this is the purpose of this chapter.

Extracting information from large amounts of data can be a challenging endeavor for any organization and becomes particularly difficult in the resource-constrained environment in which many infrastructure managers are forced to operate. For this reason, more traditional methods to evaluate bridge performance are often still preferred to more advanced data mining techniques. The goal of this study is to address bridge performance by comparing advantages and disadvantages of three techniques namely, Multilinear Regression (MLR), Artificial Neural Network (ANN), and Regression Tree (RT). These three supervised learning methods are chosen to cover a wide range of skills. MLR, widespread in civil engineering, is chosen for its interpretability and explanatory skills. ANN, commonly used to solve a wide variety of problems in civil engineering, is chosen for its outstanding predictive power. RT, not commonly adopted in civil engineering, is chosen for its user-friendly structure, which can result to be an excellent tool for practitioners to justify to a non-technical audience (i.e. management, public, clients, etc.) how decisions are made. After comparing strengths and weaknesses of the three methods, this study also aims to contribute with an alternative predictive strategy where users do not need to trade-off between different capabilities, but they can finally benefit from interpretability, explanatory

capabilities, and predictive power by using only one method. The proposed strategy combines influential observations analysis (IOA) with regression models, in this case, MLR is preferred. It is shown that, once influential observations are identified and discarded, the predictive power of MLR is significantly improved and outperforms the well-known predictive capabilities of ANN. Owing to the bridge configuration, which is representative of many others in the country, the results presented in this study have a broader applicability to many other real cases.

4.1.1 Technical Background

Many civil engineering problems have been studied using machine learning methods, in particular, ANN. The first application of ANN to civil engineering was published in 1989 to model the design of steel beams (Adeli and Yeh, 1989; Adeli 2001). Li et al. (1996) first applied ANN to assess bridge conditions using inspection data. An example of how ANN can be used to predict highway bridge performance was proposed by Tokdemir et al. (2000). ANN is also considered a promising tool for damage detection (Wu et al. 1992; Li et al., 2011) and to reduce subjectivity in structural assessment traditionally performed by visual inspection (Elkordy and Chang, 1993; Williams, 1994; Cattan and Mohammadi, 1997). A number of studies have shown that ANN has advantages over traditional statistical methods (Elhag and Boussabaine, 2002) where the statistical distributions are unknown, in the presence of outliers and noise (Burke, 1991) and when the sample size is large (King 1999; Nguyen and Cripps, 2001). More recent examples of the use of ANN in civil engineering include analysis of seismic data (Kerh et al., 2011; Adeli and Panakkat, 2009; Youd et al., 2002), estimates of dynamic displacement of bridges under dynamic loads (Masri et al., 1993; 1996; Ok et al., 2012), and prediction of bridge health condition using acceleration and displacement data (Suryanita and Adnan, 2013). Gagarin et al. (1994) used strain data and weigh-

in-motion data to compute truck attributes such as velocity, axle loads, and axle spacing, using ANN. MLR models have been specifically used in the analysis of dams since the 1950s. A comparative study of ANN and MLR to interpret dam behavior is presented by Mata (2011). Other studies compare ANN and MLR to determine future maintenance priorities based on the analysis of bridge score risk (Wang and Elhag, 2007; Elhag and Wang, 2007). While extensive literature suggests that ANN outperforms traditional statistical methods, Kumar (2005) found that for skewed data, MLR performs better than ANN. According to Dunlop and Smith (2010), MLR produces positive findings in the analysis of concrete delivery and placement process. RT is not widely present in the literature, however, Reich (1997), suggested RT as a potential machine learning technique to solve civil engineering problems.

Although this study does not aim to provide a comprehensive survey of the existing literature review on the topic, a list of representative papers, selected from above, which discusses MLR and ANN in structural engineering is organized in an easy-to-read table (Table 4.1). The table sorts the studies by methodology, monitored structure, data collected and goal of the study. The common goal of the selected studies is to evaluate the structure's conditions by predicting parameters which can shed a light on potential defects, risks or damage such as strain response and displacement. Among these papers, 53% of the monitored structures are bridges. 90% of the papers use ANN to make predictions about structures features and 26% of the papers present comparative studies. Based on the outcome of the existing literature review, this study aims to provide a comparison where RT is also suggested as a valid algorithm which was not highly considered over the 24-year period analyzed. Also, the analysis of the literature review identifies a clear prevalence of ANN use, with most of the studies highlighting the superiority of ANN over MLR. For these reasons,

the authors were prompted to explore a method that could provide the same predictive power of ANN and also provide interpretability and explanatory capabilities typical of MLR.

Table 4.1: chronological list of MLR and ANN applications in structural engineering.

Reference	Method	Structure	Data	Goal
Adeli and Yeh (1989)	ANN	Beams	Control Parameters and Perceptron	Design Steel Beams
Wu et al. (1992)	ANN	3-Story Building	Membrane Stiffness	Identify Structural Damage
Elkordy et al. (1993)	ANN	Steel Frame	Acceleration Data	Damage Diagnosis of Structures
Williams (1994)	ANN	Structures	Macroeconomic Data	Construction Cost Indexes
Gagarin et al. (1994)	ANN	Bridge	Strain and WIM Data	Predict Truck Attributes from Strain Response from Bridge
Li et al. (1996)	ANN	Bridge	Inspection Data	Evaluate Bridge Conditions
Masri et al. (1996)	ANN	Structures	Vibration Data	Detection of Structural Changes
Cattan & Mohammadi (1997)	ANN	Railroad Bridge	Bridge Parameters	Bridge Condition Rating
Tokdemir et al. (2000)	ANN*	Highway Bridge	Geometry, Age, Traffic, Structural Attributes	Bridge Rating
Youd et al. (2002)	MLR	Structures	Displacement, Earthquake Data	Lateral Spread Displacement
Dunlop & Smith (2003)	MLR	Concrete	Concrete Data	Estimate Efficiency of Concrete Operations
Wang and Elhag (2005)	ANN MLR*	Bridge	Risk Data (Safety, Functionality, Sustainability)	Risk Assessment
Elhag and Wang (2007)	ANN MLR	Bridge	Risk Data	Risk Score, Risk Category
Adeli and Panakkat (2009)	ANN		Earthquake Data	Earthquake Magnitude Prediction
Mata (2011)	ANN MLR	Dam	Relative Displacement, Strains, Stresses in Concrete	Characterization of Dam Behaviour under Loads
Kerh et al. (2011)	ANN	Bridge	Seismic Data	Identify Hazardous Bridges
Li et al. (2011)	ANN*	Bridge	Frequency Data	Damage and Defect Identification
Ok et al. (2012)	ANN	Bridge	Dynamic Loads	Estimate Displacement
Suryanita and Adnan (2013)	ANN	Bridge	Acceleration and Displacement Data	Predict Bridge Conditions

*Other methods are used in the comparison.

4.1.2 Data Pre-processing

To avoid potential WIM lane misclassification, the following relations are used to assign the correct lane to each truck:

$$ZTP_{\text{girder4}}(i) > ZTP_{\text{girder2}}(i) \rightarrow \text{Lane 1}; ZTP_{\text{girder4}}(i) < ZTP_{\text{girder2}}(i) \rightarrow \text{Lane 2} \quad (17)$$

Only trucks crossing the bridge on lane 1 are considered for this study because in general slower traffic like trucks tend to keep right, indeed, 95% of the recorded trucks crossed the bridge on lane 1, providing a significant sample size for the analyses. To assure the quality of the dataset some constraints are set. When a truck is crossing lane 1 girders 3 and 4 are the most loaded, however, it is important to ensure that also the fascia girders (external), although less loaded, show a non-zero perturbation:

$$SDA_{\text{girder 2}} > 0; SDA_{\text{girder 5}} > 0 \quad (18)$$

For the most loaded girders, a value of 6 $\mu\epsilon\cdot\text{m}$ is set as minimum, which is the monthly average for the external girders:

$$SDA_{\text{girder 3}} > 6 [\mu\epsilon \cdot \text{m}]; SDA_{\text{girder 4}} > 6 [\mu\epsilon \cdot \text{m}] \quad (19)$$

Finally, this study considers 14,660 class 9 trucks which represent 74% of all the trucks in the dataset. The initial variables included in the data analysis are: SDA [$\mu\epsilon\cdot\text{m}$] or [$10^{-6}\cdot\text{m}$] (in this chapter only the positive area under the curve is considered), GVW [kg or ton], length of the truck [m], speed [km/h], weight of each axle (ax1, ax2, ax3, ax4, ax5) [kg or ton], spacing between axles (s12, s23, s34, s45) [m].

4.2 METHODOLOGY

MLR, ANN, and RT are all supervised learning techniques where the focus is to predict the response of the bridge to trucks, in terms of SDA. Hereafter, their principles are explained. A goodness of fit is performed for all three methods.

4.2.1 Multilinear Regression (MLR)

Regression analysis aims to explain the relationship between variables and predict the value of target covariates by fitting a function to patterns of data. In MLR the dependent variable (or response, y) is estimated from k independent variables (or predictors, x_1-x_k) by using a linear equation.

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + \cdots + \beta_k \cdot x_k + \varepsilon \quad (20)$$

where β_0 is the intercept of the model, i.e. the expected value of y when all the predictors are zero, β_{1-k} are the coefficients which will be estimated by the model and ε is the error, which describes the effect of latent variables not included in the model. The goal is to estimate all the β coefficients by minimizing the sum of squared errors (SSE).

$$SSE = \sum_{i=1}^n [y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_{i1} + \widehat{\beta}_2 \cdot x_{i2} + \cdots + \widehat{\beta}_k \cdot x_{ik})]^2 = \sum_{i=1}^n e_i^2 \quad (21)$$

$$\frac{\partial SSE}{\partial \beta_0} = -2 \cdot \sum_{i=1}^n [y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_{i1} + \widehat{\beta}_2 \cdot x_{i2} + \cdots + \widehat{\beta}_k \cdot x_{ik})] = 0 \quad (22)$$

$$\frac{\partial SSE}{\partial \beta_j} = -2 \cdot \sum_{i=1}^n [y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_{i1} + \widehat{\beta}_2 \cdot x_{i2} + \cdots + \widehat{\beta}_k \cdot x_{ik})] \cdot x_{ij} = 0 \quad (23)$$

with $(j = 1, 2, \dots, k)$

Which means solving $k+1$ equations in $k+1$ unknowns.

$$\begin{bmatrix} n & \sum x_{i1} & \sum x_{i2} & \cdots & \sum x_{ik} \\ \sum x_{i1} & \sum x_{i1}^2 & \sum x_{i1} x_{i2} & \cdots & \sum x_{i1} x_{ik} \\ \sum x_{i2} & \sum x_{i2} x_{i1} & \sum x_{i2}^2 & \cdots & \sum x_{i2} x_{ik} \\ \vdots & \vdots & & \ddots & \vdots \\ \sum x_{ik} & \sum x_{ik} x_{i1} & \sum x_{ik} x_{i2} & \cdots & \sum x_{ik}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_{i1} \\ \sum y_i x_{i2} \\ \vdots \\ \sum y_i x_{ik} \end{bmatrix} \quad (24)$$

Once the following terms are defined in matrixial terms as follows:

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}; \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ 1 & x_{31} & x_{32} & \cdots & x_{3k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}; \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}; \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (25)$$

The model in its compact form becomes:

$$\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (26)$$

When $[\mathbf{X}^T \mathbf{X}]$ is invertible, the least square solution becomes:

$$[\mathbf{X}^T \mathbf{X}] \cdot \hat{\boldsymbol{\beta}} = \mathbf{X}^T \cdot \mathbf{Y} \quad (26)$$

$$\hat{\beta} = [X^T X]^{-1} X^T \cdot Y \quad (27)$$

A standard measure of the goodness of fit is the coefficient of determination R^2 which represents the portion of the variability in the data explained by the model.

$$R^2 = 1 - \frac{SSE}{SST} \quad (28)$$

$$SST = SSE + SSR \quad (29)$$

$$SSE = \sum_{i=1}^n e_i^2; SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2; SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (30)$$

Where SST is the total sum of squares, SSE is the error sum of squares and SSR is the regression sum of squares. It is usually good practice to use the adjusted R^2 (R^2_{adj}) which is calculated using the error mean of squares (MSE) and the total mean of squares (MST) to adjust for new predictors added to the model to prevent the problem of overfitting. However, when the sample size is large the difference is negligible; for this reason, in this study R^2 is adopted as a measure the goodness of fit.

4.2.2 Artificial Neural Network (ANN)

ANN aims to mimic how the human brain processes complicated data sets by modeling the response as a nonlinear function of various linear combinations of the predictors. The structure of ANN used in this study consists of one input layer, one hidden layer with multiple hidden nodes, and one output layer. These components are defined as follows.

- The input layer consists of k predictors (x) which are here are called input variables.

- The hidden layer consists of m hidden nodes (H), which are functions of the k predictors and m model parameters (α). Sigmoidal activation functions are used for all hidden nodes.
- The output layer consists of the response y , which is here called output variable. y is a function of H and model parameters β . A linear function is used for the output.

A single-hidden-layer neural network model is fitted to the training dataset. In order to fit the model, the response must be rescaled in a $[0,1]$ interval and the predictors must be standardized as follows, where j is the predictor and i is the observation.

$$y_i = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \quad (31)$$

$$x_{ij} = \frac{x_{ij} - \text{mean}(x_j)}{\text{std dev}(x_j)} \quad (32)$$

A potential problem with ANN is overfitting. To overcome this issue, a user-chosen tuning parameter λ , called decay, is included. The parameters are estimated by minimizing the nonlinear sum of squares (SSE*) as shown below:

$$\text{SSE}^* = \sum_{i=1}^n [y_i - g(\mathbf{x}_i, \boldsymbol{\theta})]^2 + \lambda \cdot \left(\sum_{m=1}^M \sum_{j=0}^k \alpha_{m,j}^2 + \sum_{m=0}^M \beta_m^2 \right) \quad (33)$$

In Eq.33, the second term is the shrinkage term, where λ multiplies the sum of squares of nonlinear regression coefficients and it is assigned to govern the trade-off between overfitting and underfitting. The best values for λ are usually between 0.001 and 0.1 (Kutner et al., 2004). M is the total number of hidden nodes m , k is the total number of predictors j . The first term minimizes

the SSE, where $g(\mathbf{x}_i, \boldsymbol{\theta})$ is the neural network response prediction which is also a function of the α and β , represented by θ , as shown below:

$$g(\mathbf{x}_i, \boldsymbol{\theta}) = \hat{y}_i = \frac{\exp\{\beta_0 + \beta_1 \cdot H_{i,1} + \dots + \beta_M \cdot H_{i,M}\}}{1 + \exp\{\beta_0 + \beta_1 \cdot H_{i,1} + \dots + \beta_M \cdot H_{i,M}\}} \quad (34)$$

$$\hat{H}_{i,m} = \frac{\exp\{\alpha_{m,0} + \alpha_{m,1} \cdot x_{i,1} + \dots + \alpha_{m,k} \cdot x_{i,k}\}}{1 + \exp\{\alpha_{m,0} + \alpha_{m,1} \cdot x_{i,1} + \dots + \alpha_{m,k} \cdot x_{i,k}\}} \quad (35)$$

4.2.3 Regression Tree (RT)

RT is a valid alternative model which combines interpretability and capacity to model non-linear relationships. The idea behind RT is simple, in that fitting models corresponds to growing trees one node at a time. The tree model is the following:

$$g(\mathbf{x}, \boldsymbol{\theta}) = \sum_{m=1}^M c_m I(\mathbf{x} \in R_m) \quad (36)$$

Where M is the total number of regions or terminal nodes, R_m is the m^{th} region, I is an indicator function equal to 1 or 0 depending on x belonging to R_m . The highest reduction of SSE defines the best predictor and the best place to split the data set into regions. At each split, the RT model divides the predictor's space into rectangular regions where the predicted response y is the average response over the entire region c_m . N_m is the size of the m^{th} terminal node (region) and θ represents all the parameters and the structure.

$$\text{SSE} = \sum_{i=1}^n [y_i - \hat{y}_i]^2 = \sum_{m=1}^M \sum_{x_i \in R_m} [y_i - \hat{c}_m]^2 \quad (37)$$

$$\hat{c}_m = \frac{1}{N_m} \cdot \sum_{x_i \in R_m} y_i \quad (38)$$

Since more algorithms overfit than it is important to prune back the branches which means to collapse one of the internal nodes into a single terminal node. For this purpose, the best size of the tree λ , also called complexity parameter, is obtained by minimizing the deviance (Sutton, 2004) which, for continuous responses, is equal to minimize the mean square error.

4.2.4 Influential Observations Analysis (IOA)

The concept of influential observations is pivotal to introduce the alternative predictive strategy. Influential observations are outliers which are likely to substantially change the regression model results (Stevens, 1984). One approach to identify them is to compute the Cook's distance (Cook, 1977; 1979) to assess how much the regression would change if a specific observation is omitted. A Cook's distance higher than $4/n$ (where n is the sample size) (Bollen and Jackman, 1985) is the rule of thumb for flagging an observation as influential. This study refers to this procedure as Influential Observations Analysis (IOA). The approach chosen for the IOA is based on regressions because this strategy will propose to combine IOA with MLR to achieve interpretability, explanatory skills, and short computing time along with outstanding predictive power.

4.3 RESULTS

4.3.1 Multilinear Regression (MLR)

Some preliminary actions must be taken before running an MLR model. A potential problem that might affect the results is multicollinearity, which is due to strong correlation among predictors and can cause inflation of the probability value (p-value). The p-value quantifies the level of statistical significance, and it is generally set to the cut-off level of p-value<0.05. The best way to detect multicollinearity is to calculate the Variance Inflation Factors (VIF) for each β . As a rule of thumb, a VIF higher than 10 indicates multicollinearity and therefore it cannot be accepted (Alin, 2010). For instance, a value of VIF equal to 100 means that the variance of the coefficient is 100 times larger than it would be if all predictors were uncorrelated. Eq.39 shows how VIF is calculated, where R^2_j denotes R^2 for regressing predictor x_j onto the remaining predictors $\{x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k\}$.

$$VIF_j = \frac{1}{1 - R^2_j} \quad (39)$$

To correct for multicollinearity, the predictors with an elevated VIF are excluded from the model. First, the VIF coefficients are calculated for all 12 predictors (weight, length, speed, ax1, ax2, ax3, ax4, ax5, s12, s23, s34, s45). As expected, variables found to have an elevated VIF are weight, and the individual axle weights (ax1 to ax5), given that the weight is roughly the sum of all axles' weights. As a result, the individual axles' weight will be removed from the regression model. When computing the VIF coefficients with the reduced variable-set, all predictors fall within the acceptable range (VIF<10).

Based on the previous considerations, the final MLR model can be summarized by the following equation:

$$SDA = \beta_0 + \beta_1 \cdot GVW + \beta_2 \text{length} + \beta_3 \text{speed} + \beta_4 s_{12} + \beta_5 s_{23} + \beta_6 s_{34} + \beta_7 s_{45} + \varepsilon \quad (40)$$

Table 4.2 shows the results from the MLR model where all parameters are statistically significant with intuitive signs.

Table 4.2: estimated coefficients of MLR model.

Predictors	$\hat{\beta}$ coefficients $\cdot 10^{-6}$	Std. Error	t-value	p-value
Intercept	-56.43	11.06	-5.102	3.4e-7
GVW [kg]	0.029	0.0001	260.267	< 2e-16
Length [m]	-9.676	0.9421	-10.270	< 2e-16
Speed [km/h]	-1.927	0.0739	-26.065	< 2e-16
s12 [m]	22.13	1.494	14.807	< 2e-16
s23 [m]	35.33	2.448	14.432	< 2e-16
s34 [m]	16.82	1.184	14.212	< 2e-16
s45 [m]	23.86	1.181	20.192	< 2e-16

As expected, GVW has a significant impact on SDA, for each change of 1 [ton] in GVW, the average increase in the mean of SDA is about $29 \cdot 10^{-6}$ [m], also all the spacing variables are found to be strong drivers of SDA variation, for instance, each change of 1 [m] in s23, the average increase in the mean of SDA is about $35.33 \cdot 10^{-6}$ [m]. The combination of spacing is important because it defines the shape of the strain waveform and therefore SDA. If the spacing between two axles is large, the strain response decays and rises again as the next axle approaches the sensor, resulting in a bimodal shape. If the spacing between axles is smaller, this means that axles are more equally spaced and the time between axles is not enough to allow the strain waveform to

significantly decay. Speed and length do not contribute much to SDA because by definition SDA excludes the dependence on speed and reduces the dependence on length. For MLR R^2 is equal to 0.85, meaning that 85% of the variability in the data is explained by the model. Actual SDA and Predicted SDA are plotted as an additional proof of the goodness of fit for the test dataset (Fig. 4.1a). A simple scheme of MLR is provided in Fig.4.2a.

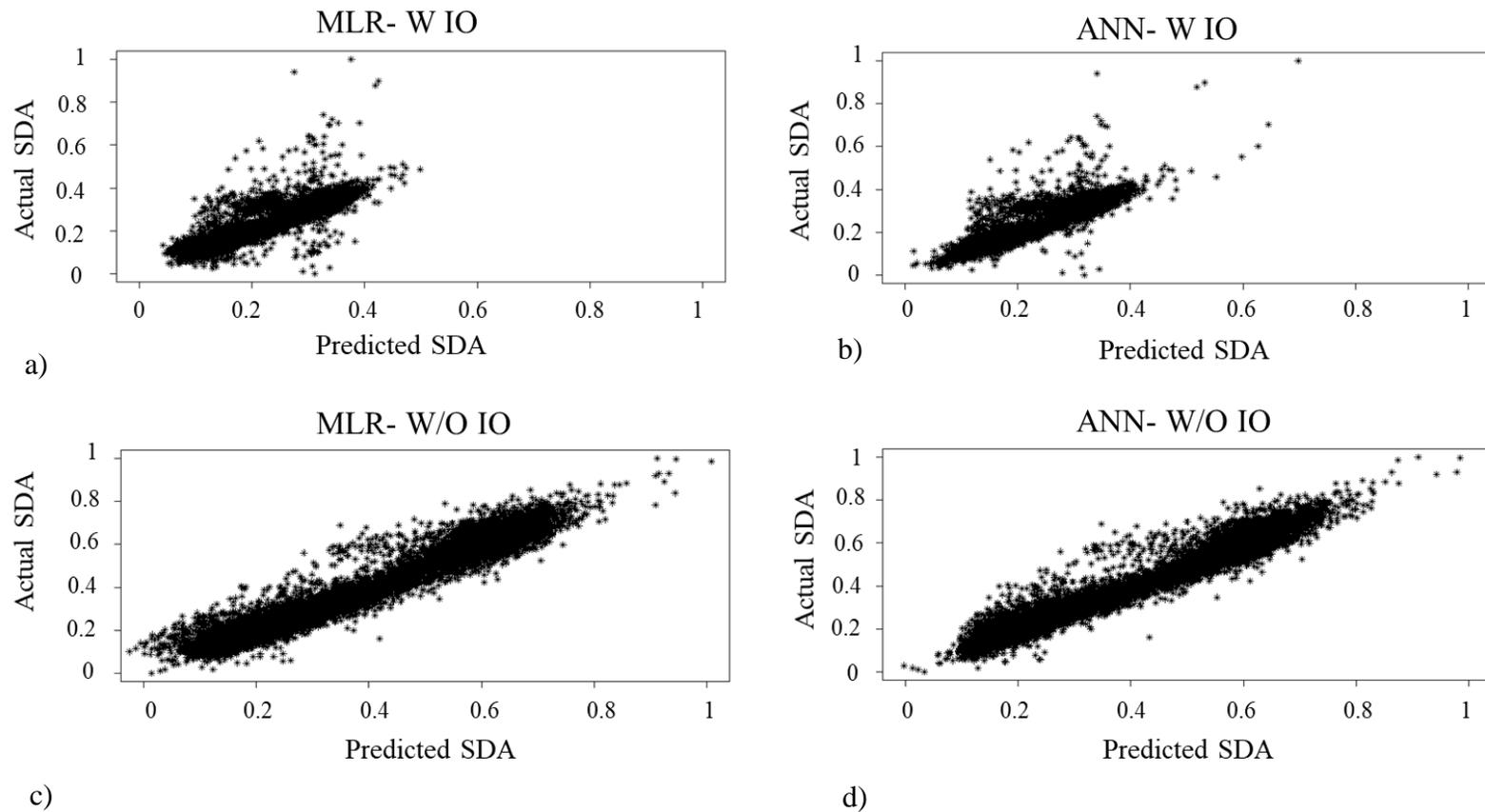


Figure 4.1: actual SDA against predicted SDA a) for MLR with (W IO); b) for ANN with (W IO); c) for MLR without (W/O IO); for ANN without (W/O IO). All SDA values are standardized ranging from 0 to 1.

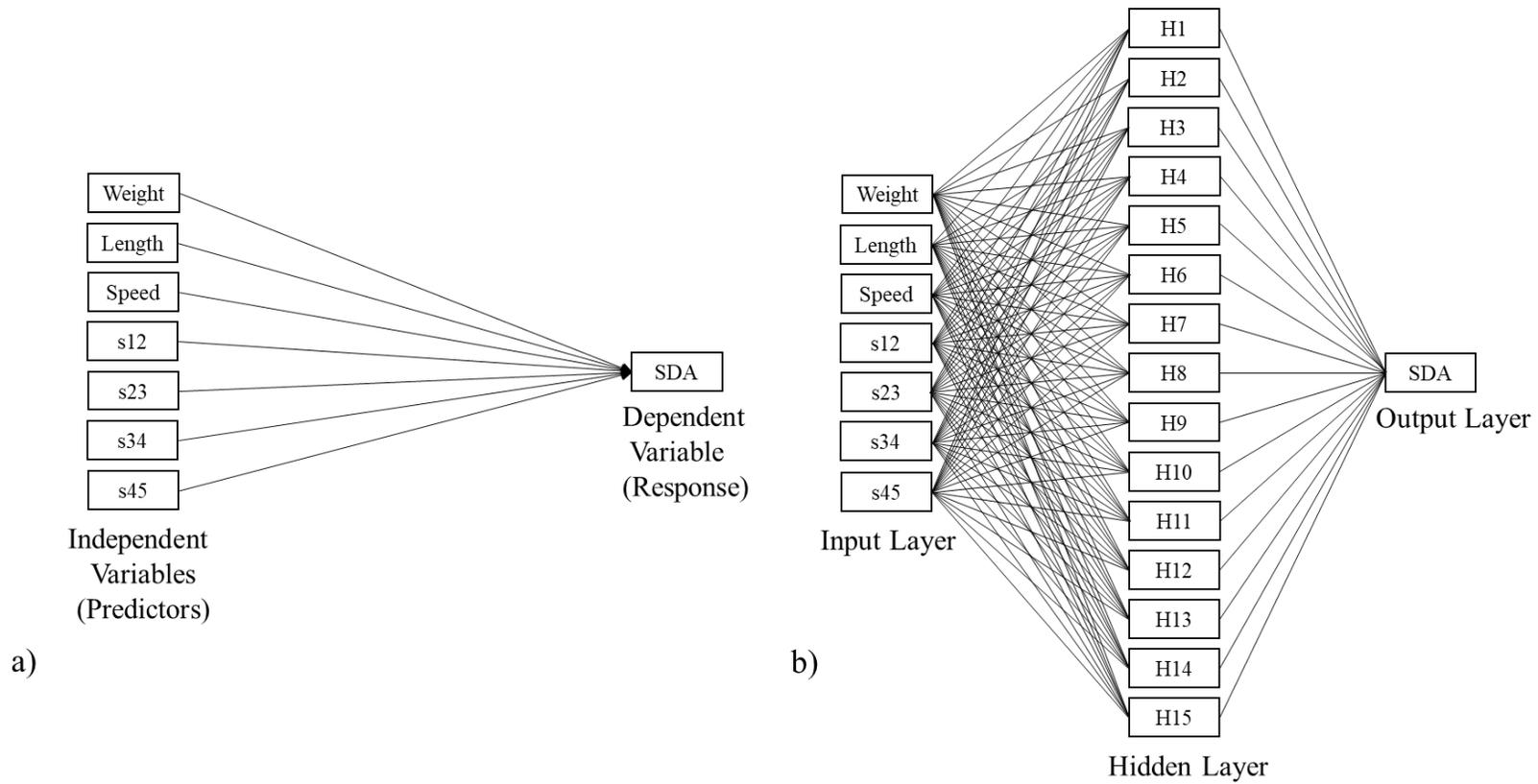


Figure 4.2: a) scheme of MLR; b) scheme of ANN.

4.3.2 Artificial Neural Network (ANN)

The ANN dataset preparation consists in standardizing the response and all the predictors as described in Sec. 4.2.2. Moreover, to avoid any subjectivity in choosing the best decay λ and size (number of hidden nodes) of the ANN, k-fold Cross Validation (CV) is used (Browne, 2000). CV is an empirical measure that does not involve any assumptions. It consists of randomly splitting the dataset into n folds where $1/n$ is left out (this method is also called leave-one-out) for the test data set while the remaining folds constitute the training dataset. At each repetition, a new data set is randomly chosen. For this case, a 10-fold CV, repeated 6 times, is used including all predictors: weight, length, speed, s12, s23, s34, s45.

The smallest value of root mean square error (RMSE) and the highest R^2 are used to select the optimal calibration parameters (Table 4.3), which are decay λ equal to 0.01 and size equal to 15. Fig.4.2b shows the scheme of the calibrated ANN model, which presents 1 input layer with 7 predictors, 1 hidden layer with 15 hidden nodes and 1 output layer with 1 response. The estimated coefficients are not reported because they lack a straightforward interpretation. To visualize the results, the predicted SDAs for the test data sets are plotted against the actual values (Fig.4.1b). The calculated R^2 is 0.88. ANN is considered more robust against outliers due to the use of a bounded function that tends to limit the influence on individual cases in comparison with the standard regression approaches (Kutner et al. 2004). This is also confirmed by the analyses shown later in Table 4.4, where improvement after removing influential observations is lower than in the other methods.

Table 4.3: ANN model calibration parameters: decay λ and size, based on the best R^2 and lowest RMSE.

λ	SIZE	RMSE	R^2
0.01	5	0.0332	0.873
0.01	10	0.0329	0.876
0.01	15	0.0327	0.877
0.1	5	0.0334	0.872
0.1	10	0.0332	0.873
0.1	15	0.0332	0.873
0.5	5	0.0339	0.868
0.5	10	0.0339	0.868
0.5	15	0.0338	0.868
1	5	0.0341	0.866
1	10	0.0341	0.867
1	15	0.0341	0.867

4.3.3 Regression Tree (RT)

RT model does not require any preparation of the dataset or predictors selection. The RT model chooses the most important predictors among GVW, length, speed, s12, s23, s34, s45. Their importance can clearly be noticed because only the most significant predictors are represented in the tree, they can be repeated multiple times and located higher in the tree plot. RT's explanatory capabilities are also accompanied by the possibility to handle non-linear relationships. K-fold CV is used to find the best tree size, by plotting the deviance against the tree size or complexity parameter λ , the minimum tree size corresponds to the lowest deviance value (Fig.4.3), which, from, in this case, is 5. Two sizes slightly bigger than the minimum are chosen to build the trees, size 8 (Fig. 4.4) and size 6 (Fig.4.5). From Fig.4.4 and 4.5 it is very intuitive which parameters are the most important because they appear first in the tree and because they are repeated multiple

times. The prediction is revealed at the end of the final branch reached by following a known path. As expected GVW is the most important predictor of SDA, and only secondarily (Table 4.2), a spacing variable (s23) has significant impact on the response. The goodness of fit measured by R^2 is 0.82.

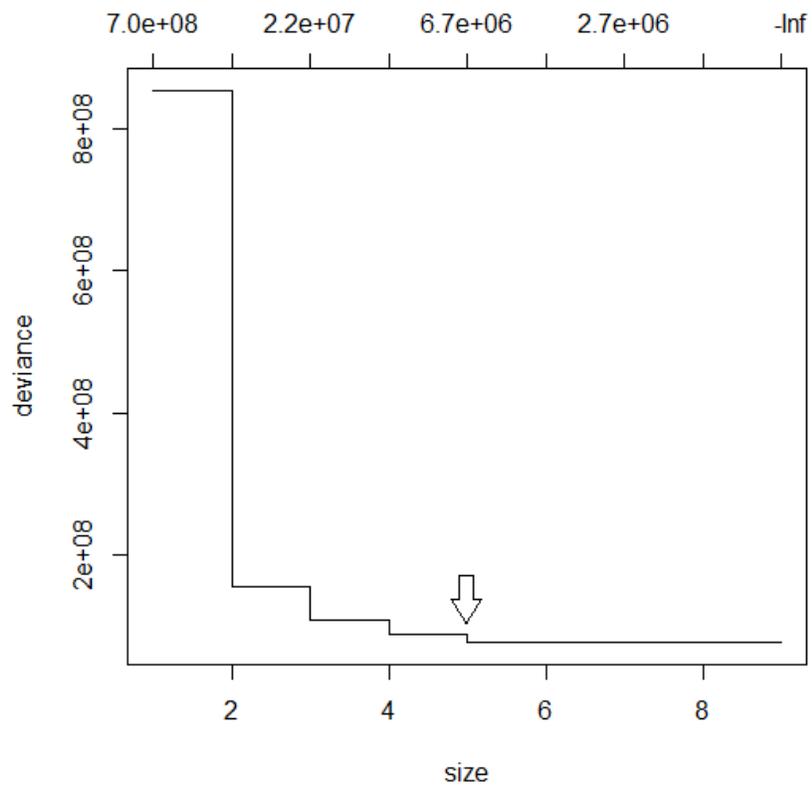


Figure 4.3: deviance vs tree size λ .

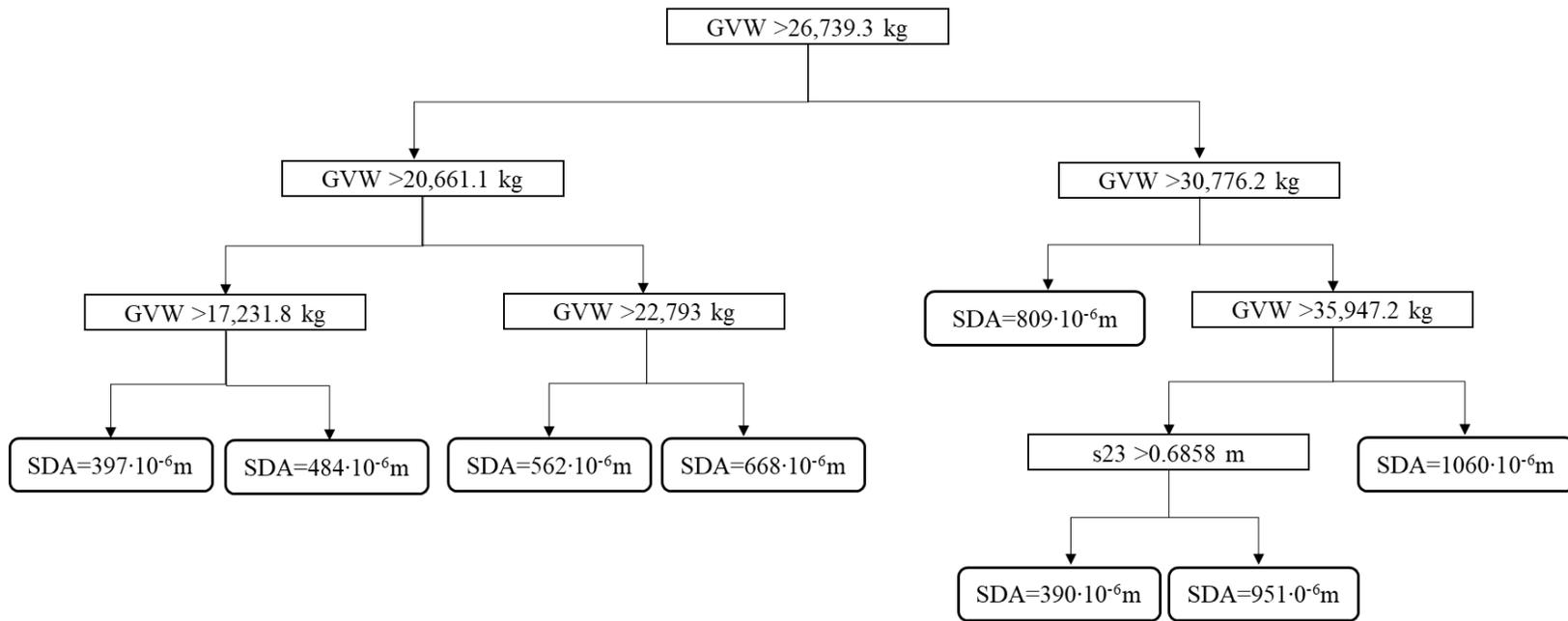


Figure 4.4: RT with 8 nodes.

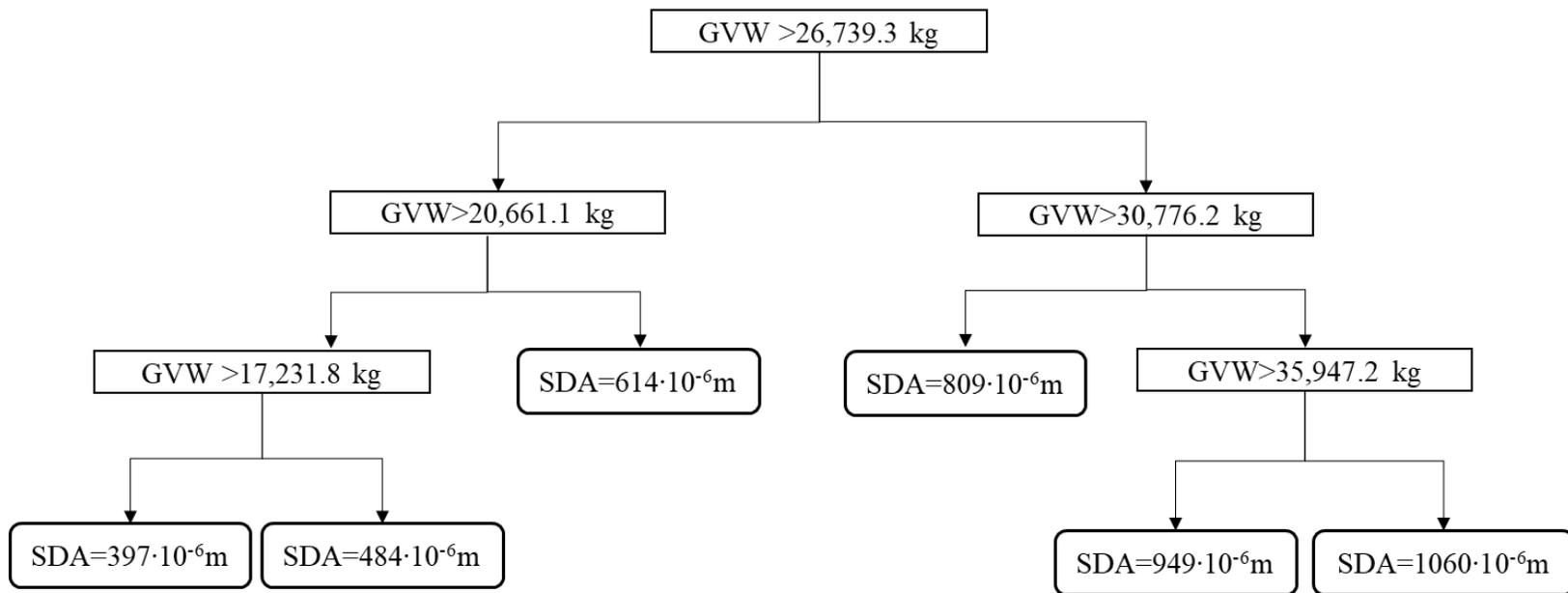


Figure 4.5: RT with 6 nodes.

4.3.4 Influential Observations Analysis (IOA)

To more rigorously analyze the predictive power of the MLR, the Cook's distance is computed and 592 (4%) influential observations are detected and discarded. The IOA resulted in a significant R^2 improvement, specifically from 0.85 (MLR) to 0.95 (IOA+MLR), showing more accurate β coefficients and marked improvement in predictions (Fig.4.1c). ANN is considered a robust method against outliers, for this reason, ANN is eligible for comparison with IOA+MLR. However, this does not imply that outliers do not have an effect on the fit of the ANN model. Indeed, IOA+ANN resulted in an improved R^2 (equal to 0.96 compared to 0.88 of ANN) and better-predicted values (Fig.4.1d). RT model is also run on the dataset with no influential observations resulting in an improved R^2 from 0.83 (RT) to 0.93 (IOA+RT).

Table 4.4: mean values of R^2 after cross-validation.

DATASET	ANN	MLR	RT
W/O IOA	0.88	0.85	0.83
W IOA	0.96	0.95	0.93
$\Delta_{\text{improvement}}$	0.08	0.1	0.1

4.4 DISCUSSION

The following four criteria are discussed to select the best model: predictive power, explanatory capabilities, interpretability, and computation speed.

- **Predictive Power**. In order to compare the goodness of fit of the three models, R^2 is computed by adopting a 3-fold CV, repeated 10 times, as opposed to splitting the dataset into test and training data. The mean of the 10 repetitions of R^2 is calculated and shown in Table 4.4. For consistency, standardized variables are used not only for ANN (where it is

necessary) but for all three models. Before removing the influential observations, all models already provided more than acceptable values of R^2 . However, on average, ANN shows better R^2 than MLR and RT, respectively 0.88, 0.85 and 0.83 (Table 4.4). The IOA led to a significantly improved R^2 for MLR and RT models (0.1, Table 4.4) and a lower improvement for ANN (0.08, Table 4.4). This confirms that ANN is more robust against outliers than the others. For this reason, it is legitimate to compare ANN to IOA+MLR. Unlike numerous studies in the literature, which state the superiority of ANN, this comparison reveals that the latter method ($R^2_{\text{IOA+MLR}}=0.95$) outperforms the former ($R^2_{\text{ANN}}=0.88$). When IOA is performed on both models, the gap between IOA+MLR and IOA+ANN is shrieked at 73% where the values of R^2 almost coincide ($R^2_{\text{IOA+ANN}}=0.956$ and $R^2_{\text{IOA+MLR}}=0.948$). RT also improves its predictive performance when influential observations are removed from 0.83 to 0.93 reducing the gap with IOA+ANN of 58% ($R^2_{\text{IOA+ANN}}=0.956$ and $R^2_{\text{IOA+RT}}=0.927$).

- Explanatory capabilities. It is well known that ANN is a machine learning black box whose output provides α and β coefficients which are used in sigmoidal functions and do not provide insight about the response. It is therefore difficult to link the magnitude of each coefficient to the predictor's importance. For MLR, instead, the p-values evaluate which predictor is statistically significant and the β coefficients suggest, on average, how much each 1-unit change would increase or decrease the response. Finally, RT provides information on the importance of predictors, by including them in the final tree, by means of the sizes of its branches and by locating the predictors in higher levels.

- Interpretability of results. ANN output lacks direct interpretability, additional plots and calculations must be done to get a sense of how the model is performing. MLR provides an output which is easy to understand. RT provides the most straightforward output, i.e. a tree that can be easily explained to non-experts by using “if...then...else” statements. RT’s output is also highly intuitive since it facilitates the ability to predict a new observation on the spot just by following the correct branches of the tree.
- Computation Speed. For ANN and RT, the user needs to make a decision on some calibration parameters. To prevent subjectivity, algorithms are used for this purpose. Due to the large size of the dataset, the time spent to calibrate the parameters for ANN is of the order of magnitude of hours. When the calibration parameters are set, running the model takes 1 second for RT and 51 seconds for ANN (time is rounded to the second). MLR does not require calibration and it also takes 1 second to run.

The four criteria are ranked in Table 4.5, from 1 to 3, with 1 being the best score, showing how the model works with and without IOA. In practice, the ranking mechanism presented here needs to be completed by also accounting for the different weights that different end-users and stakeholders assign to each criterion.

Table 4.5: ranking criteria, 1=Best.

CRITERIA	W IOA			W/O IOA		
	ANN	MLR	RT	ANN	MLR	RT
Predictive power	1	2	3	1	1	2
Explanatory capabilities	3	1	2	3	1	2
Interpretability results	3	2	1	3	2	1
Computation Speed	3	1	2	3	1	2

Notation: With Influential Observations Analysis (W IOA), Without Influential Observations Analysis (W/O IOA).

4.5 CONCLUSIONS

The first goal of this chapter is to propose a comparative study of three techniques, Multiple Linear Regression (MLR), Artificial Neural Network (ANN) and Regression Tree (RT), whose goal is to predict the bridge performance in terms SDA. Advantages and disadvantages of each method are discussed and ranked. To identify the model with the best predictive power, R^2 was calculated using CV. While all three models fitted the data well resulting in satisfactory values of R^2 , ANN displayed the highest R^2 (0.88), thus confirming its superior predictive power. MLR produced a valuable R^2 (0.85), and most notably it provided significant insights on the parameters that exert the strongest influence on SDA: GVW and spacing between axles. RT also showed satisfying values of R^2 (0.83), and similarly to MLR, it provided insights on the most meaningful predictors of bridge response, by confirming that primarily GVW and secondarily spacing between axles are the most important predictors. Nevertheless, multiple factors also need to be considered when comparing the performance of models. In particular, it was noted that the ANN's output lacks interpretability, thus not being the ideal choice if an investigation that has the primary goal of explaining the origin of certain response patterns. Another downside is the significant amount of time needed to calibrate the model's parameters. MLR results were obtained in a fraction of a second, its output was easily interpreted, and its β coefficients provided valuable insights on how the predictors drive the changes in the response. The results of RT were obtained in very short time, provided insight on the importance of the predictors, and were easily interpreted thanks to the user-friendly tree shape. These benefits are accompanied by the possibility to perform predictions on the spot by following the branches of the tree.

The second goal of this chapter is to propose an alternative predictive strategy which aims to avoid trading off predictive power for explanatory capabilities or interpretability and vice versa. For this purpose, Influential Observations Analysis (IOA) is employed. It was proved that IOA can drive improvements to all three methods. The IOA improves the predictive power of MLR which already has the potential to enjoy more support among practitioners given its elevated interpretability and explanatory capabilities. For this reason, the combination of IOA and MLR is here suggested as the alternative strategy to ANN, which is considered to be better suited to handle outliers in the basic estimation. In this case, IOA+MLR outperforms ANN (0.95 vs 0.88) and when influential observations are discarded for ANN as well, the gap between MLR and ANN is nearly eliminated (0.95 vs 0.96).

On the whole, this work's results show how to efficiently study bridge performance and guide interventions. Ultimately, this chapter demonstrates how statistical methods are critical to the process of distilling information out of bridge monitoring data, to aid infrastructure owners in evaluating bridge performance and making data-driven management decisions. While the subject bridge in this study is largely performing safely, the parameters of the IOA+MLR investigation could provide a metric that could be used for long-term monitoring. Shifts in the regression parameters could indicate a condition change on the bridge that warrants further investigation. A practical challenge in proof-of-concept research of this nature is that it is difficult to identify subject structures that will exhibit problems that can be detected with monitoring.

5 CLUSTERING BRIDGE RESPONSE AND TRUCK CLASSIFICATION

Information on truck data is important for many functions of maintaining the infrastructure and transportation network. These functions include pavement design and maintenance, enforcement, freight movement, traffic monitoring, air quality models, determining remaining life of critical fatigue details, tracking weight limits on posted bridges, and research. A two-step strategy is proposed to cluster and label the response of the bridge and to classify truck traffic. Longitudinal clustering is used to perform step 1, image processing and classification tree are used together or separately to perform step 2. This strategy aims to support decision makers with different management tasks such as traffic management policy, and regulation of heavy traffic.

5.1 INTRODUCTION

Assuming that all vehicles have a similar impact on bridges would lead to significant error in the management of a structure. Heavy and overweight trucks significantly contribute to the reduction of the service life of pavements and bridges (Fiorillo and Goshn, 2014). Having an accurate knowledge of the type and number of heavy trucks and the different responses of the bridge can be beneficial in planning effective maintenance interventions, in addition, to aiding policymakers and enforcement to promote an overall better use of the structure. As mentioned in the first pages of this thesis, the installation of the WIM and SHM systems was motivated by increased load limits on the US Highway 2 corridor including the subject bridge, where load limits for logging trucks were increased to promote transportation of timber from Michigan to a sawmill in Wisconsin. Specifically, per Michigan regulations, logging trucks loaded with greater GVW than Wisconsin are allowed to travel into Wisconsin along the US-2 corridor.

This chapter aims to propose a two-step strategy which consists of first, clustering the bridge response under different heavy trucks using longitudinal clustering and second, identifying the characteristics of the trucks which generate each specific response by using image processing and classification trees or either one depending on the type of information available. This data-driven strategy aims to support bridge managers and policymakers (from both transportation and structural engineering point of view), to regulate heavy traffic and maintain safe thresholds to avoid potential harm that they can cause to the bridge over time. Approximately 12,000 truck records and correspondent strain bridge responses, collected over the 3-year period, are analyzed. The recorded strain waveforms are clustered using a non-parametric technique belonging to the class of unsupervised learning methods: longitudinal k-means clustering (KmL) specifically

developed to identify similar patterns in longitudinal data by Genolini and Falissar (2011). The major challenge of the classification part of the study is not to classify vehicles with evident different features such as cars and trucks, but to classify all heavy trucks which belong to FHWA classes 9 to 13 (5 to 9 axles) with similar characteristics and often even similar axle configuration.

5.1.1 Technical Background

Traffic classification has always been an important area of research for infrastructure management. Previous studies on traffic classification are reported below.

- Gagarin et al., 1994 applied neural network to determine truck characteristics such as speed, axle spacings, and axle loads from strain-response records measured at the bottom flange of the bridge girder. Nine types of trucks (from 2 to 6 axles) were considered in the study.
- Sun et al., 2003 developed an inductive classifying artificial network (ICAN) for vehicle classification. This time the classification is not done by number of axles or spacing but into useful classes such as passenger cars, sport utility vehicles, vans, bus, and trucks.
- Erman et al., 2006 used two methods of clustering analysis to identify groups of traffic with similar characteristics. K-means and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) are adopted and their results are compared to AutoClass algorithm. They measured the accuracy of the clusters to classify a particular category of traffic with the DBSCAN having the highest precision for three out of four classes of traffic. They also compared the time to build the model which went from 4.5 hours for AutoClass to 1 and 3 minutes for K-Means and DBSCAN respectively.

- Shin et al., 2007 applied three machine learning algorithms: naïve Bayesian, neural network and support vector machine, with the last one outperforming the other two, to automatically classify vehicles in five classes from small vehicles to trucks. The data were collected from strain gages installed on bridge-deck panels. Principal component analysis was used for features extraction. When a vehicle was misclassified it was always incorrectly assigned to a class which neighbored the correct one.
- Yan et al., 2008 developed a strain-based vehicle classification using neural network which was compared to a manual classification developed by the authors. The dataset consists of strain records and videos from which they extracted using principal component analysis class, weight, and speed of the vehicles. They considered all vehicles from passenger car to trucks. The results obtained with neural network were compared with those obtained by using bayesian inference method.
- Fiorillo and Goshn, 2014 developed a data mining algorithm to analyze WIM data used to identify and classify overweight trucks into permit and illegal categories. They adopted a pattern recognition algorithm to cluster vehicles according to FHWA classification, number of axles, spacing and total axle length in four preliminary categories. Then, they applied Bayesian conditional probability to finally identify illegal trucks.

5.1.2 Longitudinal K-Means (KmL) in Other Fields

This section presents studies which used the KmL to perform the clustering. This machine learning technique is adopted in different fields where it is common to study longitudinal data such as criminology (Wheeler et al., 2016;), epidemiology (Hurault-Delarue et al., 2016), psychology and psychiatry (Pingault et al., 2013, Herba et al., 2015, Walton et al., 2017), sociology (Laurin et

al., 2015) neurology (Schels et al., 2013), endocrinology (Abdennour et al., 2014), and many other branches of medicine (Tanaka et al., 2017). Below some applications of KmL are presented:

- Mackelprang et al., (2012) clustered the level of exposure to HIV-1 to identify individuals with host resistance to the infection. High, low and decreasing exposure were the three clusters identified in the study. The individuals who belonged to the high exposure cluster but showed low incidence of HIV-1 infection were the host resistance. These individuals were further studied to identify their biological characteristics and understand what makes them resistant to the disease.
- Pingault et al., (2014) clustered the level of inattention across childhood to verify if the form of inattention could predict high school graduation failure. Stable, Fluctuating and increasing level of inattention were the three clusters identified by the study. The study concluded that increasing inattention level during elementary school made a significant contribution to graduation failure.
- Curman et al., (2014) clustered crime pattern trends to examine crime at the street block level. Very low, low, high and very high crime were the four clusters identified. Results showed how crime decreased over a 16-year period, they also identified some streets and intersections where the crime level is still high. This study aimed to support the local police in planning their interventions.

5.2 METHODOLOGY

5.2.1 Data Pre-processing

To develop the two-step strategy both SHM and WIM datasets were used. In the first step, the strain records collected by SHM system are used to perform the clustering analyses. Each truck

is represented by a waveform of 501 strain data-points (longitudinal dataset) resulting from 5 seconds of recording at 100 Hz (1 strain measurement every 10 ms), creating a longitudinal data set of strain measurements for each girder. For the reasons explained in paragraph 1.4.3, the average of the waveform of girder 3 and 4 is computed for trucks crossing the bridge on lane 1 and girder 2 and 3 for trucks crossing the bridge on lane 2. In the second step, the truck's characteristics (GVW, length, speed, axles weight, spacing, lane) recorded by the WIM system as well as the photos taken by the WIM cameras are used to classify trucks in both qualitative and quantitative to identify. The data processing consists of the following actions:

- i) Homogenization of raw strain records: each girder has its own zero point which changes due to the temperature and other environmental factors. In order to get rid of this bias and allow comparisons between different girders, during different time of the day and different periods of the years, all strain values are zeroed.
- ii) Conversion to longitudinal strain records: each truck was initially represented by a matrix of 501 rows and 6 columns (time, strain at girder 1-5), in order to have a longitudinal dataset each truck's matrix was transposed resulting in 5 rows (one for each girder) and 501 columns (from index 1 to 501). Then, the average of the strain waveform of girder 3 and 4 is calculated for trucks on lane 1 and girder 2 and 3 for trucks on lane 2. After taking the averages of the strain records, each truck is finally represented by a longitudinal vector [1, 1:501].
- iii) Data selection: only trucks of class 9 to 13 are considered for the analysis. The two- step strategy is fully tested on March 2012, lane 2 dataset because of the availability and

quality of photos taken by the WIM cameras. However, to assure a statistically significant sample size for most of the analyses only lane 1 trucks are considered.

- iv) Basic Data Normalization and Outlier Detection: to reduce the influence of speed on the clustering analysis, the longitudinal strain response data are normalized for direct comparison. To do so, the peak of each strain response is centered and the whole strain response is shifted accordingly. A 4-condition outlier detection strategy is adopted. As shown by Eq.41 the goal is to capture first, null GVW and length resulting from WIM reading malfunction; second, trucks with GVW less than 9 ton (20 kips) which were erroneously classified BY THE WIM as class 9; third, trucks with a summation of all axle weights not within 10% of GVW; fourth, data entries with an additional waveform on the left or right of the strain truck's response. This last condition is performed by taking the absolute value of the mean of the first ten strain readings (left side of strain response) and the absolute value mean of the last ten strain readings (right side of strain response), data entries with either of these values greater than $1 \mu\epsilon$ were removed from the dataset. This method tries to discard most of the cases of multiple vehicles on the bridge.

$$\text{outlier detection: } \left\{ \begin{array}{l} GVW = 0 \vee length = 0 \\ GVW < 9 \text{ ton (20 kips)} \\ \sum_{i=1}^{\#axles} axle\ weigh_i > 1.1 \cdot GVW \vee \sum_{i=1}^{\#axles} axle\ weigh_i < 0.9 \cdot GVW \\ \left| \frac{1}{10} \cdot \left(\sum_{i=1}^{10} \mu\epsilon_i \right) \right| > 1 \vee \left| \frac{1}{10} \cdot \left(\sum_{i=491}^{501} \mu\epsilon_i \right) \right| > 1 \end{array} \right\} \quad (41)$$

- v) Time-Window Data Normalization and Outlier Detection: an alternative combination of normalization and outlier detection is also considered (Eq.43). The first three conditions of Eq.41 are kept, and an alternative approach to treat the presence of multiple vehicles is adopted. Instead of discarding those trucks, a temporal window called time-to-cross (*ttc*) is computed using the speed, the length of the truck and the length of the bridge (36 m, 120 ft). This normalization includes cases of multiple vehicles on the bridge, but it just considers the response to the truck. By considering only the time-window necessary to cross the bridge, only the truck's waveform is included. This method is clearly not immune from errors so all the trucks with an extremely slow recorded speed are discarded, because this would mean that the truck needed more time than the 5-second recording interval to cross the bridge, in practice, it is possible that either the speed reading was wrong or special machinery with a very low speed were crossing the bridge for snow removal or maintenance. Once the time-window is identified the peak values are again centered. It is also presented a variation of this normalization which does not align the peaks but the time-frame the truck is entering the bridge (left SIDED normalization). The imperfection of the time-window normalization is due to the change of speed which the truck might have while crossing the bridge, therefore, the *ttc* is not completely accurate and predictable.

$$ttc = \frac{L_{bridge} + l_{truck}}{speed} \quad (42)$$

$$\text{outlier detection: } \left\{ \begin{array}{l} GVW = 0 \vee length = 0 \\ GVW < 9 \text{ ton (20 kips)} \\ \sum_{i=1}^{\#axles} axle\ weigh_i > 1.1 \cdot GVW \vee \sum_{i=1}^{\#axles} axle\ weigh_i < 0.9 \cdot GVW \\ ttc < 5 \text{ sec} \end{array} \right\} \quad (43)$$

5.2.2 Longitudinal Clustering

K-Means is an algorithm commonly used for clustering unlabeled data. The goal of K-Means clustering is to group data into clusters with similar characteristics. The R package (2016) *KmL* is used to perform the longitudinal clustering. This process provides a way to analyze trends and patterns in a dataset that may not be through traditional grouping and comparisons. Each strain waveform is called single trajectory or trajectory, several trajectories, which made a dataset, are called joint trajectories and can be written as a matrix where, in this case, each row is a single strain waveform (Genolini et al., 2015). K-Means works by computing the distance between two objects, in this case, it would be the distance between two matrices. Considering a set S of n objects, for each object an outcome variable Y at different time t is measured. y_{il} is the value of Y for the object i at time l and the sequence $y_i=(y_{i1}, y_{i2}, \dots, y_{it})$ is the trajectory. The goal of clustering is to divide a set S in k homogeneous subsets by assigning each trajectory to the nearest cluster by calculating the various distances, this case uses the Euclidean distance E defined as follows:

$$Dist^E(y_i, y_j) = \sqrt{\sum_{l=1}^t (y_{il} - y_{jl})^2} \quad (44)$$

The challenge consists of choosing the best number of clusters. The Calinski-Haravbatz (1974) criterion is used to select the optimal number of clusters by combining the within-clusters and between-clusters covariance to evaluate the quality of the partition. Defining n_m as the number of

trajectories in cluster m , \bar{y}_m is the mean trajectories of clusters m , \bar{y} is the mean of the whole set

S. The between-clusters covariance matrix B is defined as follows:

$$B = \sum_{m=1}^k n_m (\bar{y}_m - \bar{y})(\bar{y}_m - \bar{y})' \quad (45)$$

Where $\text{trace}(B)$ is the sum of the diagonal coefficients of B . High values of $\text{trace}(B)$ indicate well-separated clusters, low values of $\text{trace}(B)$ indicate that the clusters are close to each other.

The within-cluster W covariance matrix is defined as follows:

$$W = \sum_{m=1}^k \sum_{l=1}^{n_m} n_m (y_{ml} - \bar{y}_m)(y_{ml} - \bar{y}_m)' \quad (46)$$

Where $\text{trace}(W)$ is the sum of the diagonal coefficients of W . High values of $\text{trace}(W)$ indicate heterogeneous clusters, while low values of $\text{trace}(W)$ indicate compact clusters, which is the ideal scenario. The optimal number of clusters corresponds to the value k which maximize the following expression:

$$C(k) = \left(\frac{\text{trace}(B)}{\text{trace}(W)} \right) \cdot \left(\frac{n - k}{k - 1} \right) \quad (47)$$

Overall, the best solution is the one which maximizes the between-matrix variance and minimizes the within-matrix variance. It is worthy to note that even if this criterion helps find the optimal number of cluster, it was shown that not always find the correct solution (Shim et al., 2005). The first step of the K-Means algorithm is to choose an initial configuration. Different methods with different starting conditions are considered to choose the initial configuration:

- RandomAll: all individuals are randomly assigned to a cluster and the mean of each cluster is the cluster center.

- RandomK: k individuals are chosen randomly, they are the initial cluster center not an average of several individuals. The other individuals are not assigned. This method produces initial conditions which are not close to each other which eventually might be from different clusters, this method will speed up the convergence.
- maxDist: this is an incremental method. The matrix of the distances between all individuals is computed, the two farthest individuals are chosen to be the two centers c_1 and c_2 , the following individuals are added one at a time and they are the farthest individuals from those already selected. The farthest individual is the one with the greatest distance from the selected individuals.
- allMethods: it combines all three methods above.

The starting condition chosen by default uses allMethods for $k=2, 3, 4, 5,$ and 6 clusters 20 times each. Fig.5.1 shows how the user can visualize the options and see which one is selected as the best according to the Calinski-Harabatz criterion, also called “active criterion”. The x-axis is the number of iterations (20) and the y-axis is the value of the specific quality criterion, $C(k)$ (Eq.47). In the plot, the partitions with the same cluster k are ordered in decreasing order, the best coming first, for all the partitions the best is selected with a black dot. This plot is useful, not only, because it suggests the best option according to the selected criterion, but also, it provides an overview of how all the other cluster-options perform, and, since quality criteria are not always efficient, this may help the user to support a cluster-choice which is different from the one selected by the active criterion (Genolini and Falissard, 2011).

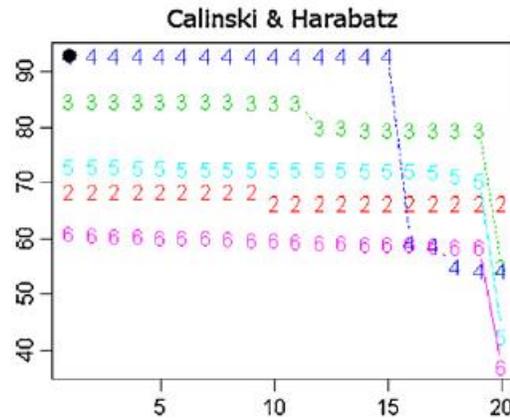


Figure 5.1: selection of the number of clusters according to the Calinski-Harabatz quality criterion.

5.2.3 Truck traffic classification

Clustering bridge response is step 1 of the two-step strategy. Step 2 starts with the clustering output and aims to define which category of trucks is more likely to produce each bridge response. Two methods are here proposed: manual image processing and classification tree. The two methods provide two different types of truck classification based on the different set of information they process. The former aims to classify the truck's type by extracting this information from the images taken by the WIM cameras, the latter, aims to classify the trucks based on their characteristics captured by the WIM system (GVW, length, speed, class axle weight, spacing). The combination of both methods makes a strong asset in the truck-classification, however, the two-step strategy still works with either method.

5.2.3.1 Image processing

The image processing method consists of manual analysis of the photos taken at each truck crossing by the WIM's cameras. This method can be automatized using artificial intelligence

algorithms when the volume of images to visualize and classify is significant. The manual analysis included the additional searches such as the carrier's company when visible to gain more information on the load carried. Since the photos only picture the front of the truck, the axle configuration information from the WIM system are pivotal in the truck classification.



Figure 5.2: day-time photo and night-time photo, March 2012, lane 2.

5.2.3.2 Classification Tree

Fitting and using a classification tree with k -category response is similar to fitting and using a regression tree, described in paragraph 4.2.3 of this thesis. For classification trees, the probability of the response Y being the category or class k ($k=1,2,\dots,K$) given the predictors \mathbf{x} is modeled as a constant over each region R_m .

$$p_k(\mathbf{x}) = \Pr\{Y = k|\mathbf{x}\} \quad (48)$$

The fitted class probability and the best class prediction, which is the most common class in region R_m , are respectively shown in Eq.49 and Eq.50.

$$\hat{p}_{m,k} = \frac{1}{N_m} \cdot \sum_{x_i \in R_m} I(y_i = k) \quad (49)$$

$$k_m = \arg \max_k \{\hat{p}_{m,k}\} \quad (50)$$

At each step in the fitting algorithm, the best split is the one that most reduces the “impurity” within the regions. Different impurity measures can be used, such as misclassification error or Gini index, however, the one that most software adopt is the deviance, already mentioned in chapter 4, also called “-2 log-likelihood”. High deviance indicates more impurity. CV-deviance plot, as the one shown in Fig.4.3, is used to choose the best size of the tree is again used the tree.

$$deviance = -2 \sum_{m=1}^M N_m \sum_{k=1}^K \hat{p}_{m,k} \log(\hat{p}_{m,k}) \quad (51)$$

5.3 ANALYSIS STRATEGY (PROOF OF CONCEPT)

Studies have shown that the GVW, axle weight, and axle configuration of heavy trucks directly affect the service life of highway bridge superstructures. Damage typically occurs in the bridge deck and in the main superstructure elements, including floor beams and girders, diaphragms, joints, and bearings. With the rapid growth of highway transportation, the increasing frequency of passing heavy trucks contributes to fatigue damage (FDOT, 2000). The two-step strategy gives a complete picture of the situation regarding the heavy traffic which regularly crosses the Hurley

bridge for different activities. Bridge owners, bridge managers, policymakers and more in general decision makers can all benefit from this data-driven strategy.

This section aims to show how each component of the two-step strategy works and the contribution that each step provides to reach the final goal of understanding bridge performance under heavy traffic and subsequent truck-traffic classification.

STEP 1: Clustering of Bridge Response (iteration 1)

The KmL is used on March 2012, lane 2. The dataset is normalized according to Eq.41 with a final sample size of 237 trucks. Fig.5.3 and 5.4 show the output of the clustering analysis where the y- axis, identified in the plots by “V”, is the strain value ($\mu\epsilon$). 145 trucks (61.2%) fall into cluster A, 55 trucks (23.2%) into cluster B and 37 trucks (15.6%) into cluster C. Fig.5.3 shows the three clusters resulting from averaging all the waveforms assigned to each cluster. Fig.5.4 shows on the left, the Calinski-Harabatz quality criterion, which suggests that 3 is the optimal number of clusters and the second-best option is 4 clusters, in the center, the clusters which overlap the waveforms, and, on the right, the waveforms color-coded according to the respective cluster.

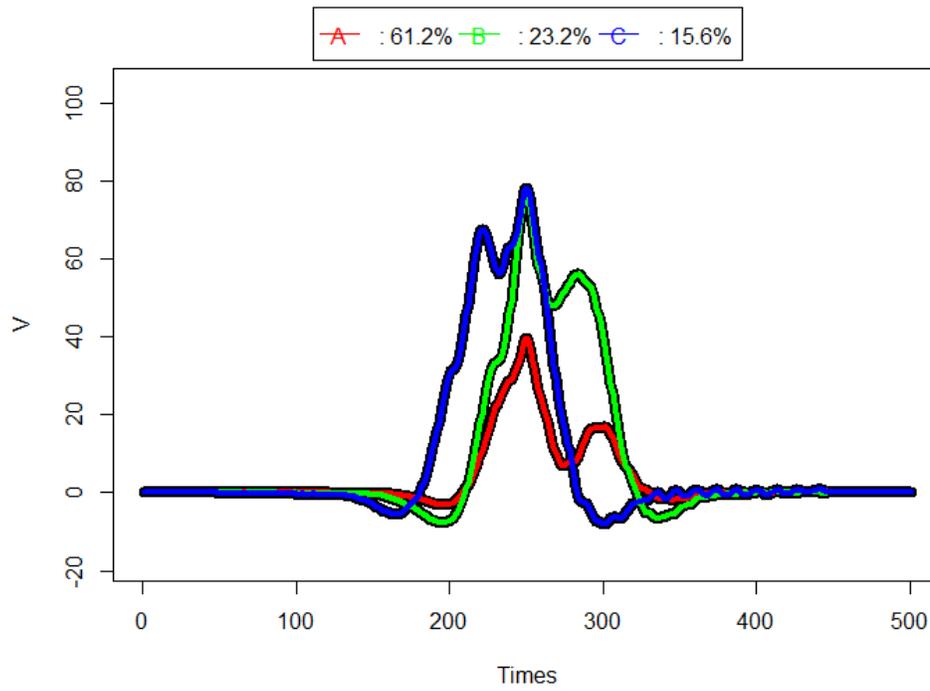


Figure 5.3: clustering output, 3 clusters.

**Calinski.Harabatz
Sorted**

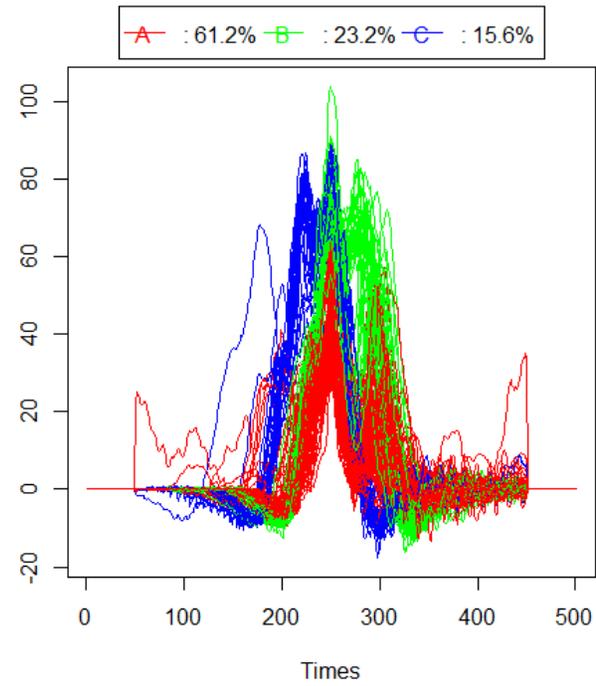
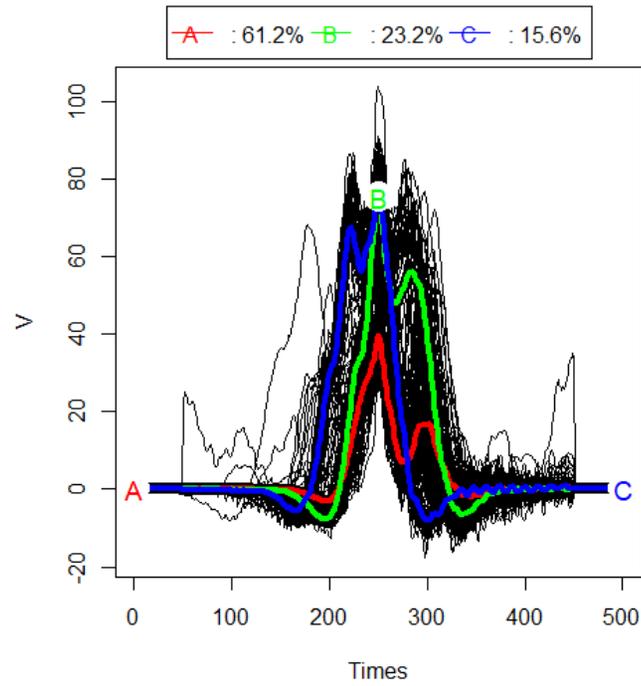
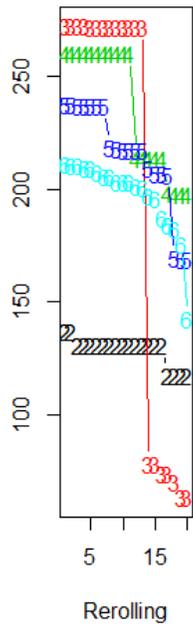


Figure 5.4: from left to right, Calinski-Harabatz criterion (y-axis is the $C(k)$ of Eq.47, the x-axis represents the 20 iterations); 3 clusters and waveform; waveforms color-coded by cluster.

STEP 2: Truck Classification - Image Processing

Photos taken between March 27 and March 31, 2012, of trucks crossing the bridge on lane 2 are envisioned to understand which type of truck belongs to each of the three clusters identified in step 1. Since three clusters are defined as the best option by the KmL method, the images are, therefore, divided in clusters A, B and C. The images are crucial to identifying trucks with specific characteristics such as logging trucks or tank trucks. Some challenges were due to the fact that the image only pictures the front of the truck not its whole length and no information is available on whether or not all the other regular trucks were traveling empty or just carrying a light load. It noteworthy to mention that the data were collected under live traffic conditions and not using surveys. 43 pictures are analyzed, and they are respectively clustered as summarized in Table 5.1.

Table 5.1: 43 trucks clustered.

Clusters	A	B	C
Number of trucks	29	9	5
Percentage of trucks	67.44%	20.93%	11.63%

The pictures revealed the following results which are summarized in Fig.5.5 and Table 5.2:

Cluster A includes trucks with a lighter GVW, always below the 80 kip-limit, a bimodal waveform due to the presence of a large spacing (25-30 ft), usually between axle 3 and 4 and ZTP which never exceeds 50 $\mu\epsilon$. They are:

1. Tank trucks with $GVW \leq 40$ kip (18 ton), Class 9, with identical axle configuration.

2. Flatbed, empty or light weighted, with $GVW \leq 50$ kips (~22 ton), or just slightly above 50 kips; class 9 (5 axles) and 10 (6 axles). Sometimes they can be confused with empty logging trucks, however, they do not have a typical configuration of that type of truck.
3. Regular trucks (dry-van or refer) with $GVW < 50$ kips, suggesting an empty or partial cargo or light weight cargo, class 9.
4. Regular trucks (dry-van or refer) with GVW between 50 and 70 kips, class 9 and 10.

Cluster B includes two different waveforms, a bimodal one, and a non-bimodal one, very heavy trucks, often overweight, ZTP is therefore very high as well. They are:

1. Full logging trucks, their GVW is about 100 kips or more, class 12 (6 axles), axles more equally spaced although there is at least one larger spacing between axle 3 and 4 of 12-18 ft (3.7-5.5 m), they present a non-bimodal waveform and $ZTP \geq 80 \mu\epsilon$.
2. Regular trucks (dry-van or refer) with high GVW , sometimes overweight but never reaching 100 kips, class 9, they present a bimodal waveform due to the large spacing, >25 ft (>7.6 m).

Cluster C includes logging trucks, with a non-bimodal waveform, extremely heavy trucks, always overweight, ZTP is therefore very high as well.

1. Full logging trucks, their GVW is around 100 kips or more, class 12 (6 axles), axles more equally spaced although there is still a large spacing is between axle 3 and 4 of (12-18 ft, 3.7- 5.5 m), they present a non-bimodal waveform and $ZTP \geq 80 \mu\epsilon$.

Table 5.2: summary of trucks characteristics within each cluster.

Cluster	Truck type	Weight	Waveform	Class
A1	Tank trucks	Light	Bimodal	9
A2	Flatbed	Medium- Light	Bimodal	9-10
A3	Regular trucks	Light	Bimodal	9
A4	Regular trucks	Medium	Bimodal	9-10
B1	Logging trucks	Overweight	Non-Bimodal	12
B2	Regular trucks	Heavy	Bimodal	9
C	Logging trucks	Overweight	Non-Bimodal	12

A1



A2



A3



A4



B1



B2



C



Figure 5.5: sample images of trucks by cluster.

Table 5.2 shows two that two clusters contain the same type of truck, they are cluster B1 and cluster C. This is believed to be a consequence of the normalization process which, as described in section 5.2.1, takes the peak of each waveform and aligns them to the center (Fig.5.6) to reduce the influence of speed.

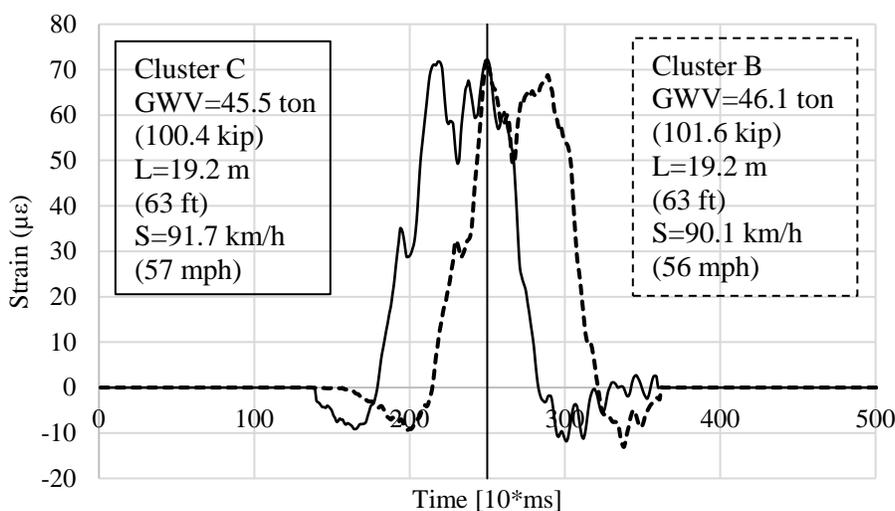


Figure 5.6: two identical logging trucks belonging to two different clusters.

To prove that those two clusters can contain waveforms generated by similar trucks, two identical logging trucks from the same company are studied in depth (Fig.5.7). To do so, 18 characteristics of the two trucks are listed in Table 5.3 to show that there are no significant differences to highlight. The question is what is causing those two identical trucks to be separated into two different clusters. The clustering technique is an unsupervised machine learning method where clusters are defined based on the strain waveform only. In this specific case, the peak's position is what influences the normalization, and, as it can be observed in Fig.5.6, the peaks can be multiple and can differ from one another of extremely small quantities, this might be due to small dynamic effects such as discontinuity in the pavement or truck's suspension. Since the road is the same for

both trucks which are crossing the bridge within 24 hours, it is highly possible that the suspensions of the trucks, in particular, their type, age, and wear, can cause that dynamic fluctuation in the waveform. It is crucial to be aware of a phenomenon which, although not measurable, and therefore uncontrolled, can play a small role in the explanation of the results.

Table 5.3: comparison of 18 truck's characteristics.

	Imperial System		Metric System	
	B	C	B	C
GVW	101.6 [kip]	100.4 [kip]	46.1 [ton]	45.5 [ton]
Length	63 [ft]	63 [ft]	19.2 [m]	19.2 [m]
Speed	56 [mph]	57 [mph]	90.1 [km/h]	91.7 [km/h]
Class	12	12	12	12
Axles	6	6	6	6
Axle 1 weight	12.1 [kip]	12.2[kip]	5.5 [ton]	5.5 [ton]
Axle 2 weight	17.3 [kip]	17.1[kip]	7.85 [ton]	7.76 [ton]
Axle 3 weight	17.2 [kip]	18[kip]	7.8 [ton]	8.16 [ton]
Axle 4 weight	17.3 [kip]	17.5[kip]	7.85 [ton]	7.94 [ton]
Axle 5 weight	19.1 [kip]	18.1[kip]	8.66 [ton]	8.21 [ton]
Axle 6 weight	18.4 [kip]	17.4[kip]	8.35 [ton]	7.89 [ton]
Spacing 1-2	15.7 [ft]	15.8[ft]	4.79 [m]	4.82 [m]
Spacing 2-3	4.3 [ft]	4.2[ft]	1.31 [m]	1.28 [m]
Spacing 3-4	18.2 [ft]	18.2[ft]	5.5 [m]	5.5 [m]
Spacing 4-5	10 [ft]	10[ft]	3.1 [m]	3.1 [m]
Spacing 5-6	10.3 [ft]	10.3[ft]	3.14 [m]	3.14 [m]
ZTP (ave_g3, g4)	72.2 $\mu\epsilon$	72.25 $\mu\epsilon$	72.2 $\mu\epsilon$	72.25 $\mu\epsilon$
SDA (ave_g3, g4)	4207.3 [ft]	4287.5 [ft]	1282.4 [m]	1306.8 [m]



Figure 5.7: (above) March 29 at 23:19:18, cluster B; (below) March 28 at 22:05:07, cluster C.

The image processing provides the important contribution of assigning a truck typology to each cluster. This step can also provide a pivotal contribution to the identification of the best number clusters. Based on these results shown in Tab.5.3, it seems that cluster B can be split into two clusters, therefore, the results might be suggesting that 4 clusters would better group the dataset. Also, the option with 4 clusters was the second-best option according to the quality criterion Calinski-Harabatsz. Therefore, a new iteration of step 1 is performed.

STEP 1: Clustering of Bridge Response (iteration 2)

The best option identified in the first iteration of step 1 was 3 clusters, however, a deeper analysis with the support of photos suggests that 4 clusters might be more representative of more truck types crossing the bridge. The results of the clustering analysis considering 4 clusters are shown in Fig.5.8 and 5.9.

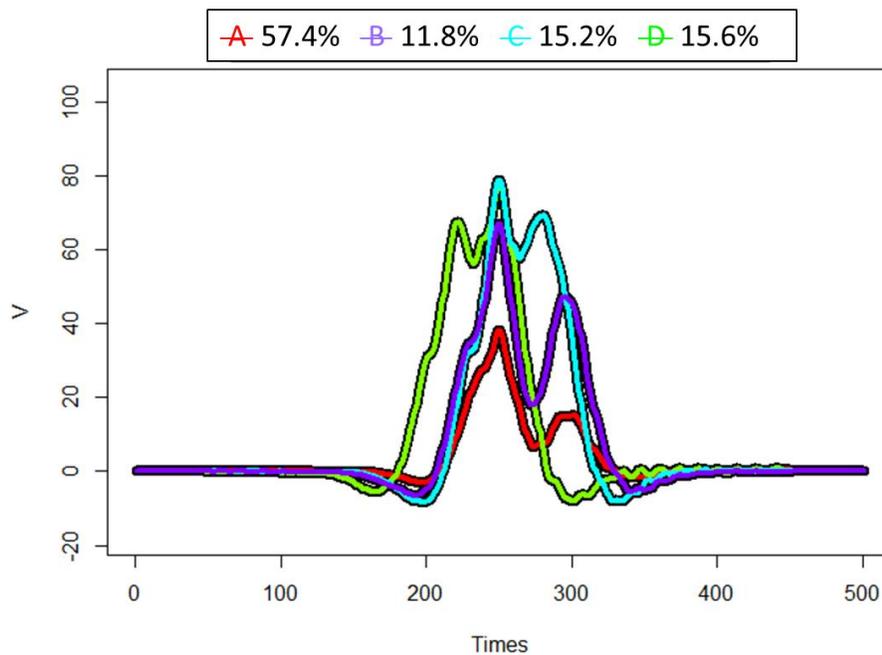


Figure 5.8: clustering output, 4 clusters.

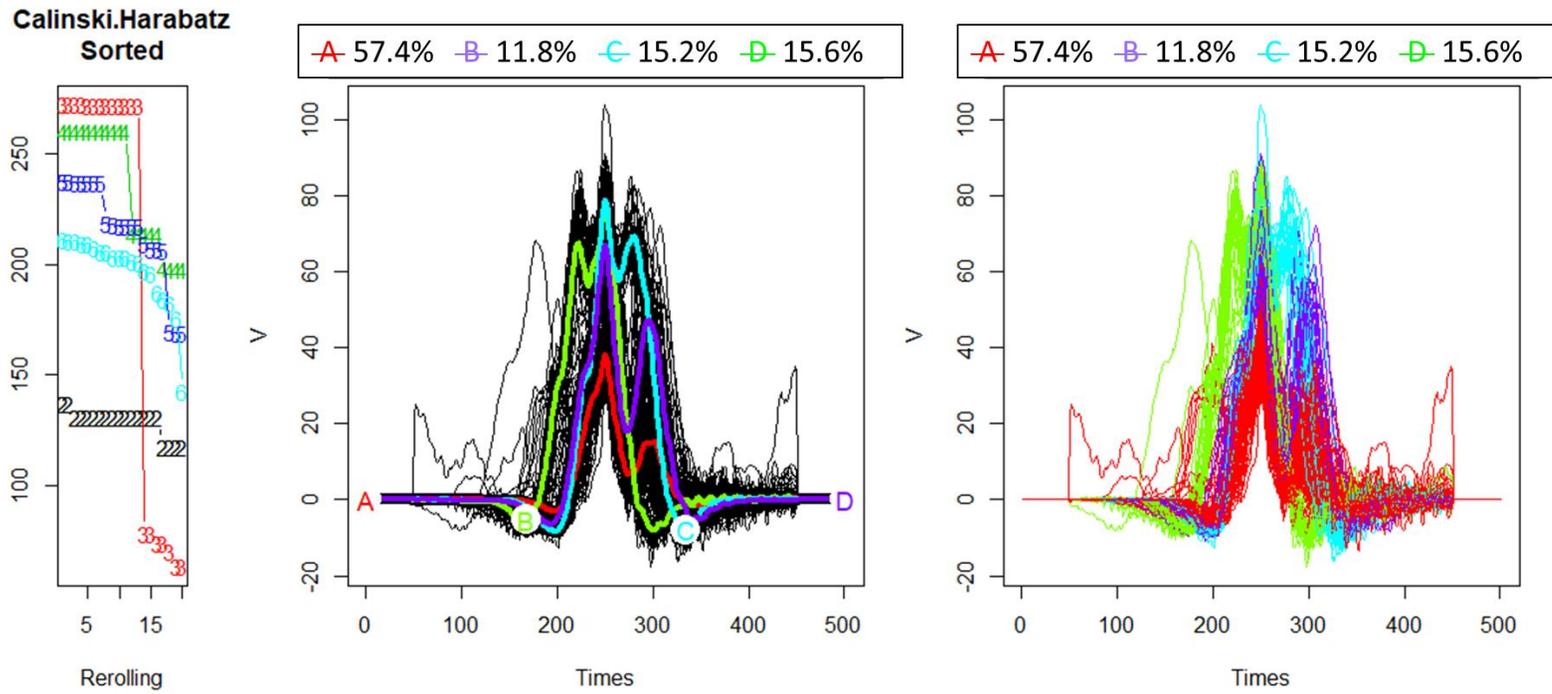


Figure 5.9: from left to right, Calinski-Harabatz choice method, 4 clusters, waveforms color-coded by cluster.

The clustering analysis in the second iteration confirms the presence of two different types of trucks in cluster B (B1: logging trucks, B2: heavy bimodal trucks), which are then split into B (heavy bimodal) and C (logging non-bimodal) in the 4-cluster option. Type A4 is also reassigned from cluster A (3-cluster option) to cluster B (4-cluster option), this makes sense because type A4 was a borderline group of trucks, with higher GVW compared all the other trucks belonging to the same cluster (A) and it is assigned to cluster B with other heavier trucks causing a bimodal response. The better reassignment of trucks into 4 clusters is shown in Fig.5.10 and Table 5.4.

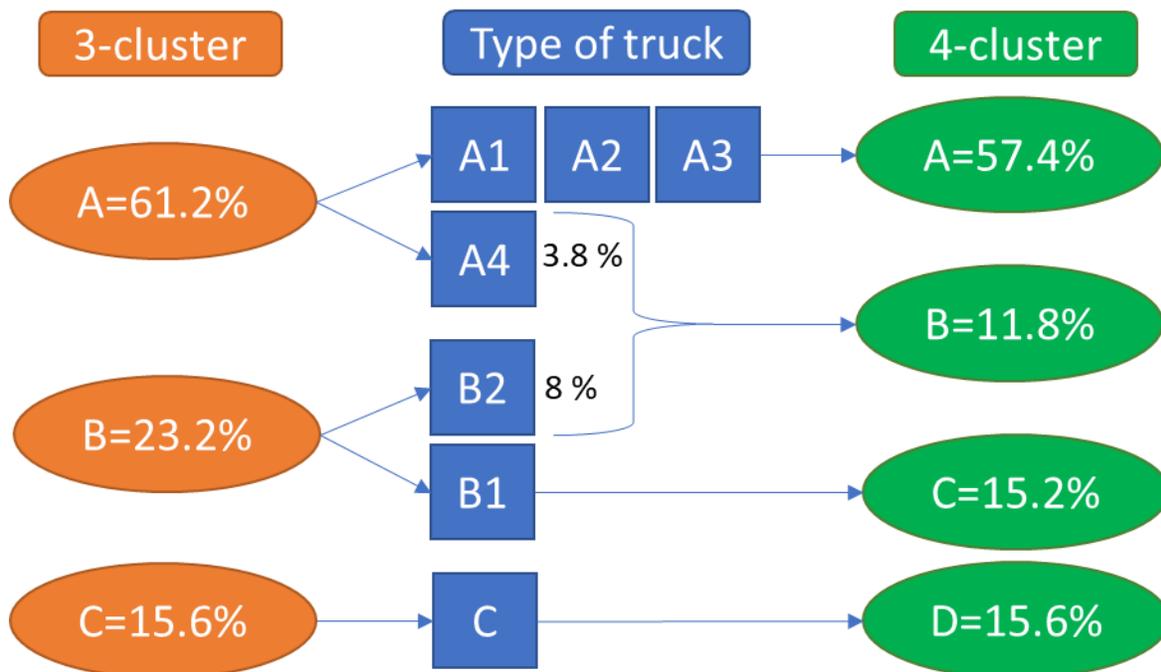


Figure 5.10: cluster reassignment.

Table 5.4: reassignment (4 cluster-option).

March 2012, lane 2		4 CLUSTERS			
		A	B	C	D
3 CLUSTERS	A	136	9	0	0
	B	0	19	36	0
	C	0	0	0	37

STEP 2: Truck Classification - Classification Tree

Classification trees, unlike the study of images which provides a qualitative classification of the truck-traffic, provide a quantitative classification based on the truck's characteristics measured by the WIM system. The tree identifies which predictors are the ones which most influence the categorical response $y=\{A, B, C, D\}$

When 3 clusters are considered (step 1, first iteration) (Fig.5.11), the classification tree identifies GVW and the weight of axle 6 as the most significant predictors of the response. It does not come as a surprise that the predictors which most influence the cluster-category, which groups similar strain waveforms, are related to weight. The first split is GVW which divides lighter trucks (A) from heavy trucks (B and C), the second split provides an important piece of information to help support the controversial distinction of logging trucks belonging to clusters B and C. According to the tree, trucks in cluster C show a heavier load in the rear axle 6, resulting in a peak shifted to the right (rear of the truck), while cluster B shows its peak right after the first axle.

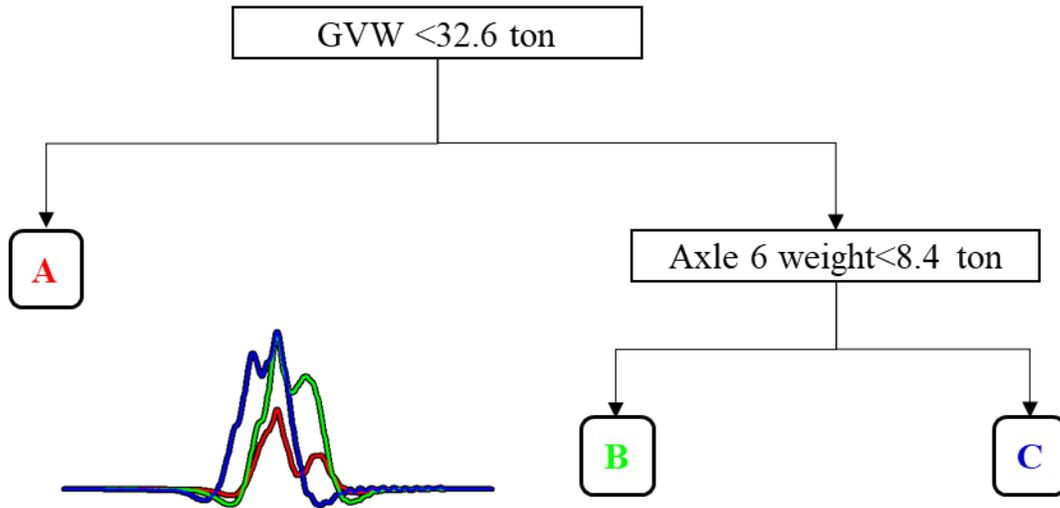


Figure 5.11: classification tree, 3 clusters.

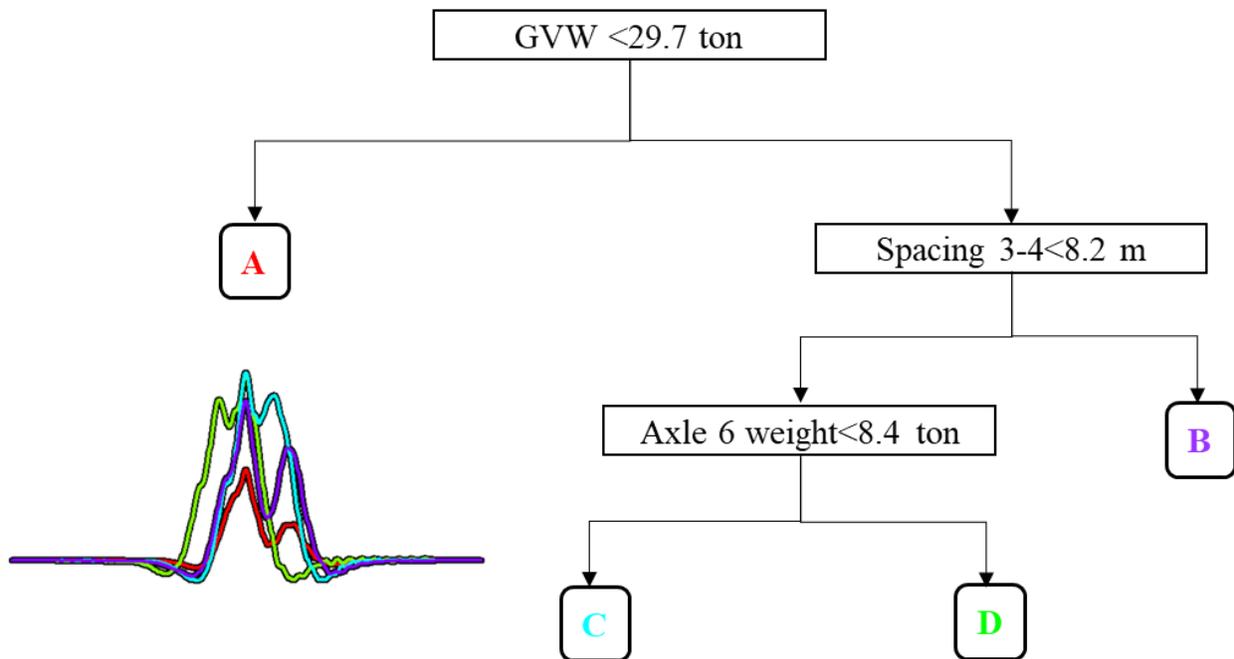


Figure 5.12: classification tree, 4 clusters.

When 4 clusters are considered (step 1, second iteration) (Fig.5.12), the tree first, separates the trucks in heavy and lighter trucks (respectively A and B, C, D), then, it separates heavy trucks

according to the shape of their waveform (bimodal B or not bimodal C, D) which is strictly connected to the dimension of spacing between axle 3 and 4. A large spacing indicates a bimodal waveform, meaning that after axle 3 hits the bridge there is enough time for the bridge response to reach low values of strain before axle 4 hits the bridge. This type of behavior is not ideal for the bridge because one truck-crossing is perceived by the bridge as almost two cycles, this can cause fatigue problems over time. Finally, the heavy trucks with more equally spaced axles (clusters C and D) are separated by the weight of axle 6, with D having heavier rear than C, as explained before. Finally, a flow chart of the complete two-step strategy is provided to summarize the procedure in Fig.5.13.

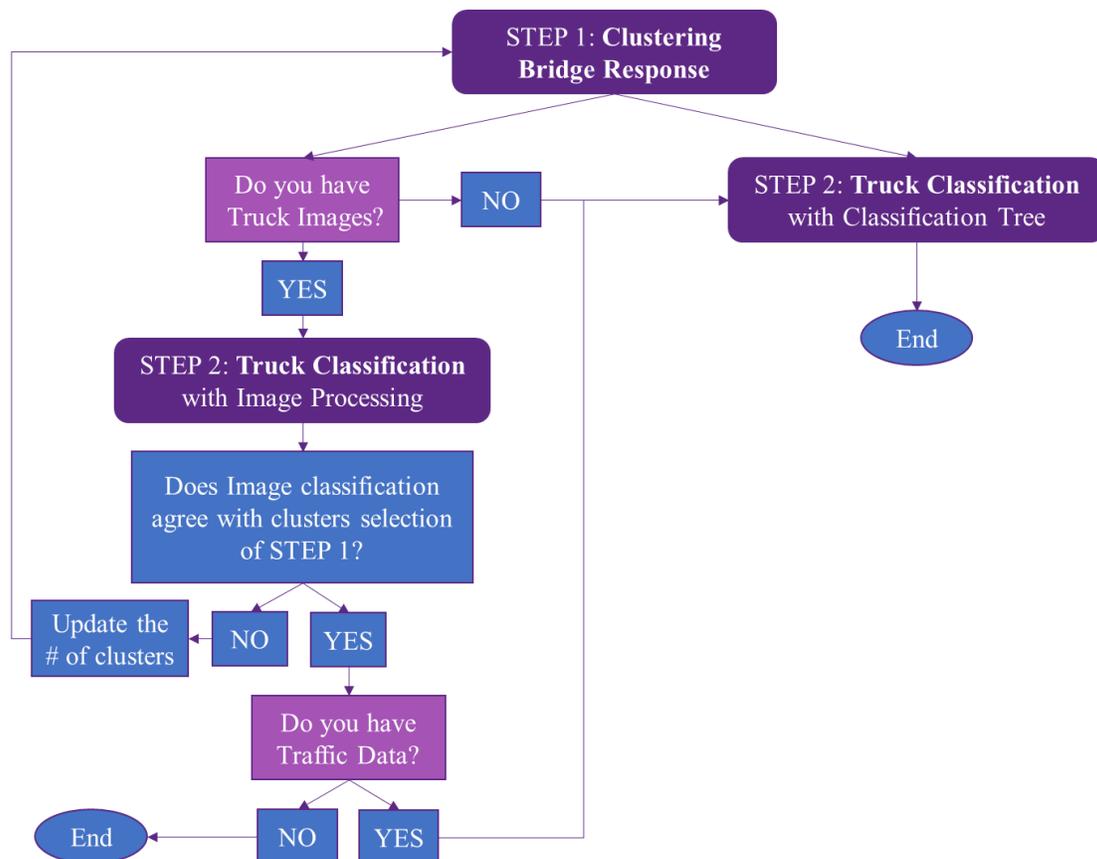


Figure 5.13: flow chart of the two-step strategy.

5.4 RESULTS

GVW and axle configurations are the key characteristics, not only to group trucks but also to keep under control their potential effects to the bridge and its components. GVW is certainly the first most important parameter. For this reason, trucks in cluster A are considered the least dangerous, with a bridge response which never exceeds $50 \mu\epsilon$. Even if they present a bimodal strain waveform due to the presence of a large spacing, typical of class 9 trucks, their limited weight assures their allocation among the safer traffic for the bridge. Unlike trucks in cluster A, trucks which fall in cluster B, C, and D should all be taken under control. Going in order of increasing GVW, trucks in cluster B are much heavier than A and often overweight (80-kip limit), in addition to their excessive weight they also present a large spacing which produces a bimodal waveform. This axle configuration makes the bridge perceive one truck crossing almost as two trucks crossings, this would speed up fatigue damages in areas which are particularly prone to it. Increasing further the GVW, there are the logging trucks with their extremely heavy weight, logging trucks are a constant presence in the Hurley bridge traffic. Both trucks in cluster C and D show a very high GVW sometimes exceeding 100 kips (limit 98-kip on 6 axles) and the same axle configuration with more equally spaced axles. For this category, the concern is due to the peak strain that they can reach, up to $100 \mu\epsilon$ and more. There are also some exceptions which should be taken under control, though very limited in number, where the bridge response can reach up to $250 \mu\epsilon$. They resulted to be all trucks with 9 axles, and, since they were found mostly in summer months and in some cases in winter. A visual summary of the truck classification is provided in Fig. 5.14.

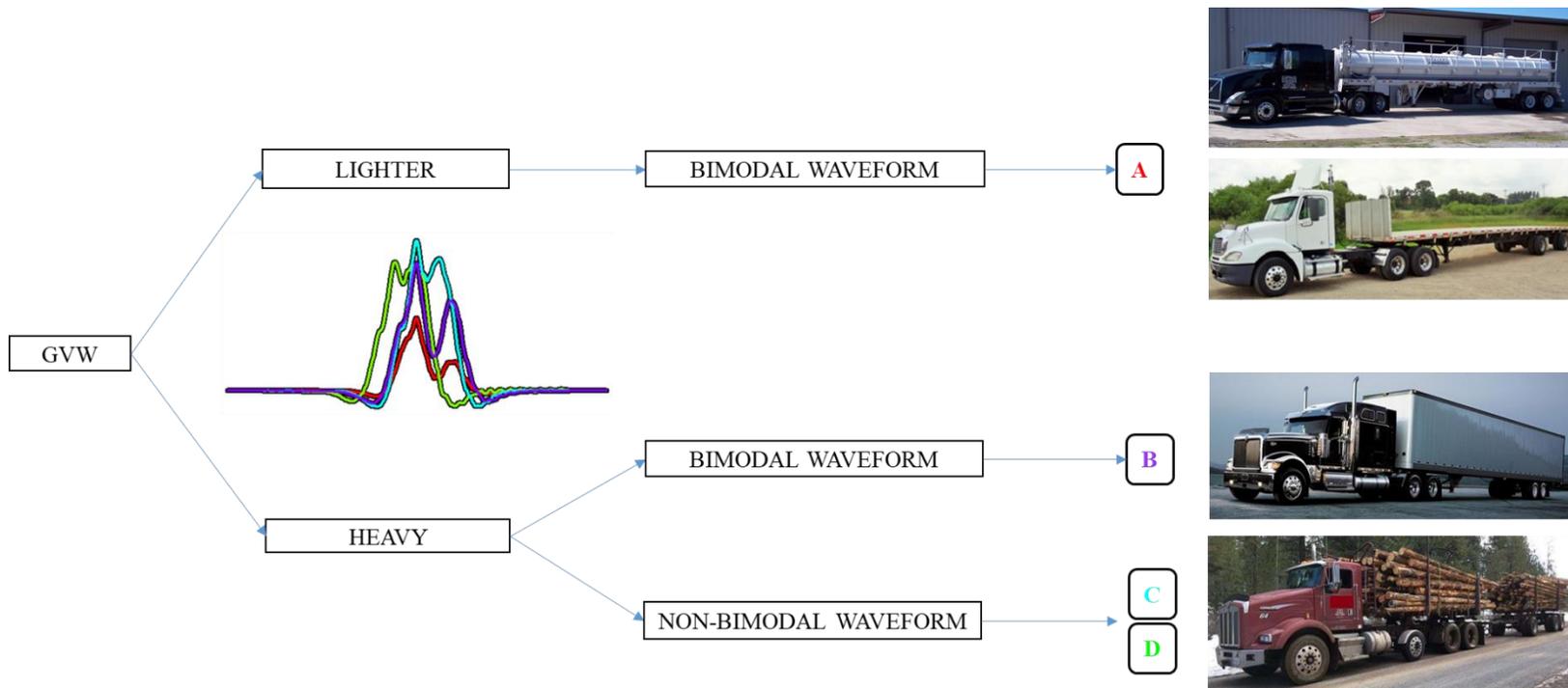


Figure 5.14: visual summary of trucks classification.

5.4.1 Monthly clustering

This section is going to show the clustering of 10 months over the 3-year period when the bridge was monitored. Only the months with a significant sample size are selected. The database is subjected to the basic normalization and outlier detection (Eq. 41). Per previous considerations, to allow direct comparison, for each month, 4 clusters are considered. Table 5.5 lists the months studied, the monthly average. Fig.5.15 shows the monthly and the yearly average, where trucks in cluster A are always between 40% and 50%, and they represent the majority of trucks. This result can be seen as encouraging since they are the least harmful to the bridge. However, Fig.5.16 and 5.17 show combined scenarios to have a more accurate overview of the situation. The former combines cluster C and D to for a better understanding of the magnitude of logging activity in the area, which are slightly less than 30% of the whole truck traffic, Table 5.6 shows that on average the probability of a logging truck to be clustered as group C is slightly higher (0.56) than group D (0.44). The latter, instead, combines all the heavy trucks B, C and D (close to limit and overweight) to have an estimated magnitude of the traffic which would require more attention in order to plan strategic maintenance of the structure, indeed these trucks are slightly more than 50%. Among the heavy trucks, the most common truck is the logging truck which concerns because its weight often exceeding 100 kips. Slightly less numerous than logging trucks are the trucks belonging to cluster B, heavy truck with large spacing with the potential to cause fatigue issues, these kinds of trucks show a peak in 2011 with 30% (Fig.5.15), in correspondence of a downward trend of trucks A. This result could be a consequence of the Wisconsin statutory change (Owusu-Ababio and Schmitt, 2014) which allowed heavy trucks to cross the area even during the spring thaw.

Table 5.5: percentages of trucks in each cluster.

Month	Size	A	B	C	D	B&C&D	C&D
Aug-10	1310	48.1%	22.7%	15.8%	13.4%	51.9%	29.2%
Sep-10	1342	49.8%	25.0%	14.5%	10.7%	50.2%	25.2%
Oct-10	1791	48.2%	23.3%	15.2%	13.3%	51.8%	28.5%
Nov-10	1585	49.6%	22.0%	16.9%	11.5%	50.4%	28.4%
Dec-10	1688	48.1%	23.2%	16.8%	12.0%	52.0%	28.8%
Mar-11	2029	40.8%	28.4%	18.9%	11.9%	59.2%	30.8%
Apr-11	1705	42.1%	31.1%	15.7%	11.1%	57.9%	26.8%
Mar-12	1955	45.3%	25.1%	17.3%	12.3%	54.7%	29.6%
Apr-12	1884	47.0%	28.5%	13.3%	11.2%	53.0%	24.5%
Jun-12	2293	48.7%	28.1%	10.7%	12.5%	51.3%	23.2%
	Average	46.8%	25.7%	15.5%	12%	53.2%	27.5%

Table 5.6: logging trucks probabilities.

Month	Size	Logging trucks	Pr(C)	Pr(D)
Aug-10	1310	29.2%	0.54	0.46
Sep-10	1342	25.2%	0.58	0.42
Oct-10	1791	28.5%	0.53	0.47
Nov-10	1585	28.4%	0.60	0.40
Dec-10	1688	28.8%	0.58	0.42
Mar-11	2029	30.8%	0.61	0.39
Apr-11	1705	26.8%	0.59	0.41
Mar-12	1955	29.6%	0.58	0.42
Apr-12	1884	24.5%	0.54	0.46
Jun-12	2293	23.2%	0.46	0.54
	Average	27.5%	0.57	0.44

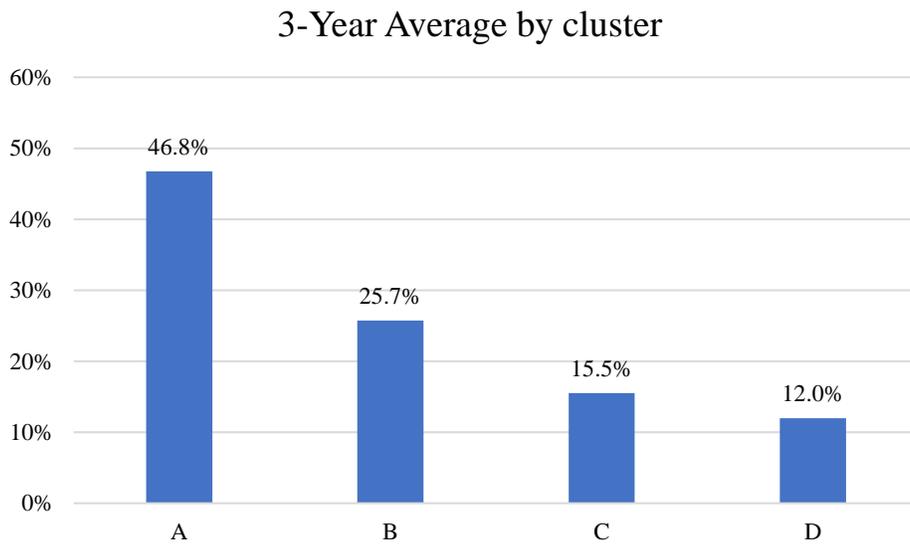
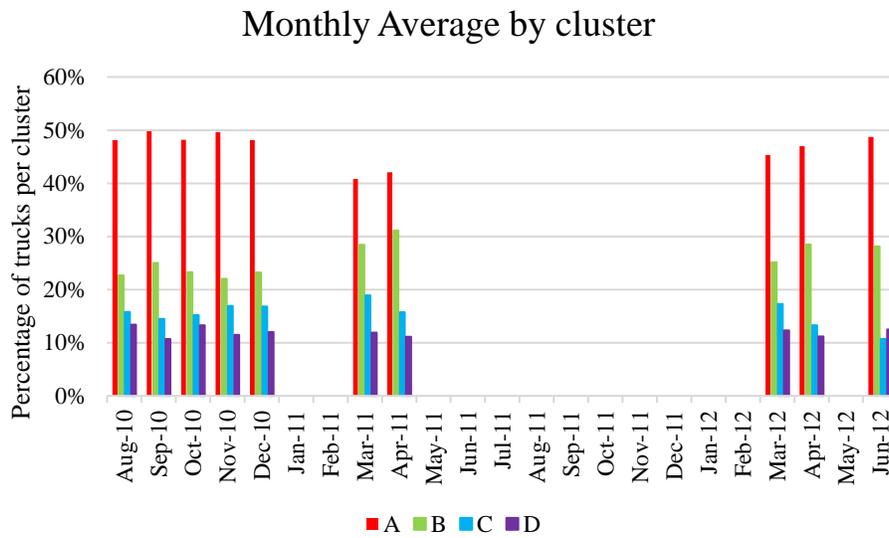


Figure 5.15: monthly and 3-year average percentages of trucks by cluster.

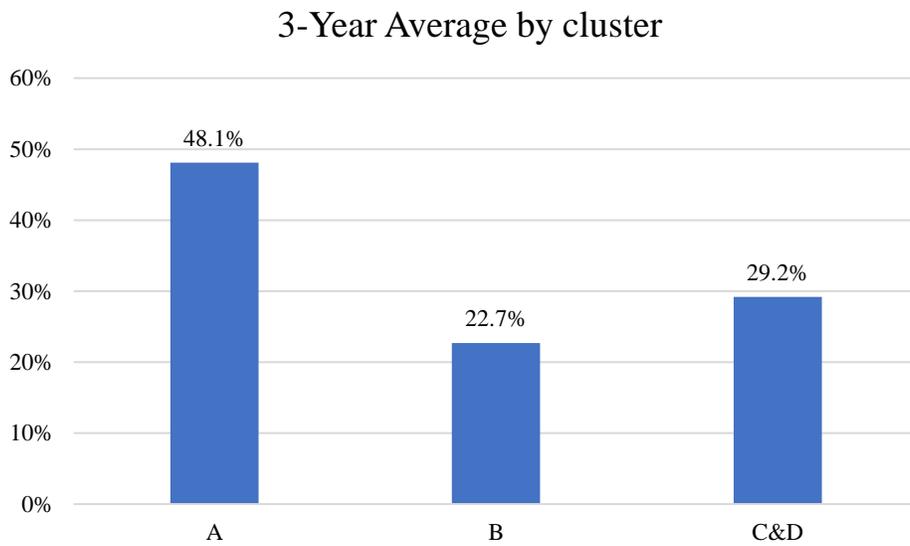
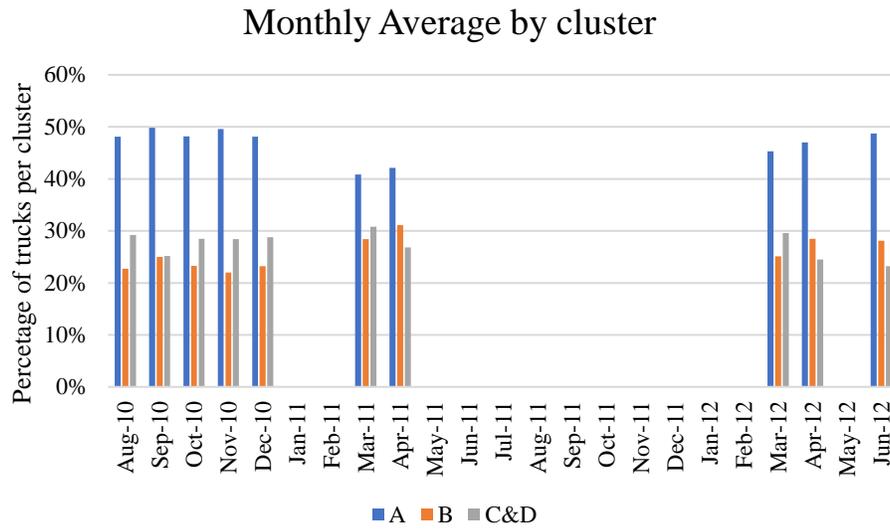


Figure 5.16: monthly and 3-year average percentages of trucks by cluster, with C&D representing logging trucks.

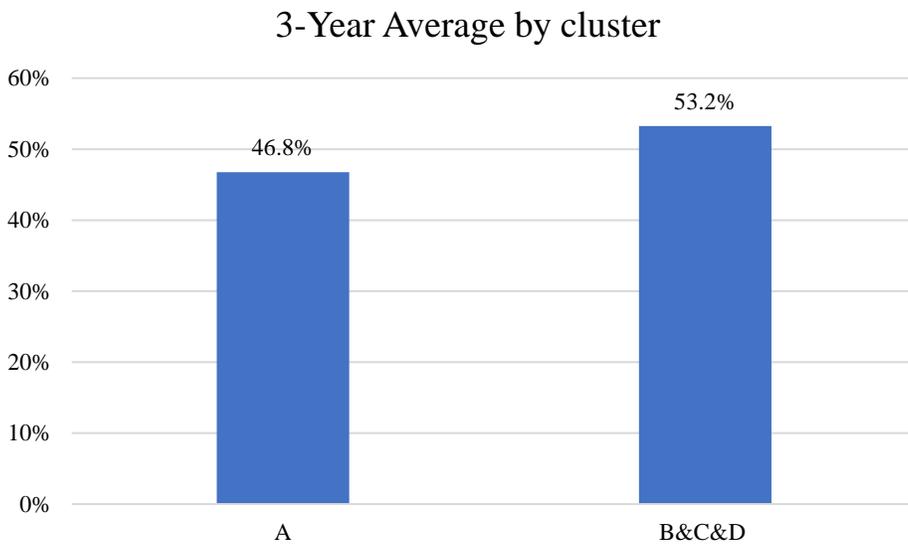
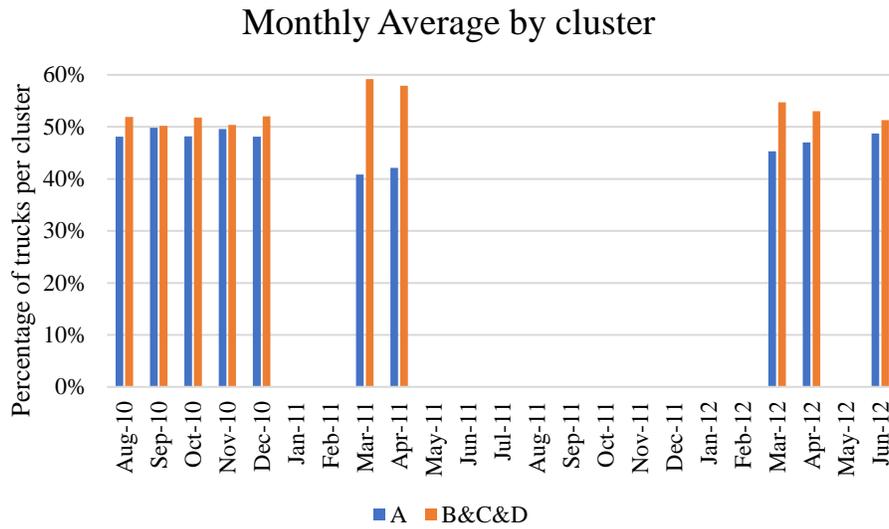


Figure 5.17: monthly and 3-year average percentages of trucks by cluster where B&C&D represents heavy trucks.

Figures 5.18 to 5.27 show, for each month considered in the study, the four clusters which are the result of the average of all waveforms included assigned to that cluster and the individual waveforms color-coded by clusters. The y-axis noted as “V” in the figures represent strain [$\mu\epsilon$]. Table 5.7 shows how the maximum value of all clusters is stable over time, with very small variations. The highest value of strain for the clusters is about 100 $\mu\epsilon$ for all months (cluster C), however, some months show individual waveforms which exceed 200 $\mu\epsilon$ (August, November and December 2010). Even if the clusters are the result of the average of all waveforms, their magnitude is not highly influenced by the presence of extreme events because of their limited number. Therefore, if extreme events become more frequent the cluster will show it and this can be used as a warning. The shape of the cluster can be determined by the number of each truck’s typology. For instance, cluster A and B maintain their shape consistent over the period while C and D tend to slightly vary their shape, in particular, D, depending on the number of trucks with a large spacing, can tend to a more bimodal shape.

Table 5.7: maximum strain values of each cluster.

	Maximum Averaged Strain Values [$\mu\epsilon$]			
Month	A	B	C	D
Aug-10	47.34	78.51	99.11	90.35
Sep-10	45.36	76.92	99.23	83.67
Oct-10	45.36	77.18	99.72	95.68
Nov-10	44.81	75.11	94.43	88.61
Dec-10	43.01	74.14	92.67	83.76
Mar-11	42.34	74.45	92.22	85.69
Apr-11	44.36	78.28	94.27	79.84
Mar-12	43.39	76.12	95.57	86.78
Apr-12	45.06	75.35	97.38	86.32
Jun-12	45.69	74.14	96.11	87.62

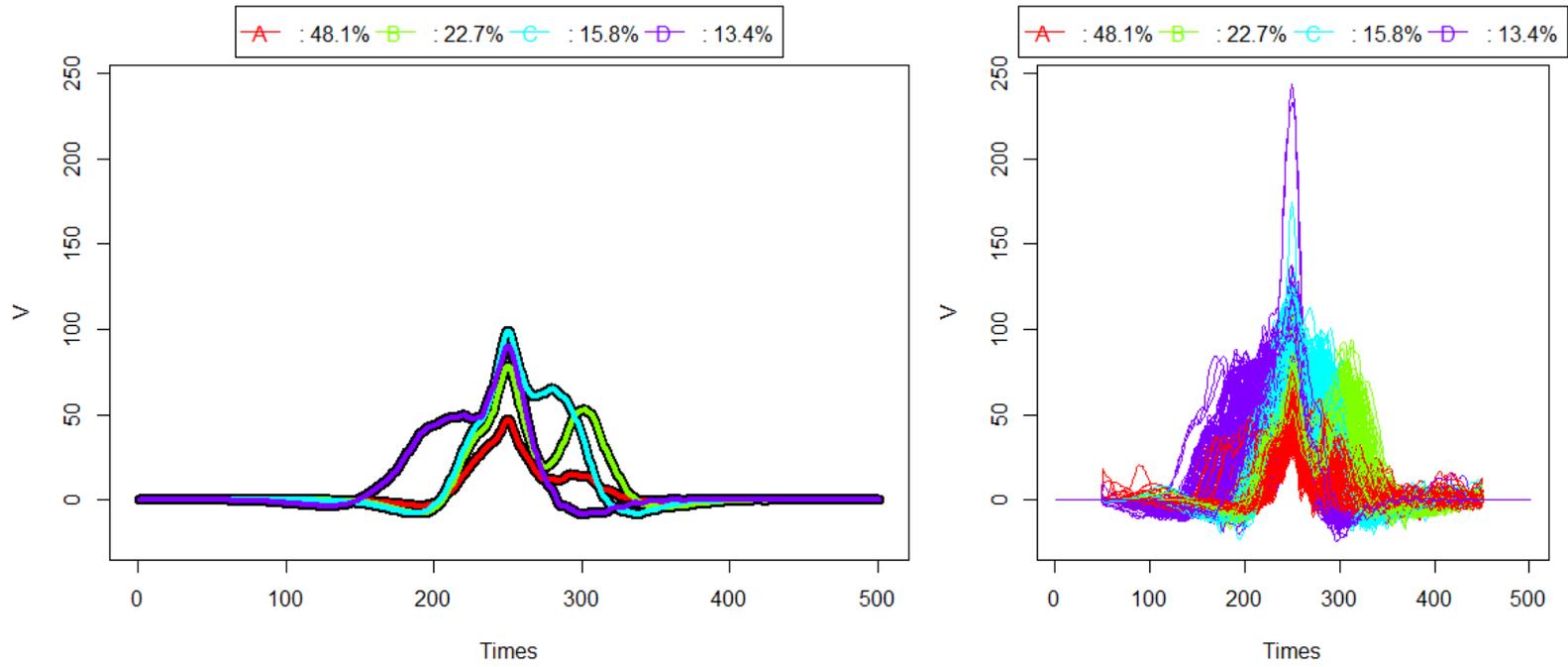


Figure 5.18: August 2010, clusters (left) waveforms sorted by cluster (right).

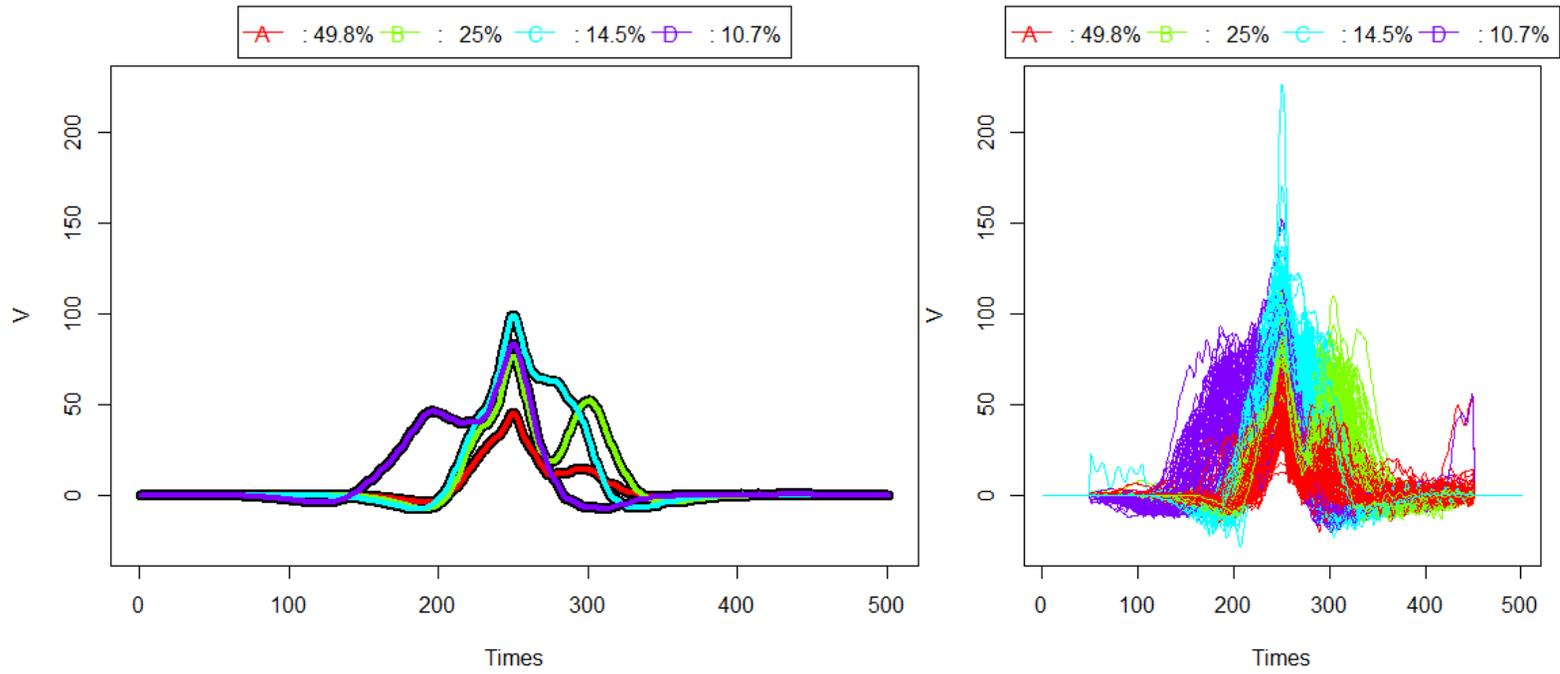


Figure 5.19: September 2010, clusters (left) waveforms sorted by cluster (right).

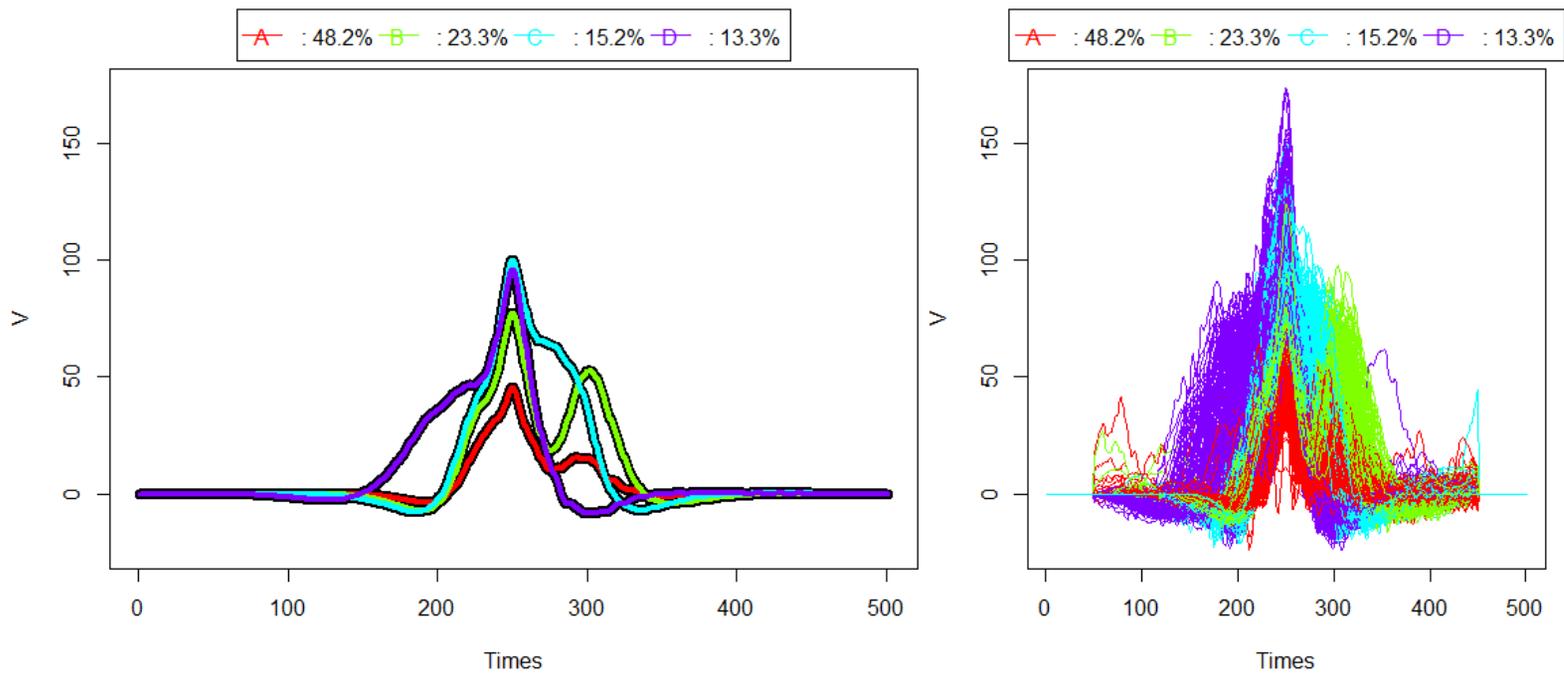


Figure 5.20: October 2010, clusters (left) waveforms sorted by cluster (right).

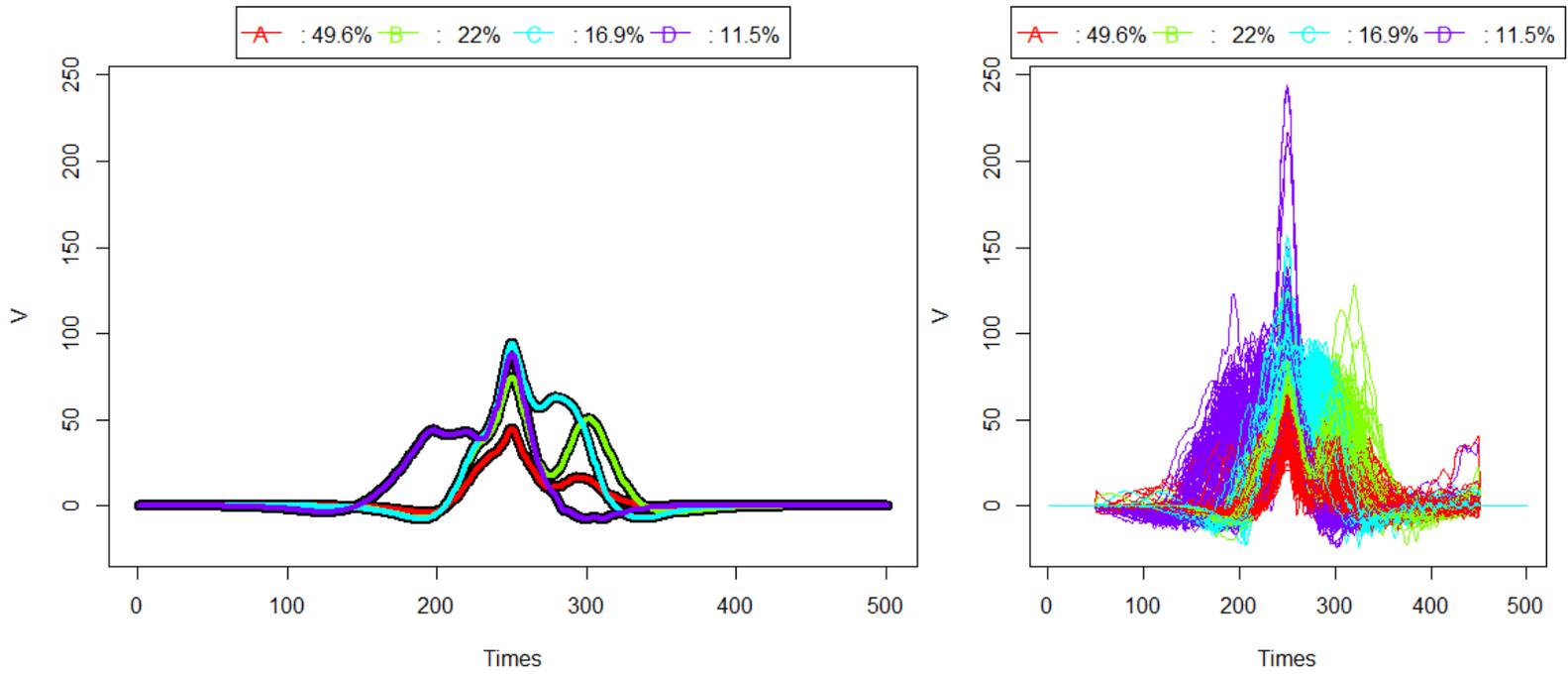


Figure 5.21: November 2010, clusters (left) waveforms sorted by cluster (right).

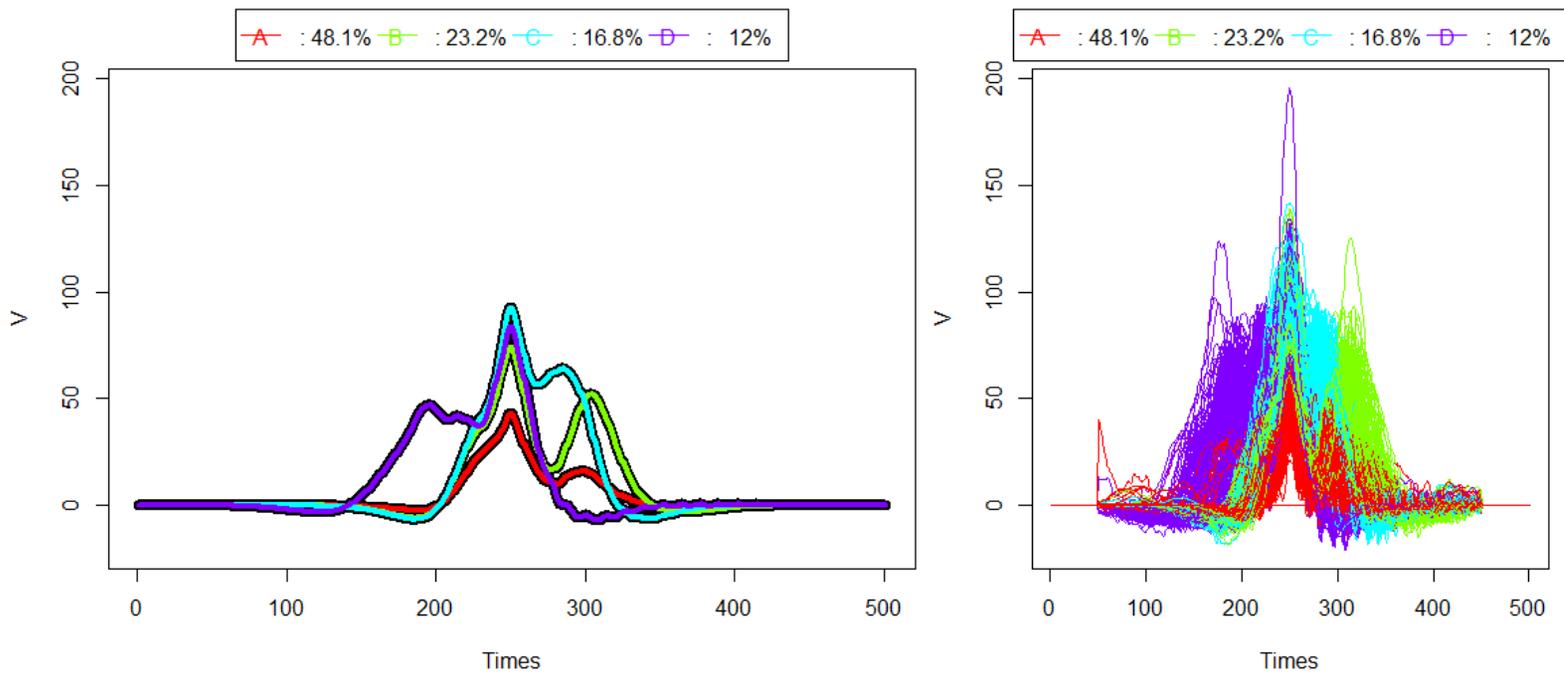


Figure 5.22: December 2010, clusters (left) waveforms sorted by cluster (right).

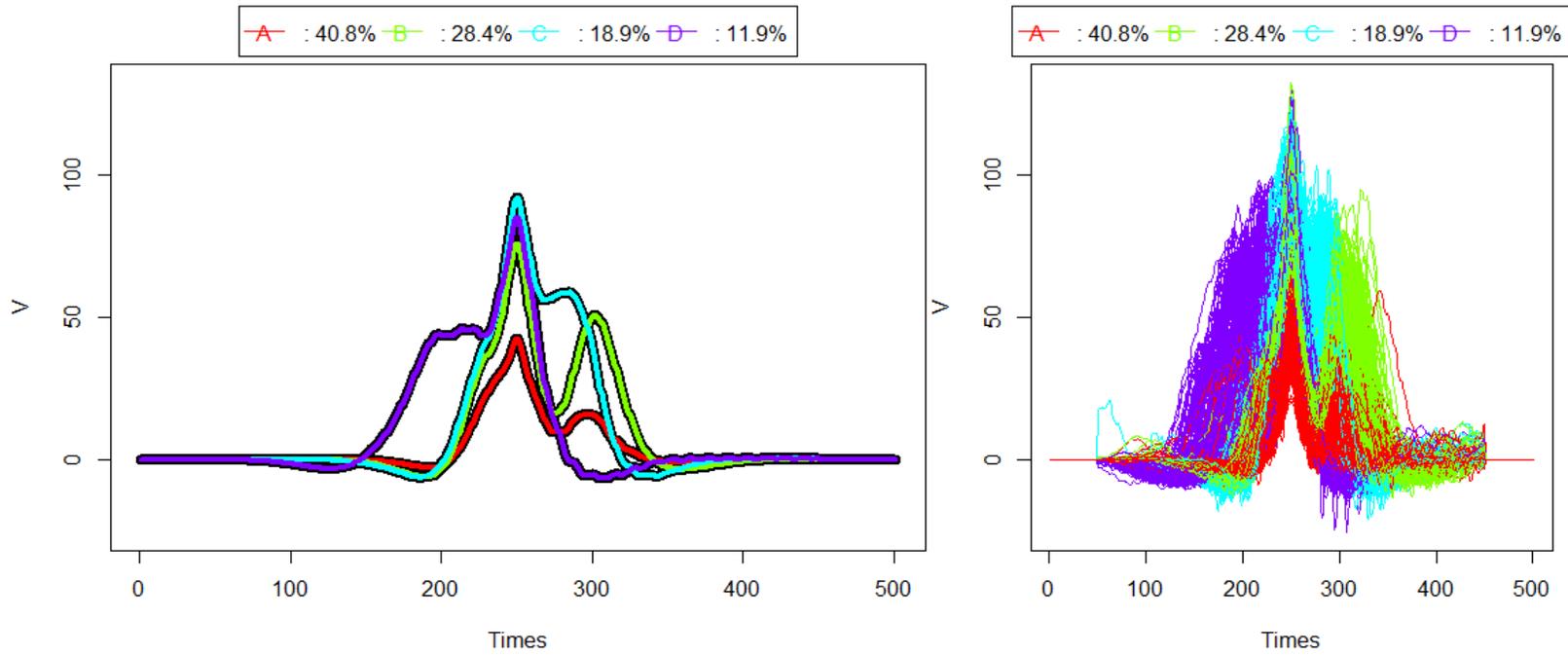


Figure 5.23: March 2011, clusters (left) waveforms sorted by cluster (right).

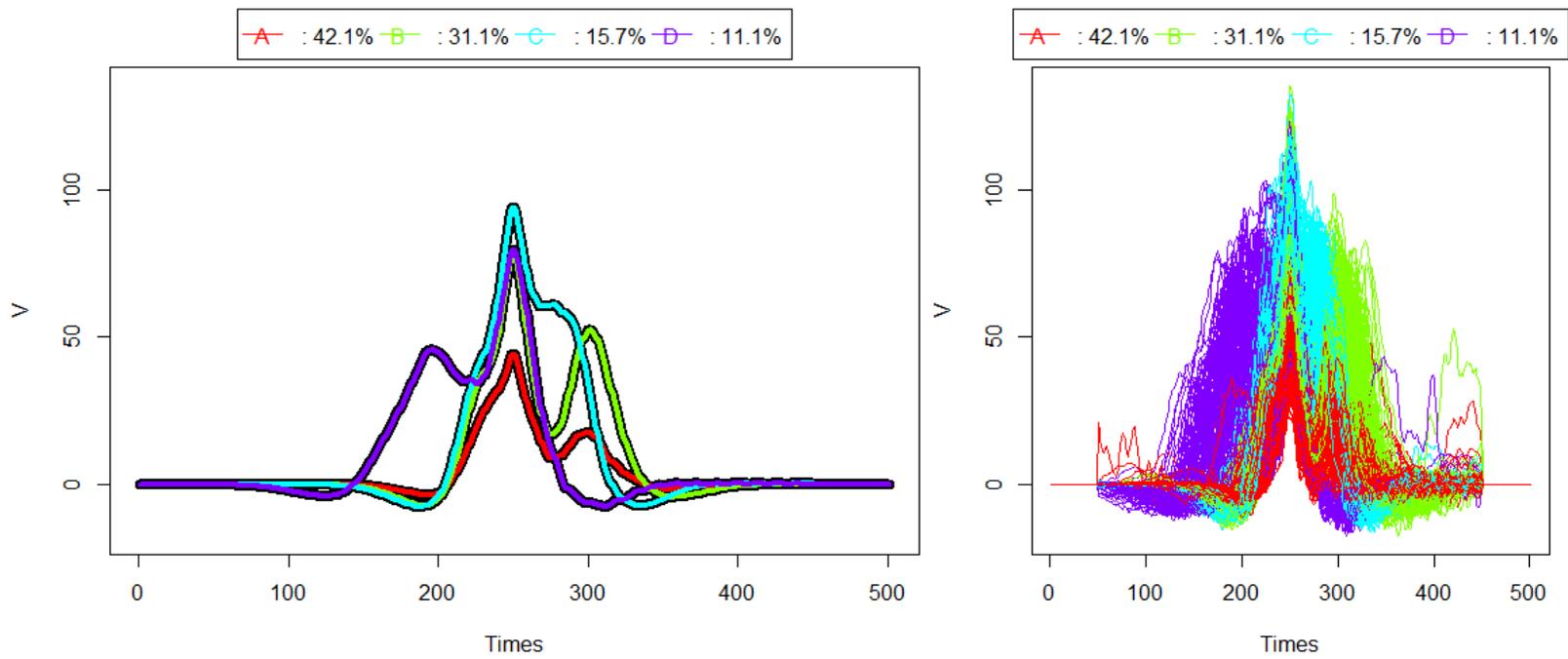


Figure 5.24: April 2011, clusters (left) waveforms sorted by cluster (right).

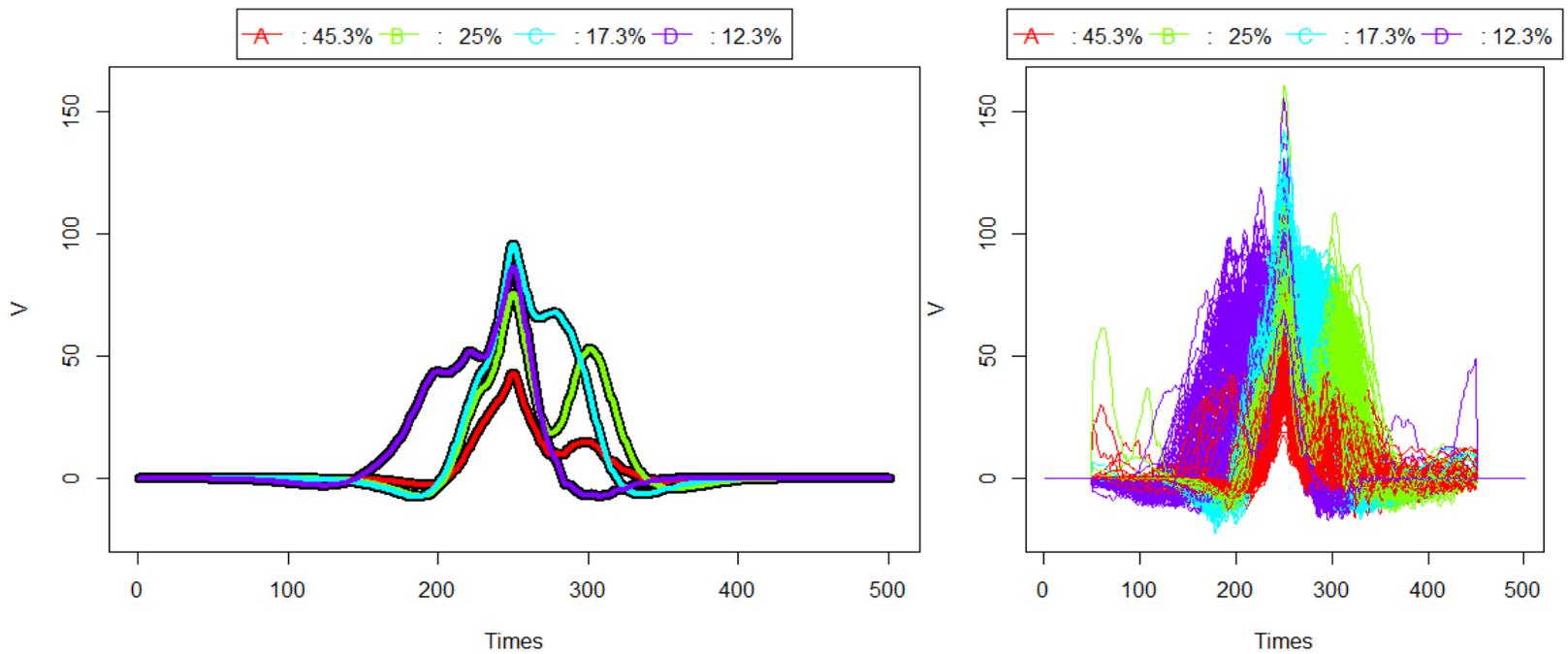


Figure 5.25: March 2012, clusters (left) waveforms sorted by cluster (right).

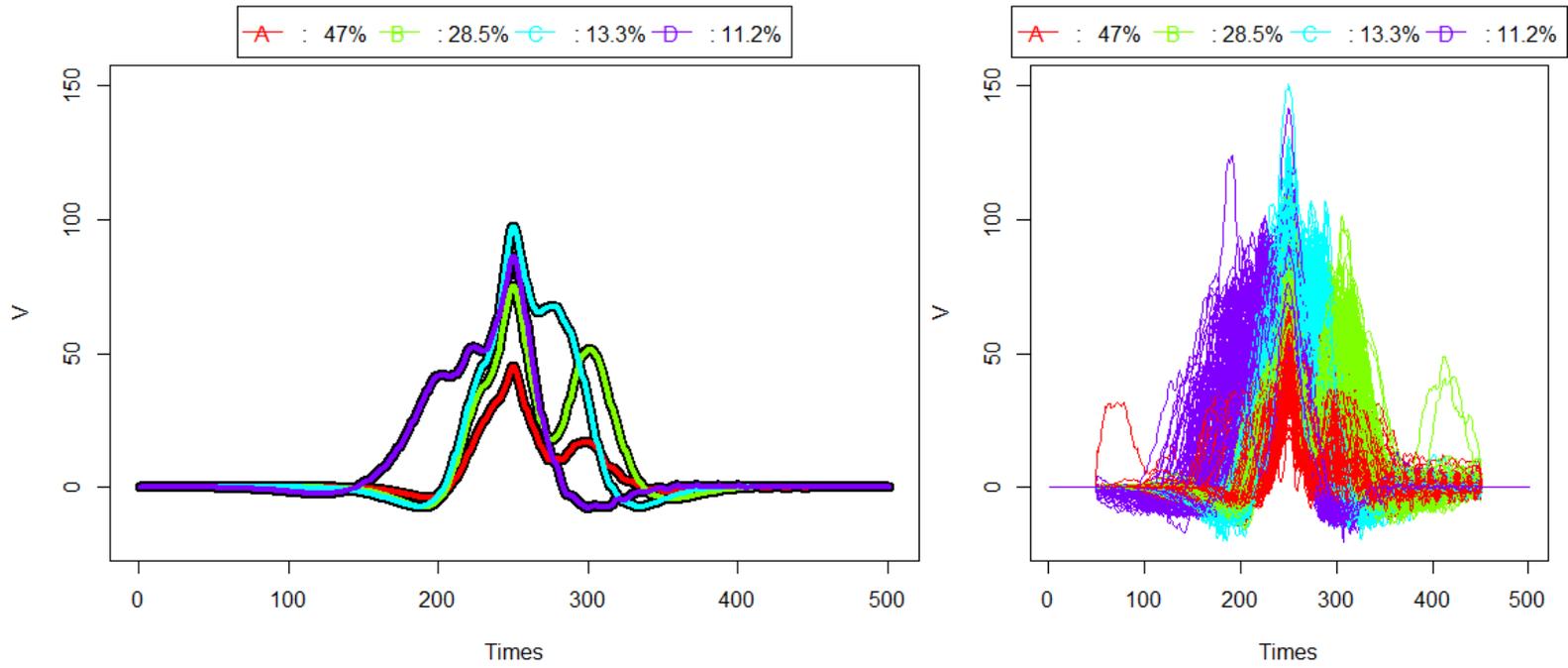


Figure 5.26: April 2012, clusters (left) waveforms sorted by cluster (right).

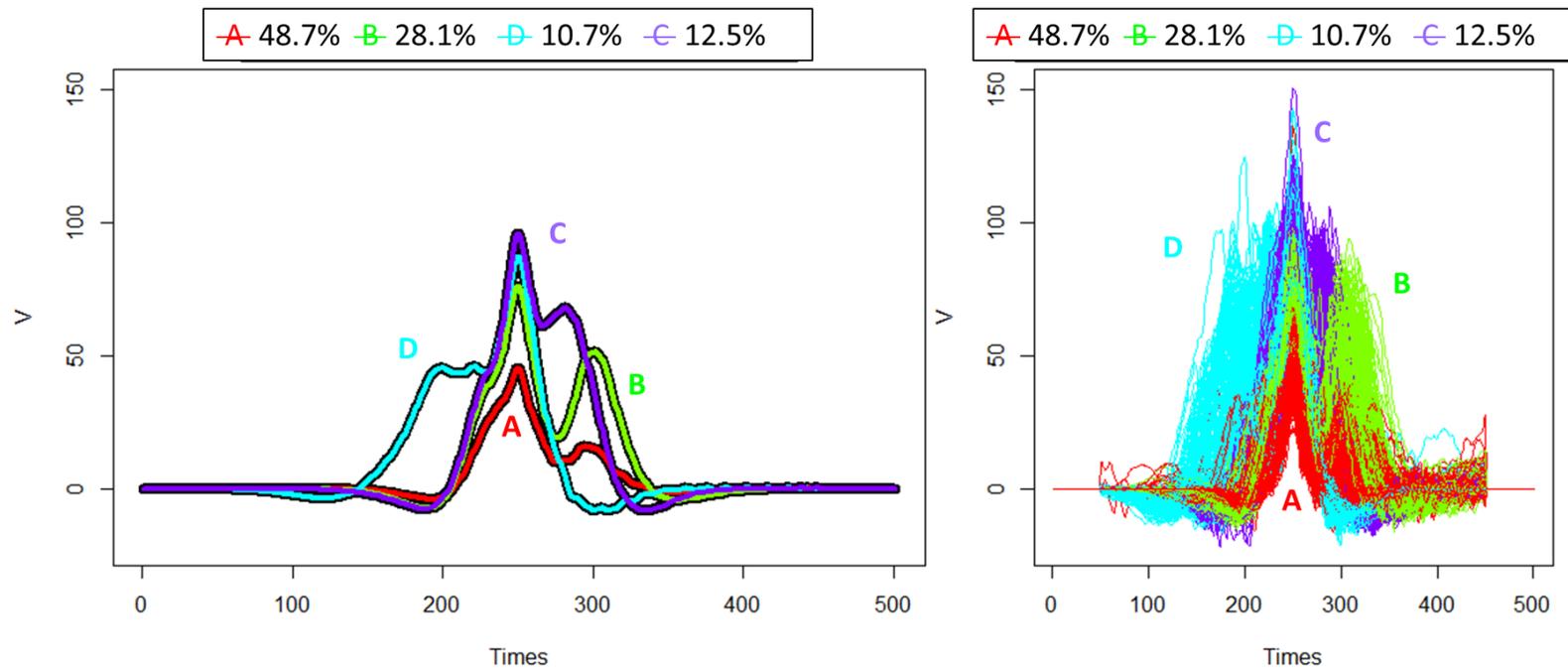


Figure 5.27: June 2012, clusters (left) waveforms sorted by cluster (right).

5.4.1.1 Sensitivity Analysis

A sensitivity analysis is performed to show how different normalization methods and outlier cleansing techniques can influence the clustering results.

Comparison 1: normalization 1

Fig.5.28 to 5.37 aim to show the difference between the two methods of normalization and outlier detection explained in Eq.41 and Eq. 43. Both methods have the following in common, all the peaks of the waveforms are centered (basic normalization), and both discard those trucks whose data present WIM incongruences.

The difference between the methods are:

- Basic Normalization and Outlier Detection (Eq. 41): tries to discard most of the multiple vehicle presences. The advantage is to have a cleaner dataset where fewer outliers can affect the clusters; the disadvantage is a reduced sample size to study.
- Time-Window Normalization and Outlier Detection (Eq.43): adopts a different normalization which does not discard those trucks whose waveform signals the presence of another vehicle but only considers the portion of waveform generated by the crossing of the truck. The advantage is that fewer trucks are discarded; the disadvantage is that the calculation of the time window (t_{tc} , Eq.42) depends on the speed measured by the WIM, which is hardly kept constant.

The choice of outlier detection and normalization can influence the number of clusters selected, for instance, April 2012 (Fig. 5.36) shows 5 clusters or 4 depending on the method selected.

Overall, the main difference is a slight redistribution of percentage.

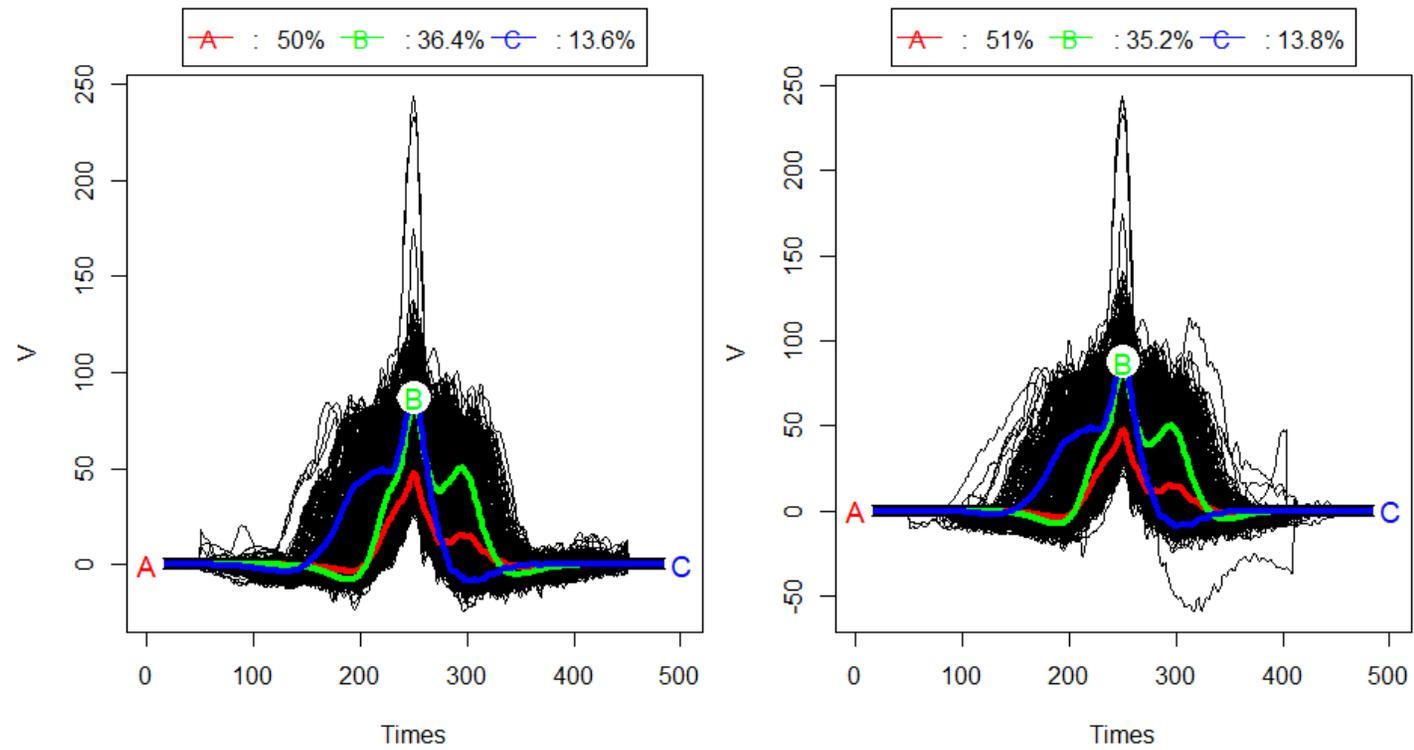


Figure 5.28: August 2010, 3 clusters, (left) basic normalization, (right) time-window normalization.

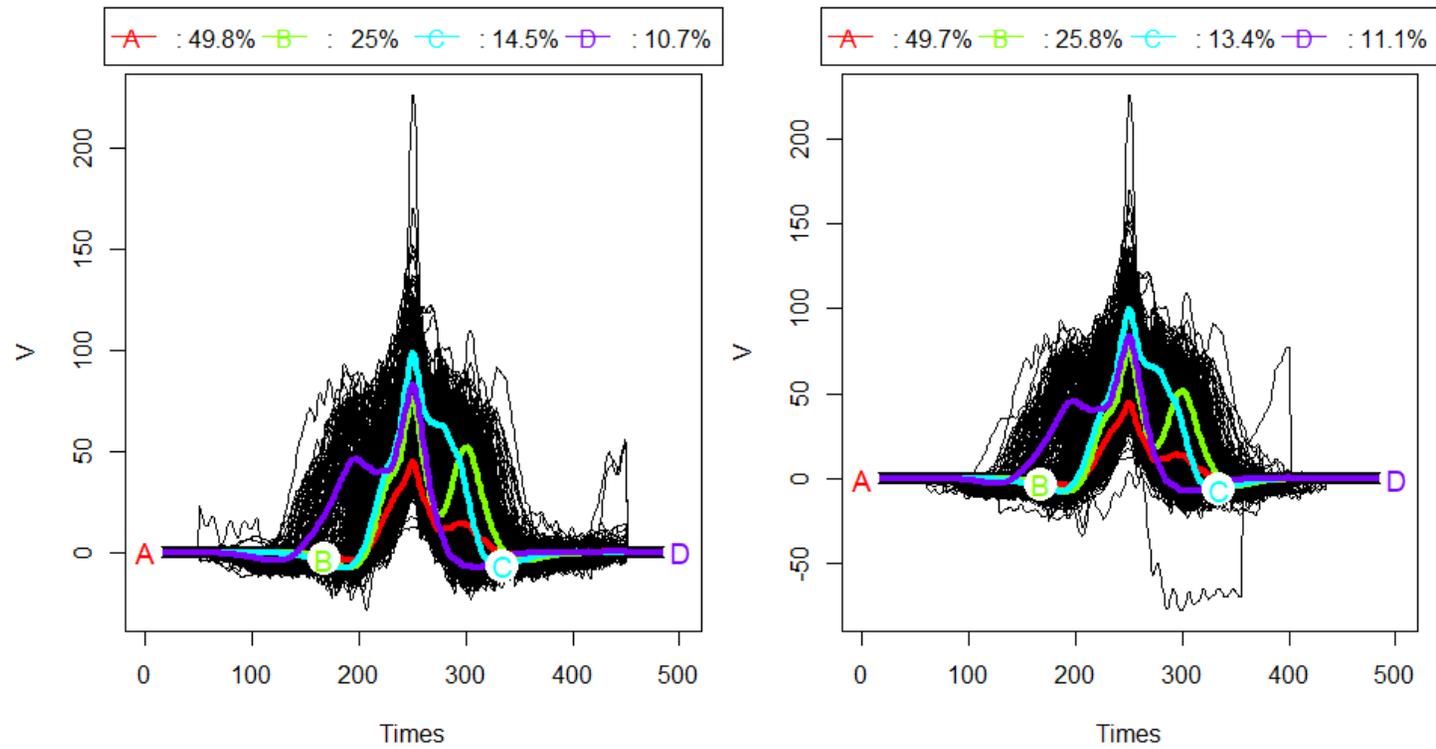


Figure 5.29: September 2010, 4 clusters, (left) basic normalization, (right) time-window normalization.

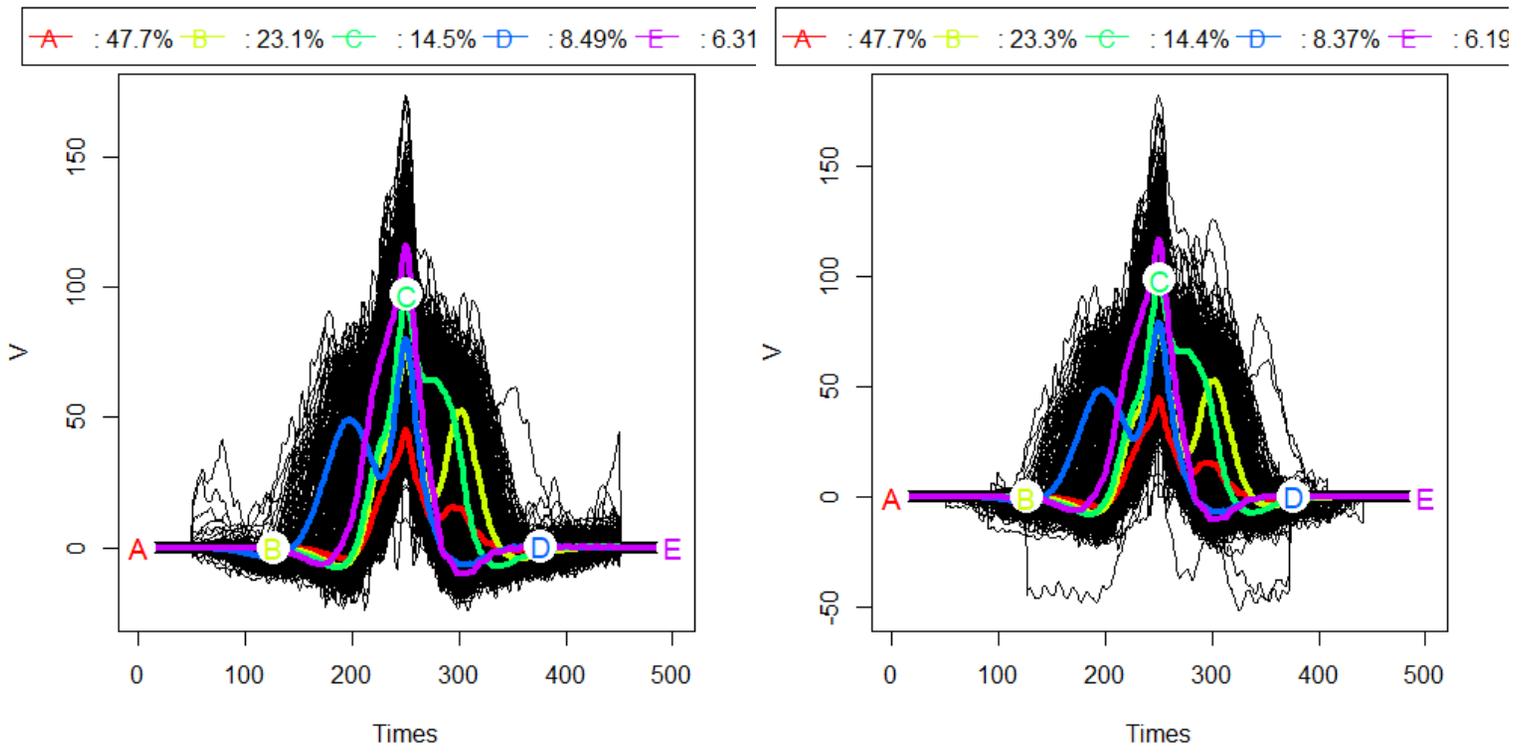


Figure 5.30: October 2010, 5 clusters, (left) basic normalization, (right) time-window normalization.

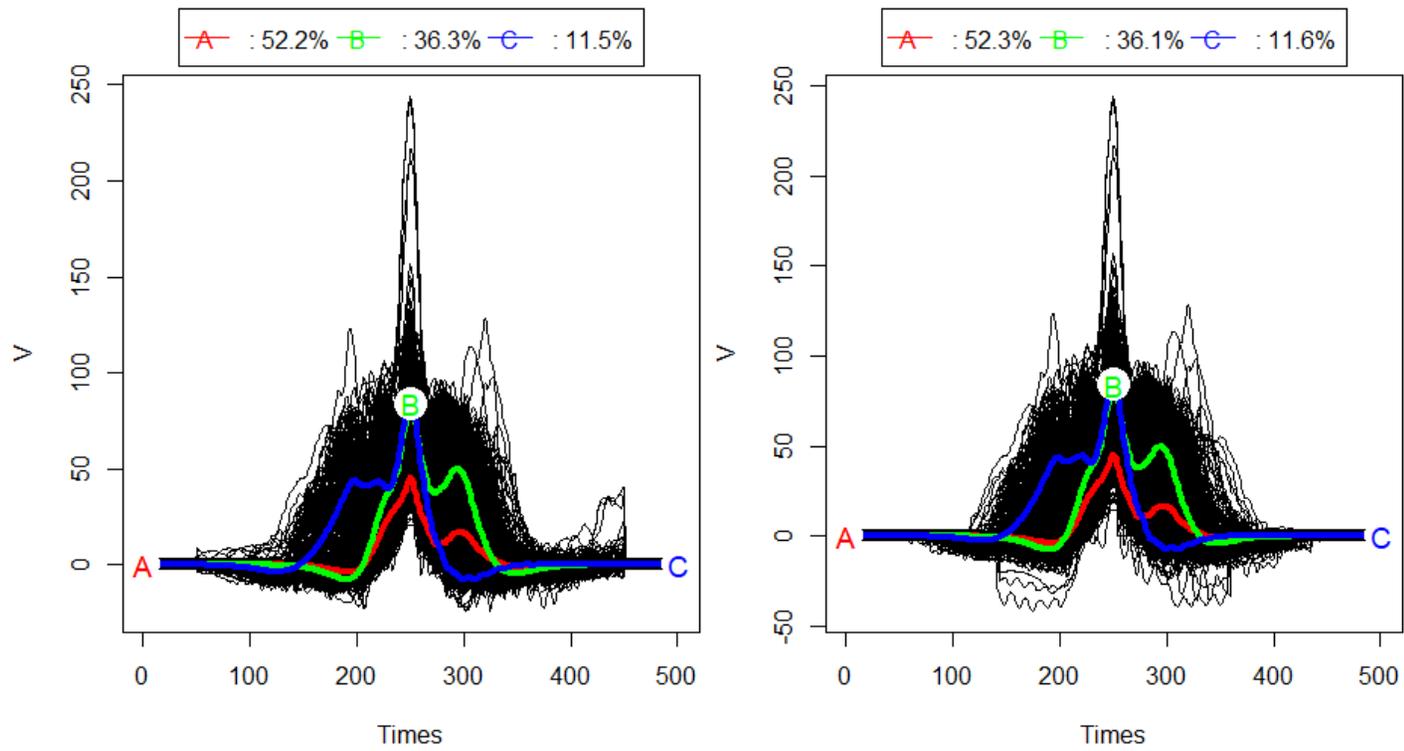


Figure 5.31: November 2010, 3 clusters, (left) basic normalization, (right) time-window normalization.

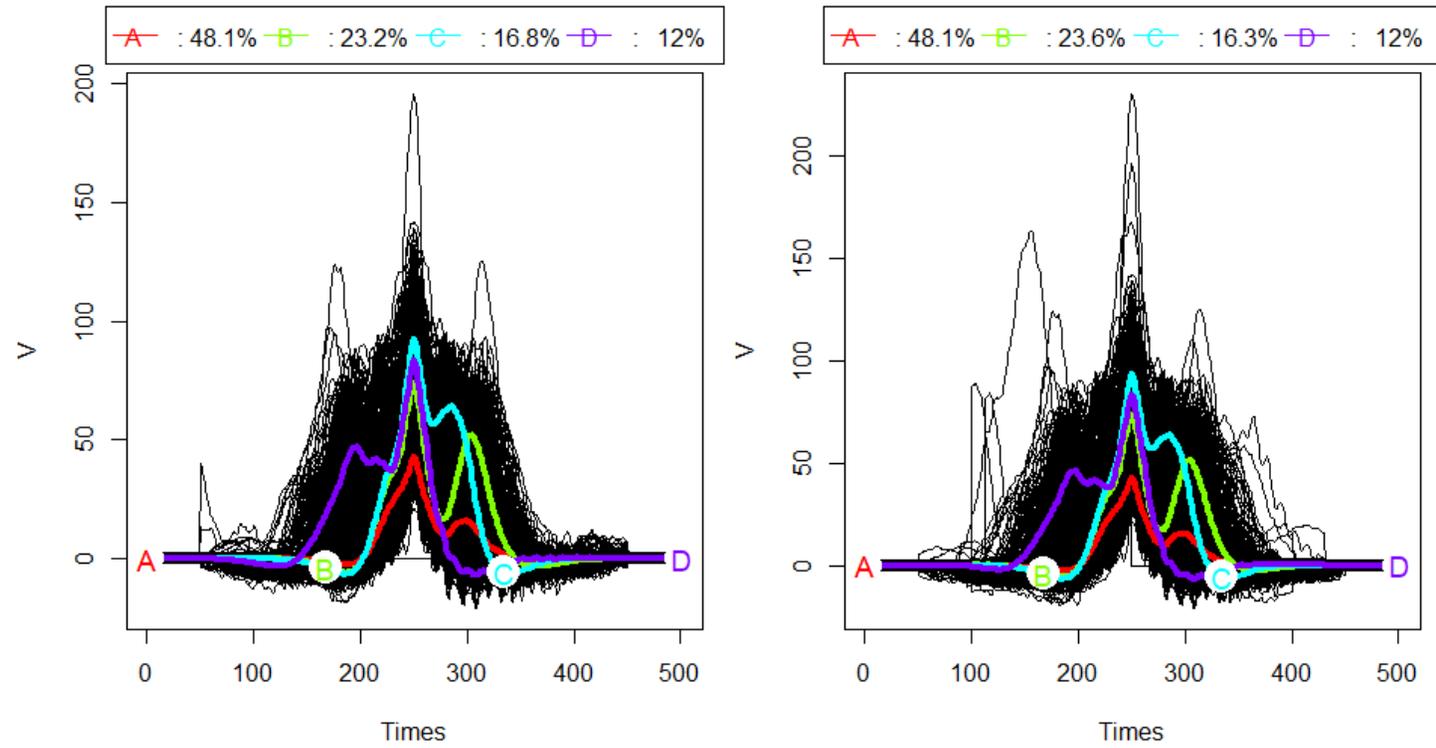


Figure 5.32: December 2010, 4 clusters, (left) 4-rule outliers detection, (right) ttc-normalization.

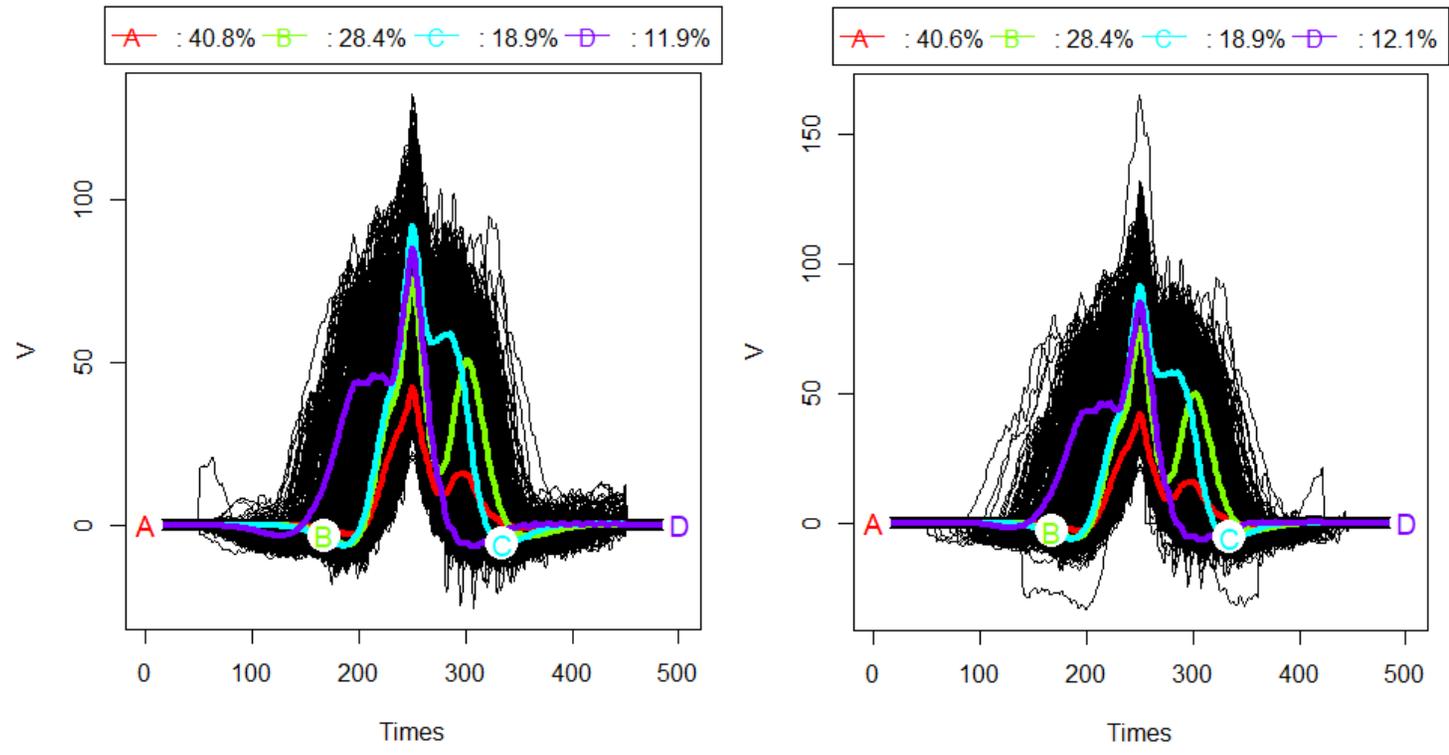


Figure 5.33: March 2011, 4 clusters, (left) basic normalization, (right) time-window normalization.

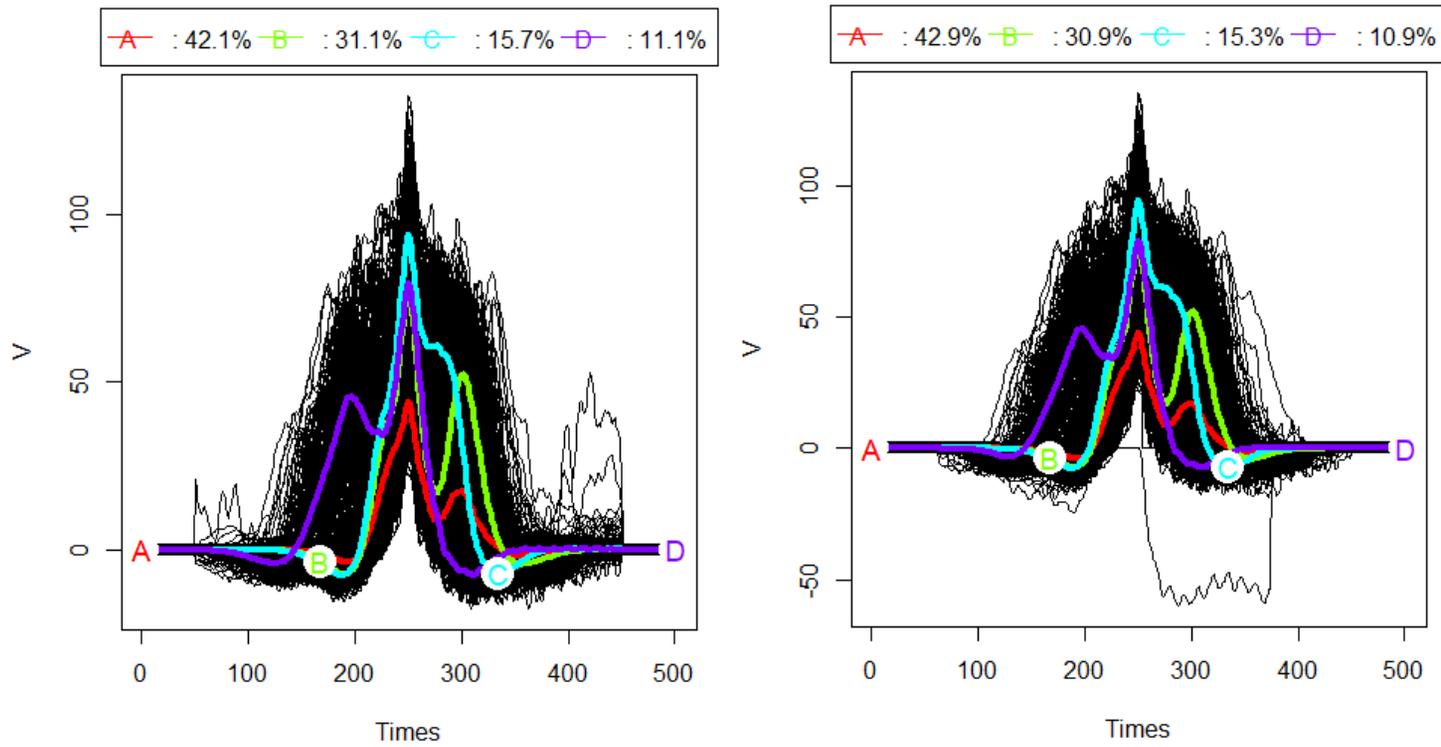


Figure 5.34: April 2011, 4 clusters, (left) basic normalization, (right) time-window normalization.

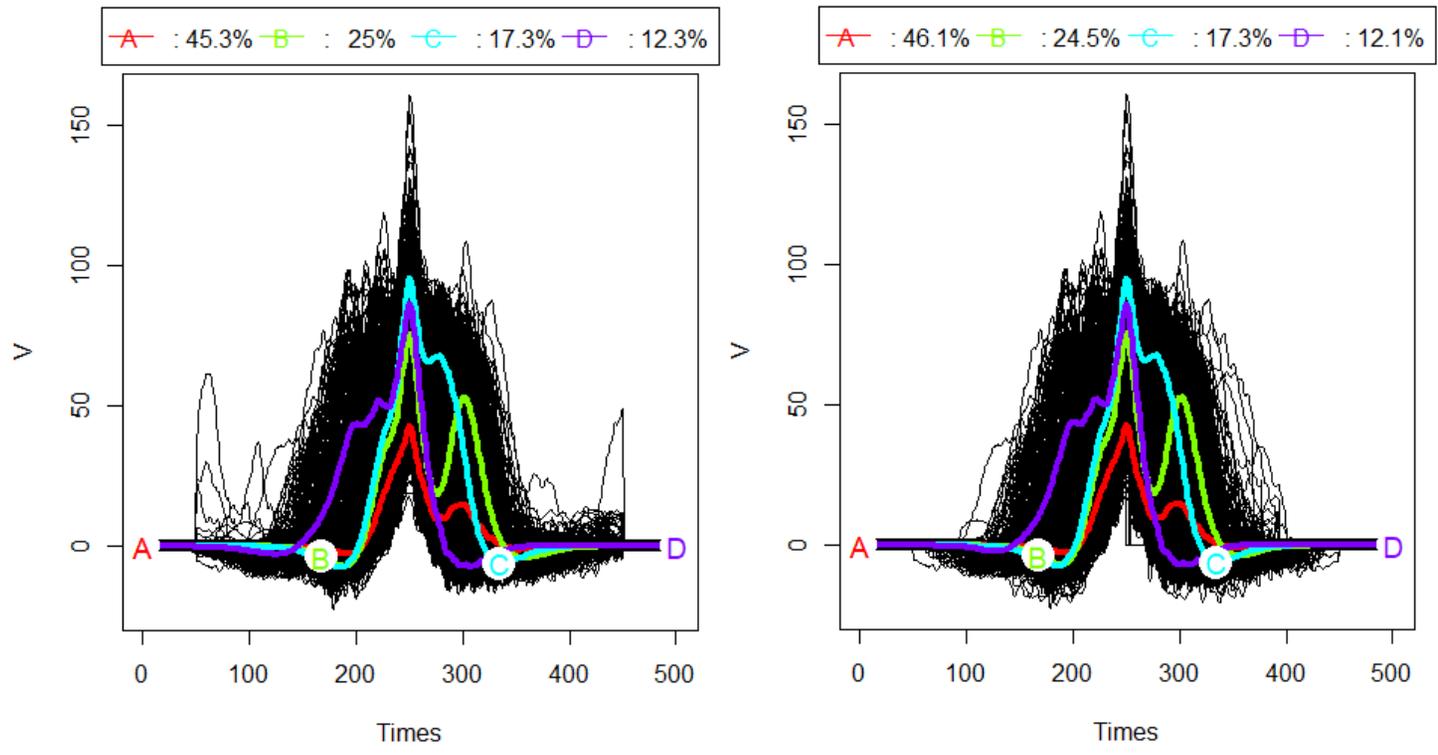


Figure 5.35: March 2012, 4 clusters, (left) basic normalization, (right) time-window normalization.

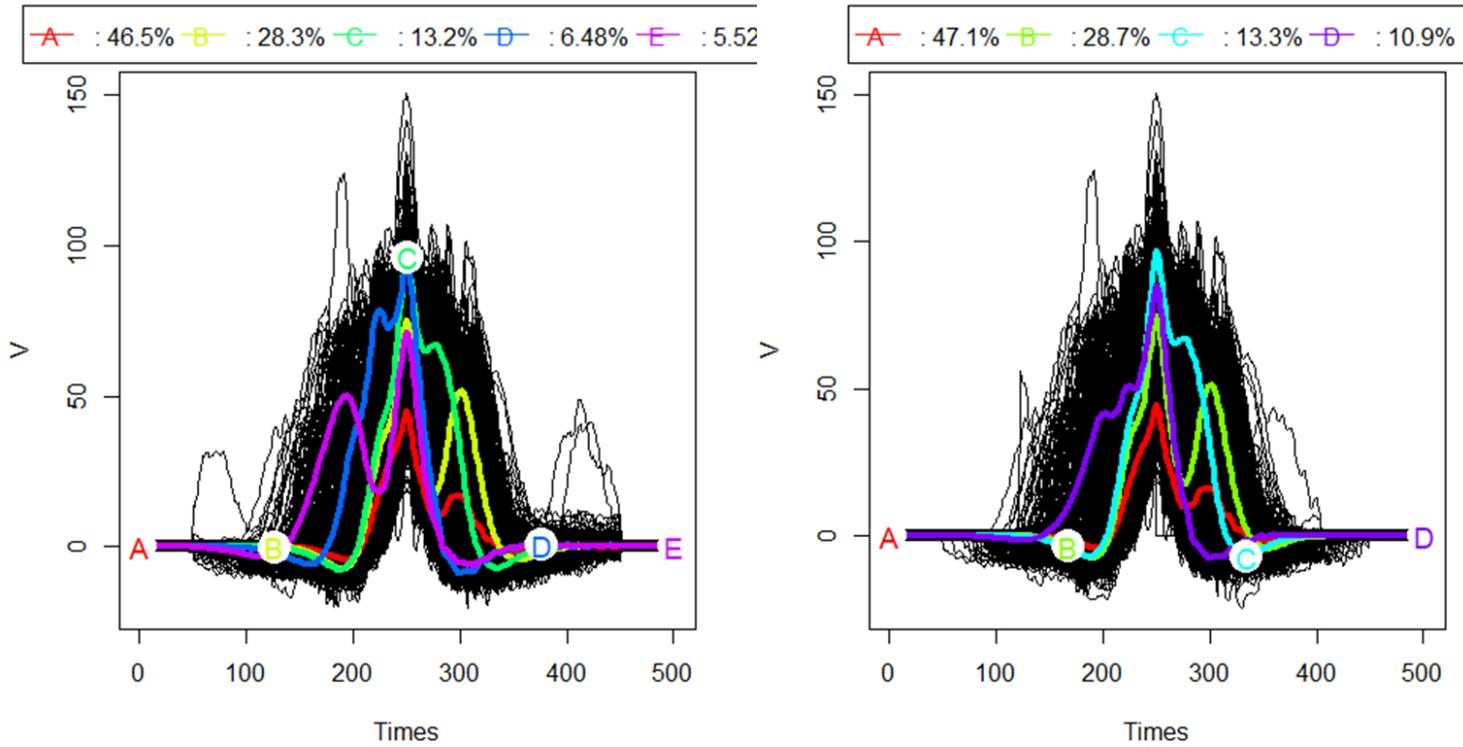


Figure 5.36: April 2012, 4 clusters, (left) basic normalization, (right) time-window normalization.

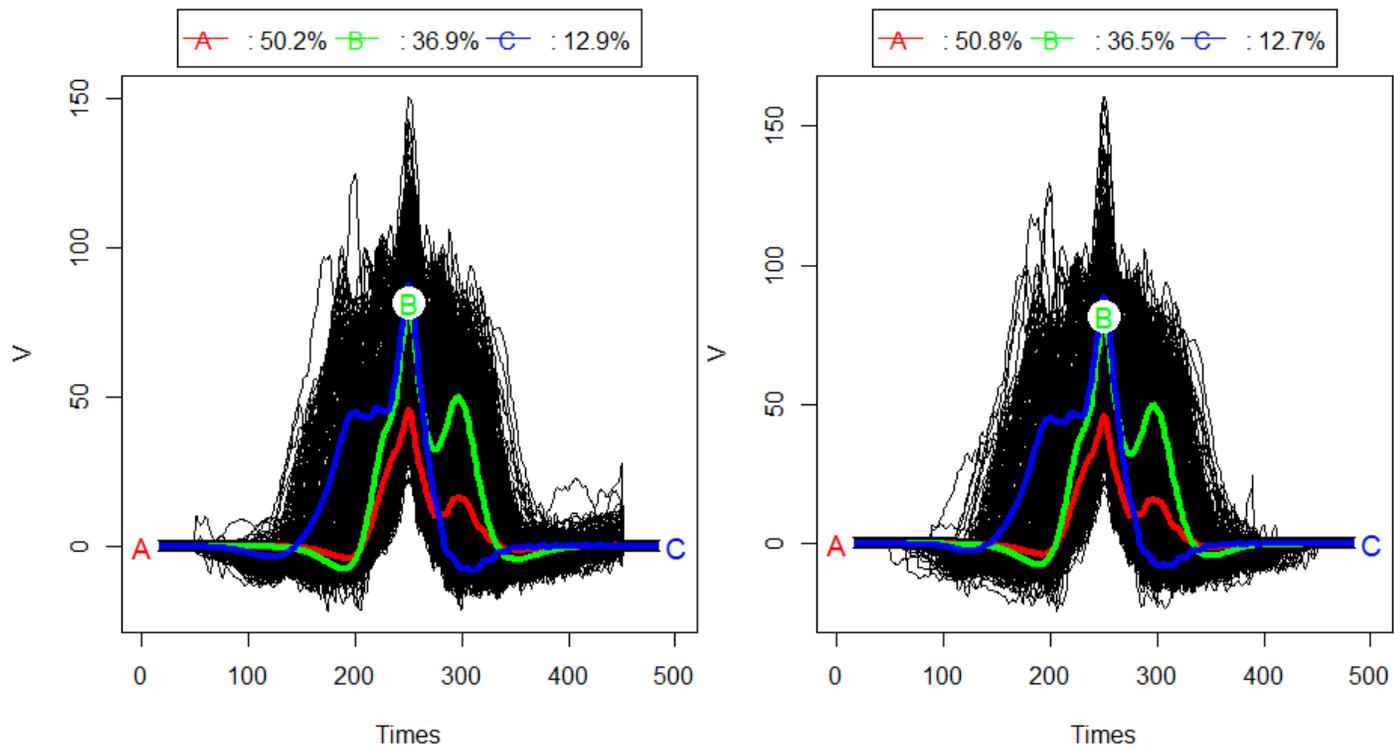


Figure 5.37: June 2012, 4 clusters, (left) basic normalization, (right) time-window normalization.

Comparison 2: normalization 2

The second comparison aims to show the importance of performing the normalization. For this analysis, March 2012, lane 2 is used including potential outliers. Dataset 1 is not normalized, dataset 2 is normalized (basic). Fig.5.38 shows how all waveforms are more spread due to the different speed of the trucks, waveforms towards the left belong to faster trucks than the waveforms towards the right. The clusters resulting from the analysis of dataset 1 are not reliable because highly biased by the speed which determines their position in the plot.

Comparison 3: normalization 3

In the third comparison, time-window normalization is used on March 2012, lane 2. The goal of this comparison is to show the alignment point used for the normalization influences the clusters. Dataset 3 is normalized by aligning the peaks, while dataset 4 aligns the time-frame when the truck enters the bridge, identified by the time at which the ttc starts (Fig.5.39). The peaks eventually result aligned in both datasets, however, the shape of the clusters is slightly different. The peak alignment is considered more reliable than the entering-moment alignment because not dependent on the speed which can change.

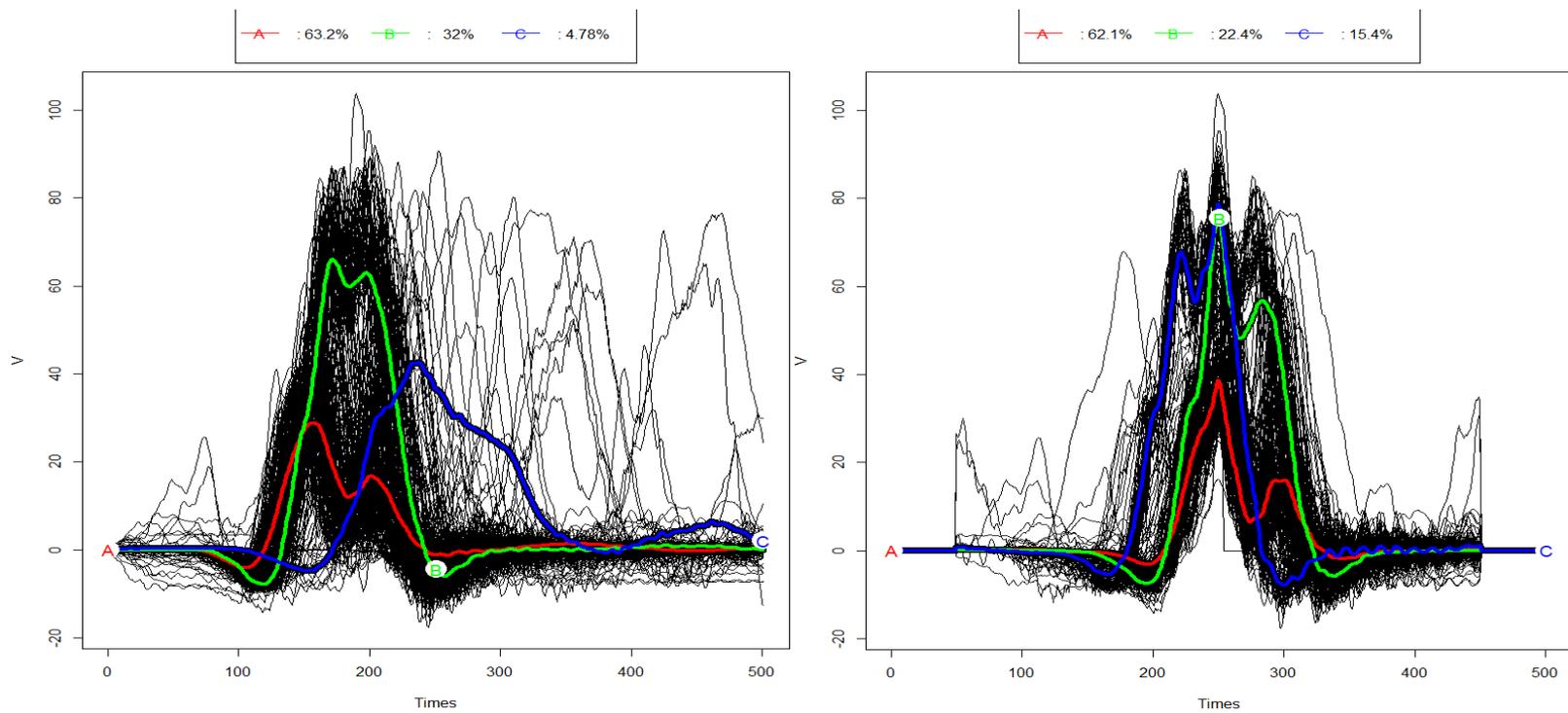


Figure 5.38: *Normalization 2, 3 cluster, dataset 1, non-normalized (left) and dataset 2, normalized (right).*

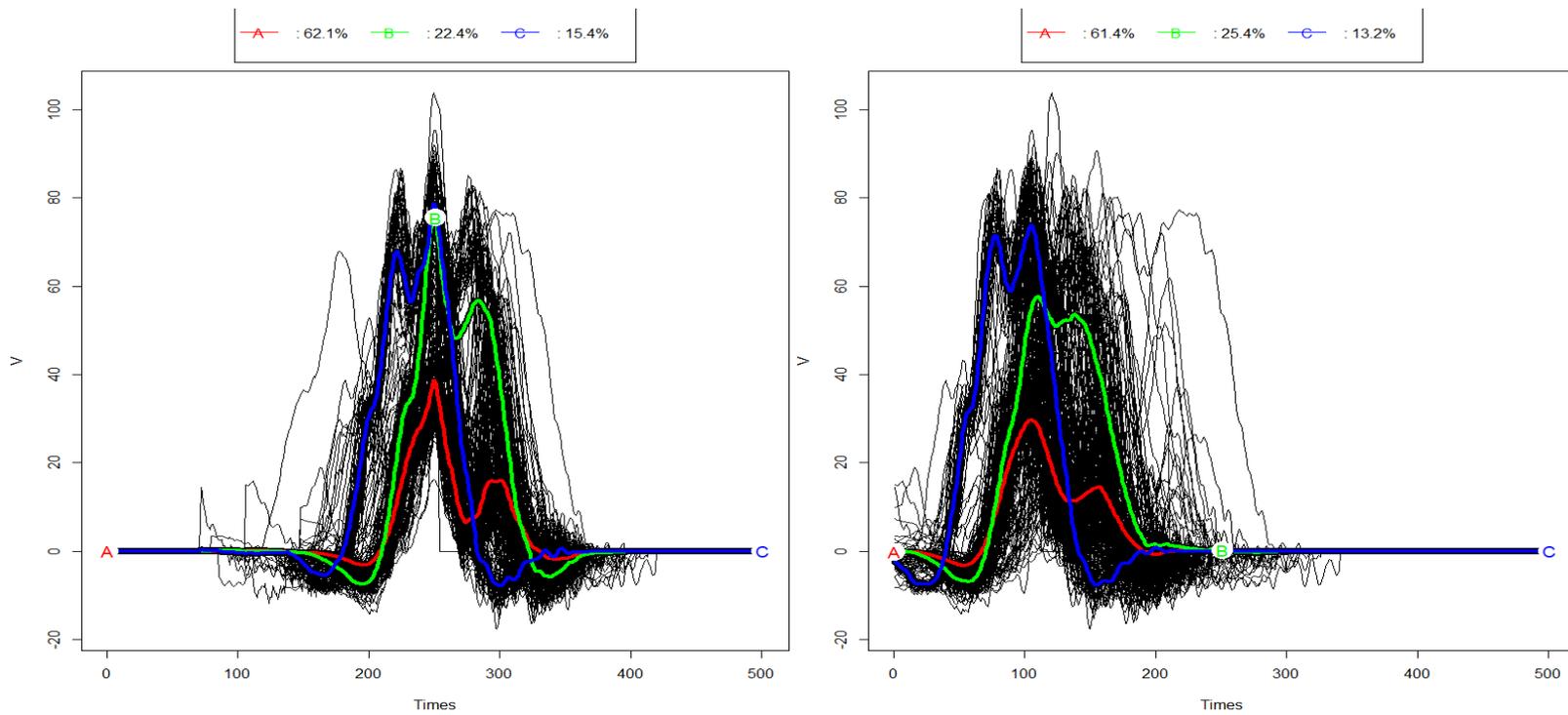


Figure 5.39: *Normalization 3*, 3 clusters, dataset 3, peak alignment (left) and dataset 4, entering-time alignment (right).

5.4.2 Monthly Classification Tree

After each monthly clustering analysis of the bridge response (step 1) follows the truck classification (step 2) using classification trees. As mentioned before, the classification tree provides a more quantitative description as opposed to a more qualitative description of the truck's typology, which, nevertheless, can be eventually deduced from the tree output. Tables 5.8 to 5.17 provide a visual overview of which truck's characteristics are used every month and which value determines each split. Fig.5.40 to 5.49 provide a visual overview of the parameters which are considered the most important based on their position and on the frequency with which they appear in the tree. Not all parameters are used for every month's analysis, some are consistently present such as GVW, speed, spacing 3-4, some are more common than others such as weight of axle 4 and length, while others only appear in one or two months. Overall, all parameters appear at least once but spacing 1-2. The features that all trees have in common are the following:

- The first split is always governed by GVW, which separates lighter from heavy trucks. GVW can also appear multiple times in a tree.
- Speed consistently separates A from B in the lighter GVW side, with A being always faster than B. However, splits governed by speed happen also for heavy trucks and not only between A and B.
- Spacing 3-4 always separates trucks with a large spacing and therefore a bimodal waveform (B) from non-bimodal waveform (C, D), but not limited to these clusters. It can also happen that sometimes D includes trucks with large spacing (Fig. 5.30, October 2010, blue cluster).
- From the two-step strategy conducted on March 2012, Lane 2, the weight of axle 6 seemed to contribute to the controversial separation of cluster C and D. This idea of having a

heavier rear part of the truck, is confirmed here where most of the time axle 4 weight and sometimes axle 5 weight are heavier in D than C.

- An additional piece of information regarding the separation between C and D is gained. Length splits C and D with D being consistently shorter than C.

Two comparisons are provided to see how practitioners can use classification tree to perform long-term monitoring and to acquire a general idea of what may happen during the following month and even during the same month in the following year. These values can be used to raise red flags in case of uncommon tree's structures.

- The first comparison is between two consecutive months (March 2012 and April 2012, Fig.5.50). Four splits stay the same with negligible differences in values and they are GVW (green) which splits all trucks in lighter and heavy, spacing 3-4 (orange) which splits bimodal from non-bimodal waveforms, speed (light blue) which separates A from B, also in the heavier side of GVW, and length (yellow) which separates D from C. Also, other splits are somehow related to each other, for example axle weight 4 in March 2012 and GVW in April 2012 (purple) agree that weight is influencing this split. Also, spacing 2-3 and axle weight 1, even if they are different quantities with different units, they are both related to the front part of the truck.
- The second comparison is between two consecutive years (March 2011 and March 2012, Fig.5.51). It is encouraging to see how the parameters which constitute the structure of the tree (GVW, spacing 3-4, speed and length) are consistent, suggesting reliability of this method to classify and monitor trucks over time.

Table 5.8: August 2010.

Aug-10	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	Length [m]
A	<27.6				<4.1								
A	<27.6				>4.1			>87.7					
A	27.6-32.1							>90.9		>8.4			
B	<27.6				>4.1			<87.7					
B	>27.6							<90.9		>8.4			
B	>32.1							>90.9		>8.4			
C	>27.6									<8.1			<16.6
D	>27.6									<8.4			<16.6
D	>27.6									8.1-8.4			>16.6

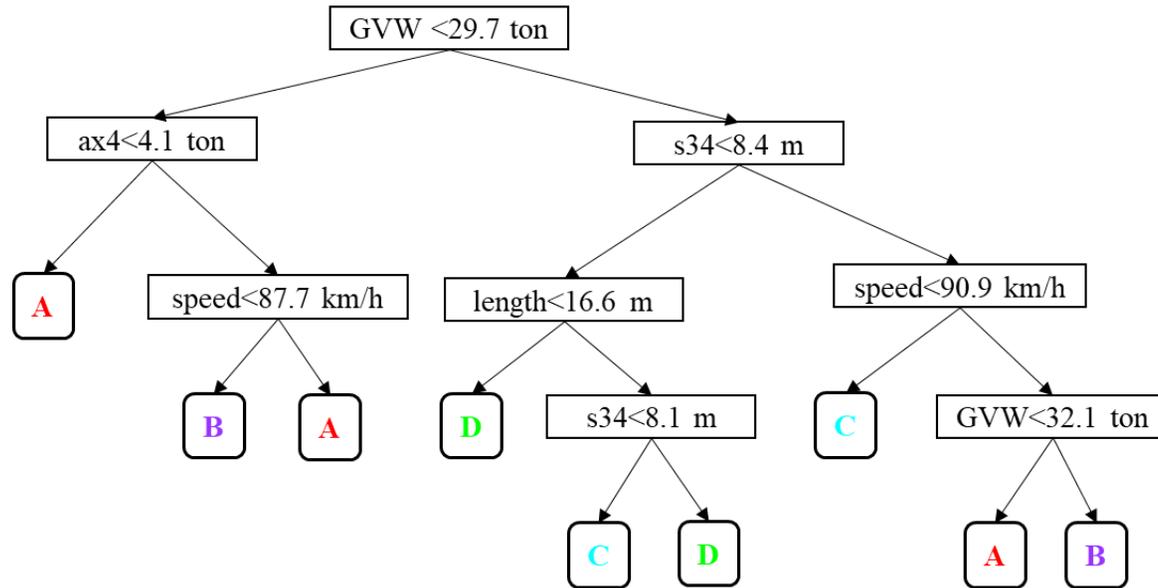


Figure 5.40: classification tree, August 2010.

Table 5.9: September 2010.

Sep-10	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	Length [m]
A	<26.7				<4.2								
A	<26.7				>4.2			>87.7					
A	>26.7			<6.4				>90.9		>8.3			
B	<26.7				>4.2			<87.7					
B	>26.7			>6.4						>8.3			
B	>26.7			<6.4				<90.9		>8.3			
C	>30.6									<8.3			
D	26.7-30.6									<8.3			

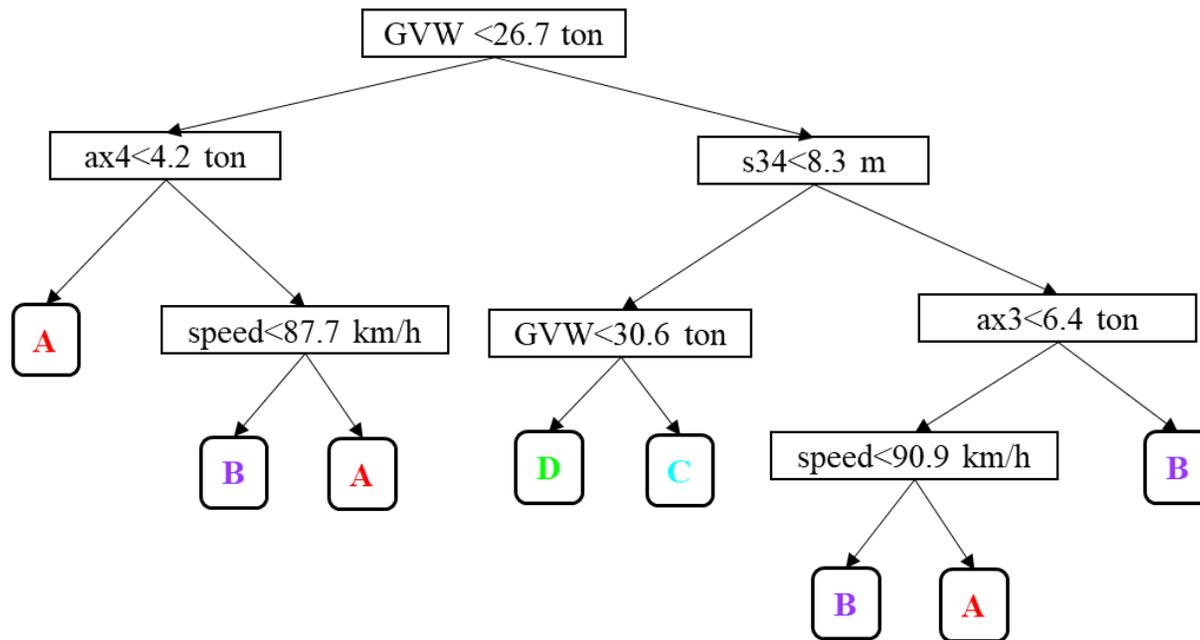


Figure 5.41: classification tree, September 2010.

Table 5.10: October 2010.

Oct-10	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	Length [m]
A	<28.7												
B	>28.7		>5.8							>8.3			
C	>28.7									<8.3	>2.8		
D	>28.7		<5.8							>8.3			
D	>28.7									<8.3	<2.8		

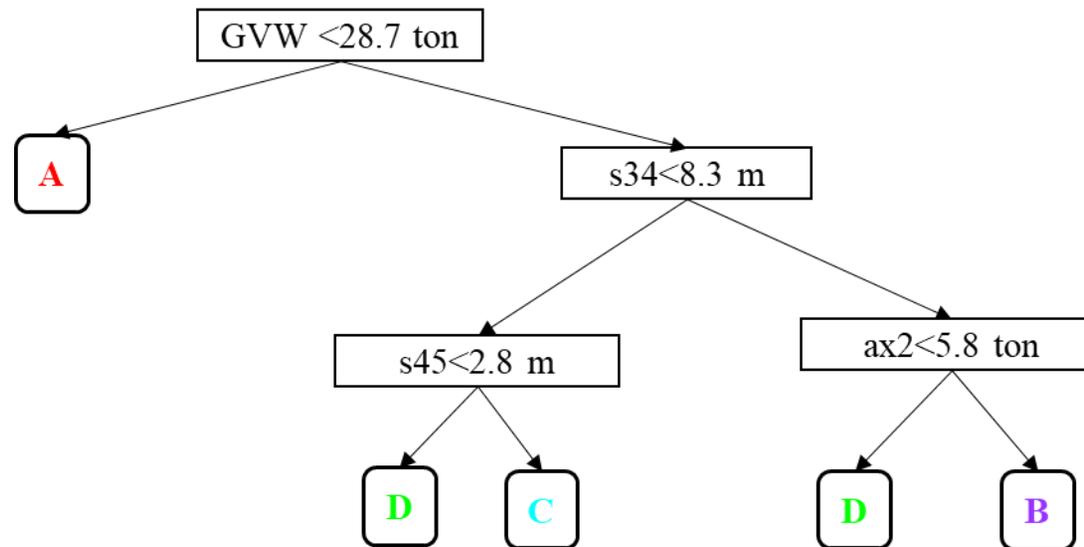


Figure 5.42: classification tree, October 2010.

Table 5.11: November 2010.

Nov-10	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	Length [m]
A	<27.5												
A	>27.5				<13.9			>90.9		>8.5			
B	>27.5				>13.9			>90.9		>8.5			>19.1
B	>27.5							<90.9		>8.5			
C	>27.5				>13.9			>90.9		>8.5			<19.1
C	>27.5									<8.5	>2.8		
D	>27.5									<8.5	<2.8		

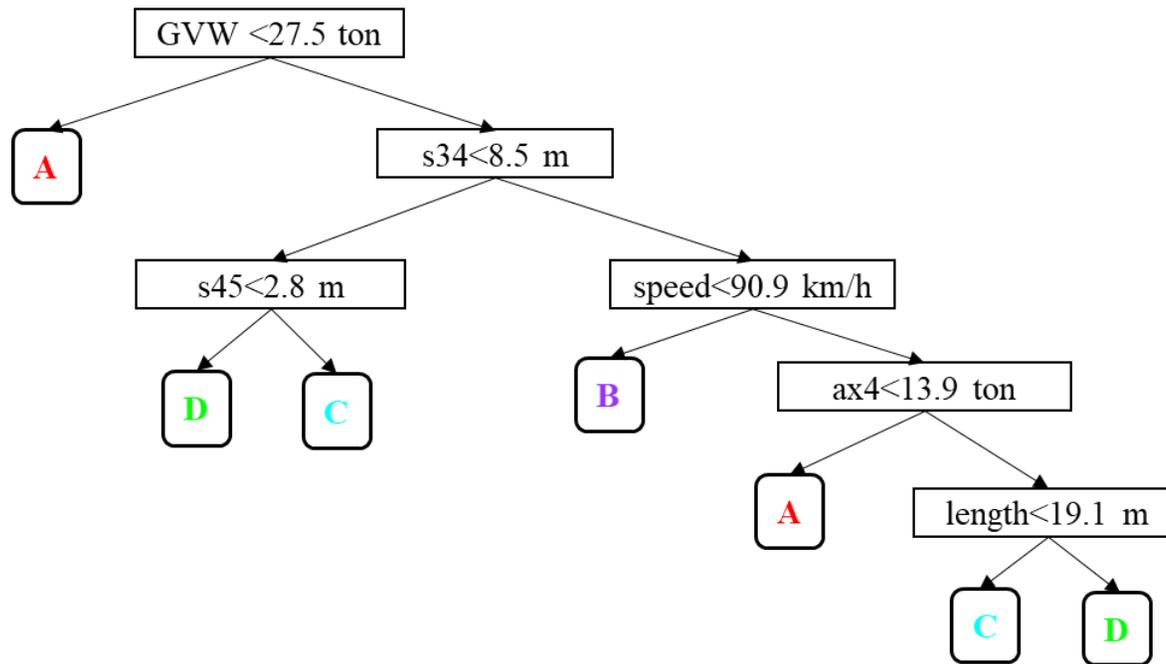


Figure 5.43: classification tree, November 2010.

Table 5.12: December 2010.

Dec-10	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	Speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	length [m]
A	<22.4				<3.7								
A	<29				>3.7			>86.1					
A	<29				>3.7			<86.1		<9.1			
B	<29				>3.7			<86.1		>9.1			
B	>29							<84.5		>8.2			
B	>29							>84.5		>8.2	>2.8		
B	>29							>84.5		>9.7	<2.8		
C	>29							>84.5		8.2-9.7	<2.8		
C	>29									<8.2		>2.9	
C	>29			>7						<8.2		<2.9	
D	>29			<7						<8.2		<2.9	

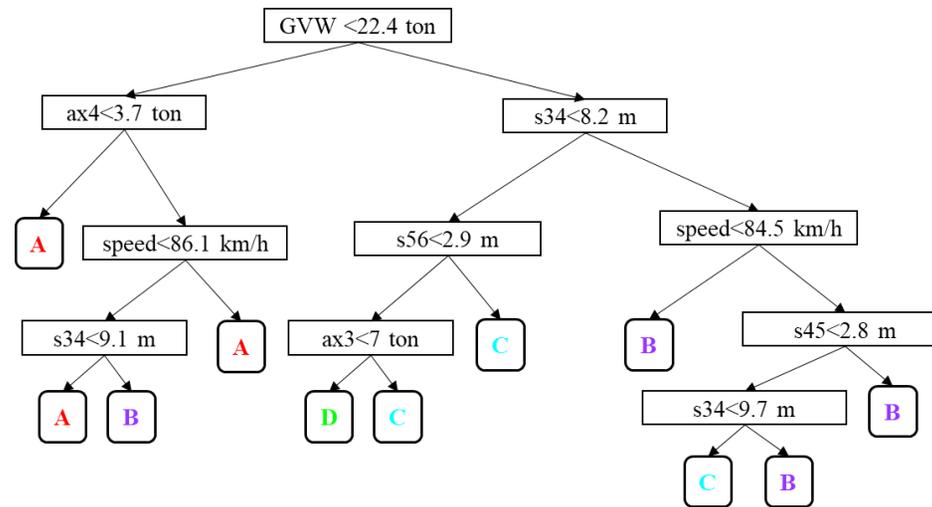


Figure 5.44: classification tree, December 2010.

Table 5.13: March 2011.

Mar-11	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	length [m]
A	<23.4												
A	23.4-28.5							>89.3					
B	23.4-28.5							<89.3					
B	>28.5					<8.3				>8.1			
C	>28.5									<8.1			>16.3
D	>28.5					>8.3				>8.1			
D	>28.5									<8.1			<16.3

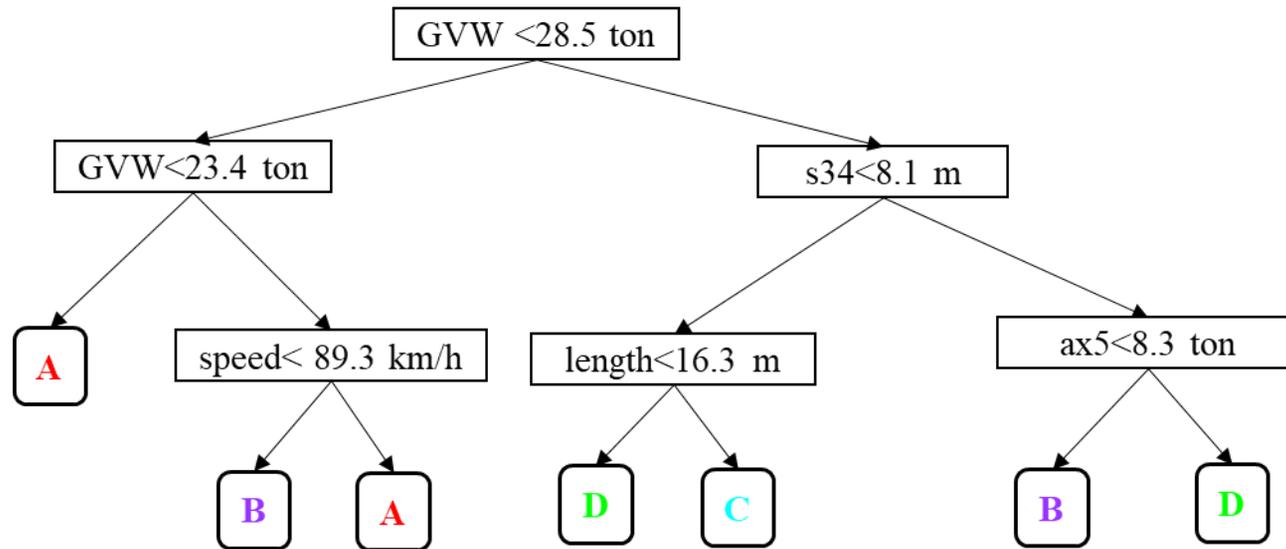


Figure 5.45: classification tree, March 2011.

Table 5.14: April 2011.

Apr-11	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	Length [m]
A	<28				<3.6								
A	<28				>3.6			>87.7					
A	<28				3.6-4.7			<87.7					
B	>28				>4.7			<87.7					
B	>28									>8.6			
C	>28									<8.6		>2.8	
D	>28									<8.6		<2.8	

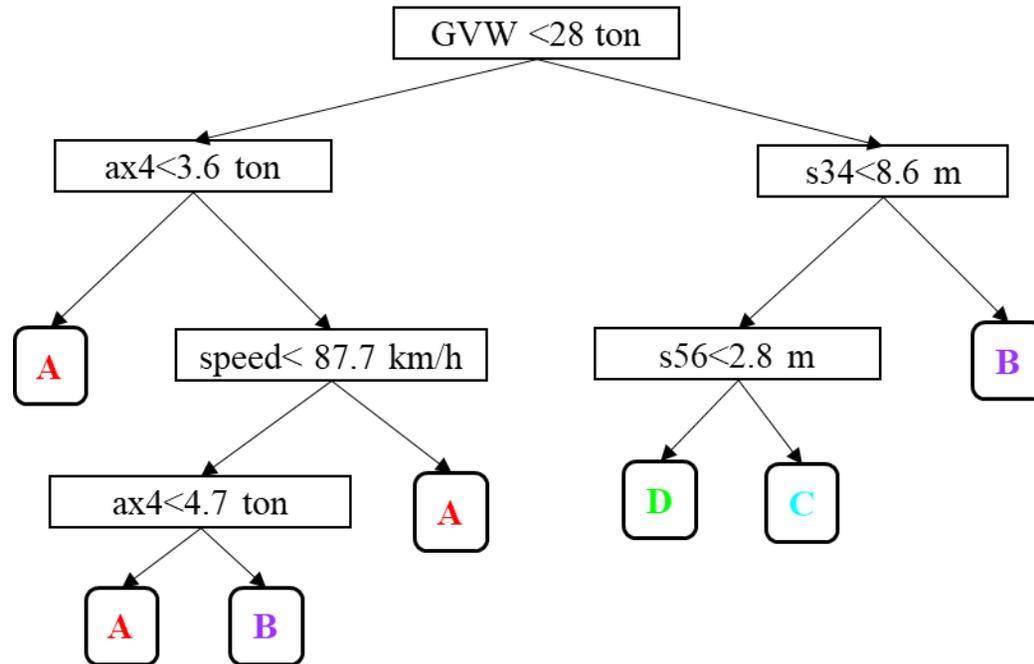


Figure 5.46: classification tree, April 2011.

Table 5.15: March 2012.

Mar-12	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	length [m]
A	<27.1				<4.1								
A	<27.1				>4.1			>82.9					
B	<27.1				>4.1			<82.9					
B	>27.1									>6.2			
C	>27.1								>2.1	<6.2			
C	>27.1								<2.1	<6.2			>17.2
D	>27.1								<2.1	<6.2			<17.2

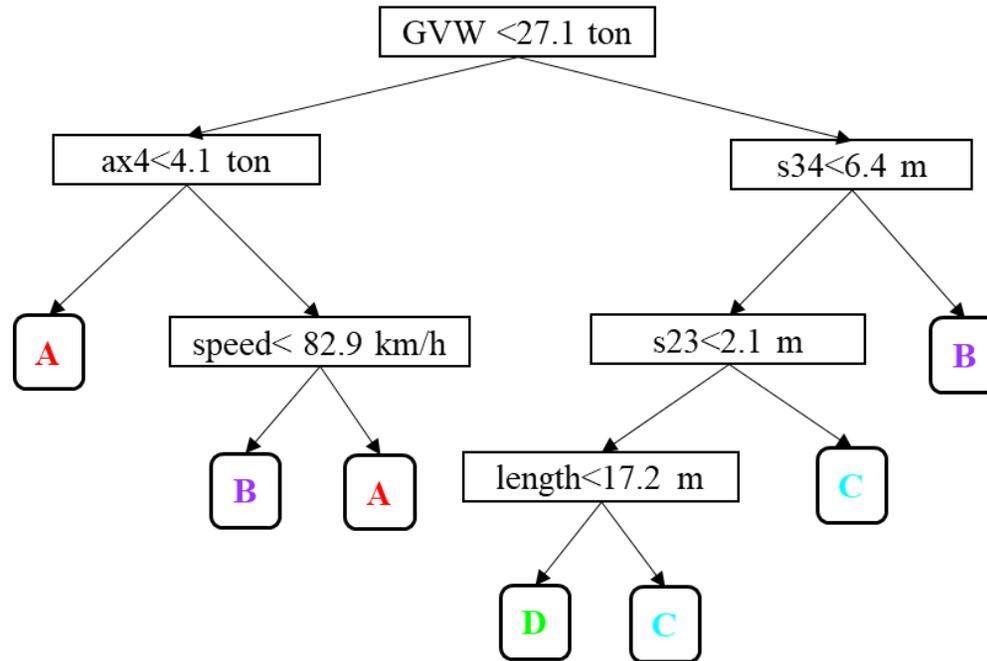


Figure 5.47: classification tree, March 2012.

Table 5.16: April 2012.

Apr-12	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	length [m]
A	<23												
A	23-26.7							>82.9					
A	26.7-30.5							>90.9		>7			
B	23-26.7							<82.9					
B	>30.5									>7			
B	26.7-30.5							<90.9		>7			
C	>26.7	>5.7								<7			
C	>26.7	<5.7								<7			>16
D	>26.7	<5.7								<7			<16

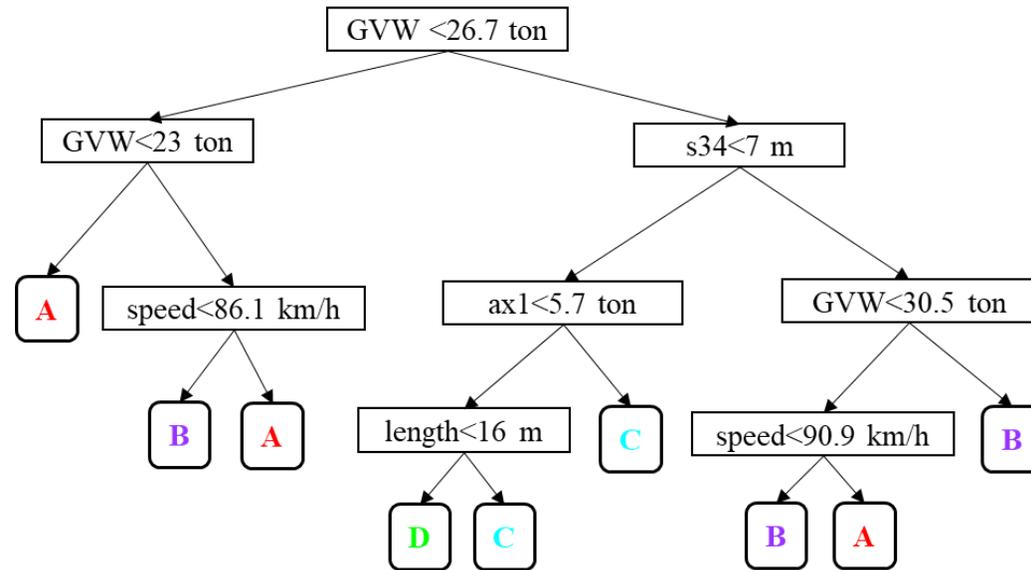


Figure 5.48: classification tree, April 2012.

Table 5.17: June 2012.

Jun-12	GVW [ton]	ax1 [ton]	ax2 [ton]	ax3 [ton]	ax4 [ton]	ax5 [ton]	ax6 [ton]	speed [km/h]	s23 [m]	s34 [m]	s45 [m]	s56 [m]	length [m]
A	<26.7				<3.8								
A	<26.7				>3.8			>78.1					
A	26.7-30.4						<6.4	>92.5		>8.1			>19.1
B	<26.7				>3.8			<78.1					
B	>26.7									>8.1			<19.1
B	>30.4						<6.4			>8.1			>19.1
B	26.7-30.4						<6.4	>92.5		>8.1			>19.1
C	>26.7						>6.4			>8.1			>19.1
D	>26.7									<8.1			

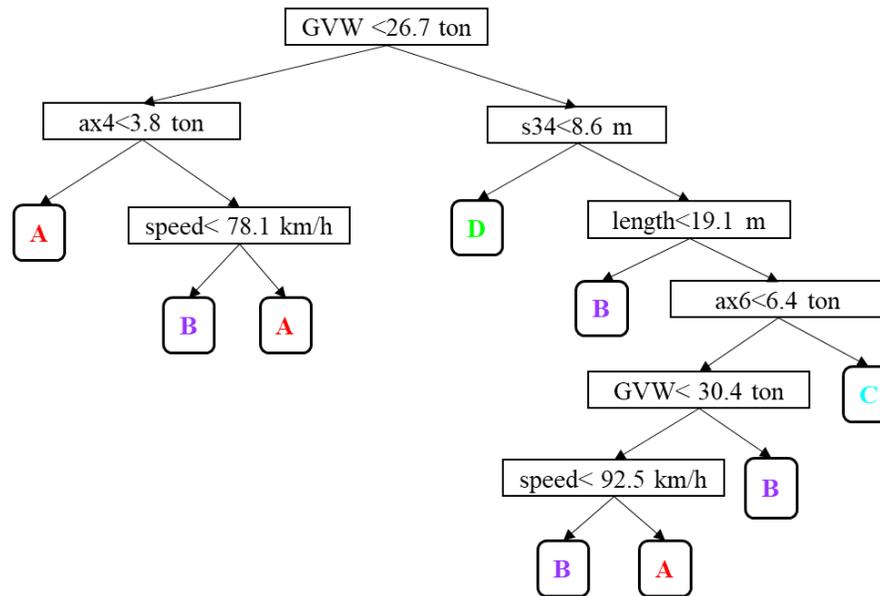


Figure 5.49: classification tree, June 2012.

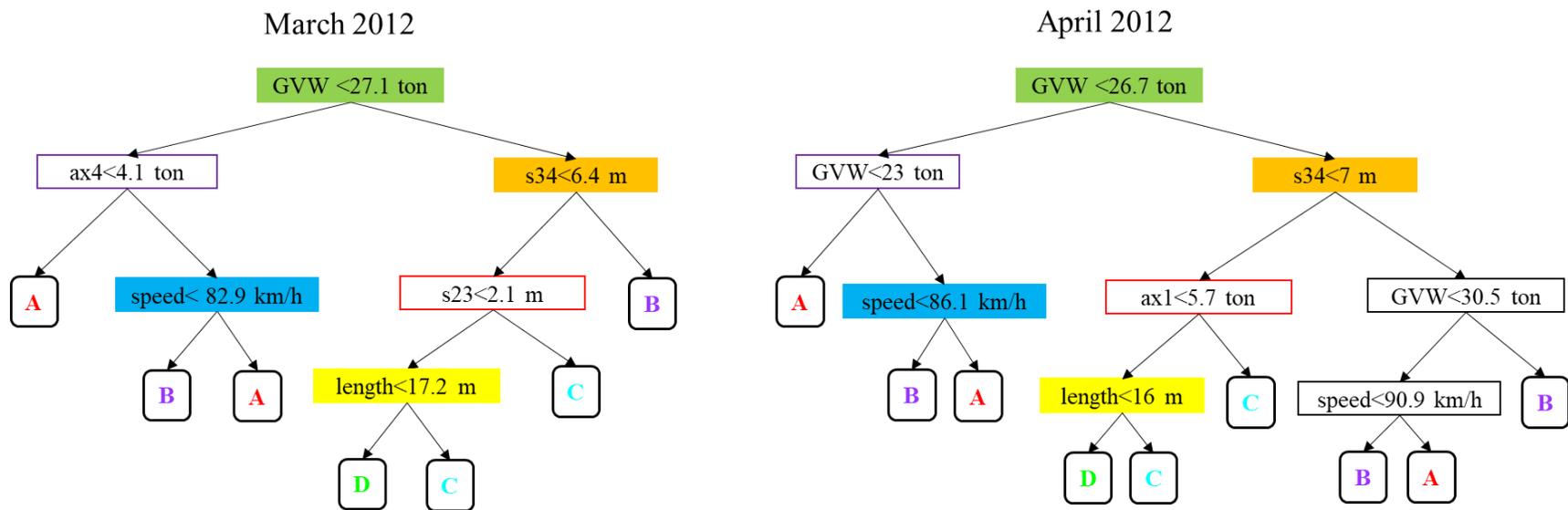


Figure 5.50: comparison of classification trees between two consecutive months.

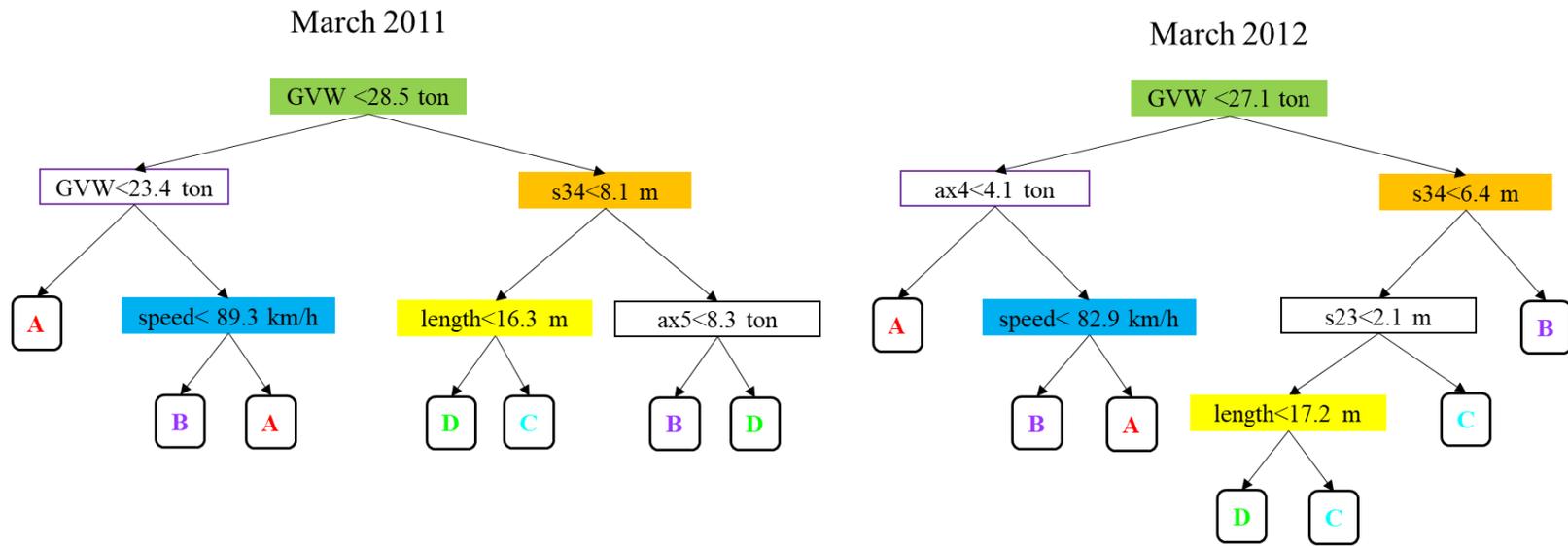


Figure 5.51: comparison of classification trees between two months in two consecutive years.

5.5 CONCLUSIONS

This chapter aims to develop a two-step strategy to cluster bridge response and classify the truck traffic which regularly crosses the subject structure. The goal of this strategy is to provide support to decision-making and planning, in order to assure better maintenance and therefore improved serviceability of the infrastructure.

Step 1 of the two-step strategy is performed using a machine learning method called longitudinal clustering (KmL). This method belongs to the class of unsupervised learning, meaning that the right answer is not known a priori. KmL assigns each strain waveform, which represents the response of the bridge to a specific type of truck, to a cluster by calculating the distance between the waveform and the cluster's centroid. Waveforms with similar shapes are grouped together. The best number of cluster to use in the analysis is an important decision to make. This decision can be supported by some quality criteria such as Calinski-Harabatz, and other consideration of practical nature.

Step 2 of the two-step strategy can be performed with different methods depending on the data availability. Here, two techniques are proposed, first, manual image processing, when photos of the vehicles are available, which aims to classify the type of truck, second, classification tree which is a machine learning method, which aims to provide a more quantitative description of the truck's characteristics.

Results show that the bridge response can reasonably be grouped in four categories, however, other ad hoc scenarios can be considered. The combined use of image processing and classification tree allowed to build a truck profile for each response of the bridge identified by the clustering

analysis. Either method can also be used individually to obtain valid truck classification. The four bridge responses are clustered as follows:

- Cluster A is caused by lighter weight trucks, with large spacing which produce bimodal strain waveforms (tank trucks, flatbed, regular trucks);
- Cluster B is caused by heavier trucks with large spacing which produce bimodal strain waveforms (regular trucks);
- Clusters C and D are mainly caused by logging trucks, with more equally spaced axles to support extremely heavy loads. Cluster D can sometimes be caused by trucks with large spacing as well.

According to the classification tree outcome, GVW, speed, spacing 3-4, length and weight of axle 4 are the most important features which determine the clusters distinction. GVW first separates lighter from heavy trucks, spacing isolates cluster B from the other clusters, speed separates cluster A from B, length, and weight of rear axles separate C from D.

Each cluster can be linked to a specific problem to the bridge. For instance, trucks which cause a bimodal waveform can cause fatigue problems because the bridge perceives one crossing as almost two different cycles, such as cluster B. These types of trucks should be taken under control. Another problem is the presence of overweight trucks which can speed up the deterioration of the bridge and its component. The two-step strategy can successfully extract pivotal information regarding the bridge performance and its traffic which can facilitate the planning of specific maintenance programs, the development of ad hoc regulation, more in general, support decision-making and the overall better management of the structure.

6 CONCLUSIONS

6.1 OVERVIEW

America's highway network is underperforming, and while not all deficient bridges are in danger of collapse, there are consequences that come with their decayed state. One of them is the need to impose weight restrictions which can have an impact on the routes of heavy trucks, and, more in general, on the economic activity. It has proven challenging for state and local governments to keep pace with the growing nation's bridge needs; for this reason, it is crucial to find more effective strategies to allocate resources by assuring timely and effective interventions to improve long-term durability and serviceability.

This goal can be achieved by understanding bridge performance and monitoring of bridges to detect structurally deficient components in time, to properly plan maintenance and promote data-driven decision-making to assure an overall improved management of the structure.

Adopting SHM techniques on critical bridges can contribute to addressing some of today's challenges and improving inspection, repair, and rehabilitation methods and reducing traffic disruption. Keeping America's roads and bridges in a state of good repair can positively contribute to a robust economy and favorably impact the quality of life. As a result, timely identification of potential problems can help mitigate their impact on structural health and reduce bridge rehabilitation costs, extending the service life and minimizing life-cycle cost of bridge networks.

6.2 CONTRIBUTION

The overall objective of this work is to support asset management and decision-making process, not to provide warning of incipient failure, but rather reduction in some aspect of

performance or serviceability. Application of structural health monitoring instrumentation and analysis can provide timely decision-support information to extend service life and promote ongoing serviceability. This thesis aims to provide all the key elements to monitor and assess bridge conditions which are: a dataset, which includes both traffic information and bridge response; appropriate metrics to monitor the bridge performance; efficient predictive models; bridge response and truck classification strategy to support planning and managing, as well as, facilitate regulation.

Some preliminary studies were necessary to understand the health condition of the subject bridge. For this purpose, the GDFs were computed and compared to AASHTO specifications. The long-term behavior of the bridge over the 3-year period was analyzed at different level of aggregation. Changes in the trends and patterns were identified, and linked to potential causes such as logging season, extreme weather conditions and Wisconsin statutory change, which allowed vehicle combination up to 98 kips on six axles to transport loads of raw forest products during the spring thaw suspension period. Overall, the bridge is in healthy conditions with no evidence of structural loss of integrity such as anomalous repartition of the load on the girders.

Long-term monitoring is performed proficiently when the best selection of instruments, thoughtfully placed, is deployed. Assessing the integrity of existing bridges requires accurate monitored bridge data and adequate interpretation. The need for an effective use of monitoring data promoted the development of novel metrics to monitor how the bridge responds to heavy traffic. The first metric (SDA) proposed aims to overcome some limitations of traditional strain measurements. It aims to provide a generalized measure of the response of the bridge as opposed to a localized measure; to show reduction of the dependency on length, resulting in significant reduction of scatter in data; to provide an alternative way to calculate GDF and to efficiently detect

reading errors of from the monitoring systems, as well as hidden anomalies which can lead to response underestimation. The second metric (α_e) aims to monitor the bridge rigidity/flexibility over time by constantly comparing it to the design values (α_d). Any detected change over time has the potential to raise red flags regarding the conditions of the bridge rigidity/flexibility.

Once a valid metric to study the bridge response is defined, the next step is to distill information out of bridge monitoring data, to aid infrastructure owners in evaluating bridge performance and making data-driven management decisions. Machine learning methods are deployed to model and predict bridge performance. First, a comparative study of Multilinear Regression, Artificial Neural Network, and Regression Tree, is presented to analyze advantages and disadvantages and to identify the most desirable features to model the bridge performance. Second, an alternative strategy is proposed for practitioners to avoid trading off important capabilities. The results of the comparative study showed that all three models fitted the data well, with ANN confirming its superior predictive power. Nevertheless, predictive power is not the only valuable capability to evaluate the proficiency of a model, indeed, four criteria are discussed to select the best model: predictive power, explanatory capabilities, interpretability, and computation speed. ANN's output showed lack of interpretability and explanatory capabilities, which is not ideal if the primary goal is to explain the origin of certain response patterns. MLR's output was easily interpreted, and its coefficients provided valuable insights on how the predictors drive the changes in the response. The results of RT provided insight on the importance of the predictors, and they were easily interpreted thanks to the user-friendly tree shape. The outcome of the comparative study provided a complete overview of each method's capabilities allowing an alternative solution to be defined which eventually combined benefits of different methods into one. It was proven that the IOA could drive improvements to all three methods. The IOA improves the predictive power of MLR

which already has the potential to enjoy more support among practitioners given its elevated interpretability and explanatory capabilities. For this reason, the combination of IOA and MLR was suggested as the alternative strategy to ANN, that could be used for long-term monitoring of shifts in the regression parameters which could indicate a change of the bridge's condition that warrants further investigation.

Planning and managing infrastructures can certainly benefit from predictive models of bridge response. Assuming that all vehicles have a similar impact on the bridge would lead to significant errors, therefore, to plan effective maintenance interventions, it is important to study in depth what is causing different bridge responses and what type of traffic should be taken more under control because of its potential impact of the structure. To achieve this goal a two-step strategy which provides a complete overview of the bridge response to the heavy traffic which regularly crosses the subject bridge, is proposed. Step 1 is performed using longitudinal clustering. The choice of number of clusters can be facilitated with quality criteria and other consideration of practical nature. Step 2 can be performed with different methods depending on the data availability, such as manual image processing, which aims to classify the type of truck and/or classification tree, which aims to provide a more quantitative description of the truck's characteristics. The identified groups of bridge responses were, bimodal-low strain caused by lighter trucks, bimodal-high strain caused by heavy trucks, non-bimodal-very high strain mainly caused by logging trucks. Each cluster can be linked to a potential problem to the bridge. For instance, high frequency of trucks which trigger bimodal waveforms can contribute to fatigue damage and heavy and overweight trucks significantly contribute to the reduction of the service life of pavements and bridges. This data-driven strategy can facilitate the planning of specific maintenance programs, the development of

ad hoc regulations, more in general, support decision-making and the overall better management of the structure.

6.3 LIMITATIONS AND FUTURE WORK

This thesis addressed multiple aspects of the bridge monitoring, using a unique dataset resulted from the fruitful collaboration between WisDOT and Northwestern ITI. The term “unique” is used with a two-fold meaning, first, unique in the sense that it is rare to have a long-term dataset with both traffic (input) and bridge (output) data of such quality, second, unique meaning that all the analyses were developed on one bridge. Although the subject bridge is very representative of many others in the country for its characteristics, it would be excellent to test these results on other bridges. A future step it would be to apply machine learning image recognition algorithms to automatize and speed up the image processing in the case of large databases. Another future plan would be to compare classification tree results obtained with other methods such as XGBoost and Support Vector Machine and compare their misclassification rate.

It is hoped that this thesis, by showing how statistical methods can be critical to the process of distilling information out of the bridge monitoring, could provide a guide for practitioners to support data-driven decision-making, improve the maintenance planning process, and overall, to increase serviceability of bridges and infrastructures.

REFERENCES

CHAPTER 1

- Agdas, D., Rice, J.A., Martinez, J.R. and Lasa, I.R. (2015). Comparison of visual inspection and structural-health monitoring as bridge condition assessment methods. *Journal of Performance of Constructed Facilities*, 30(3).
- American Road and Transportation Builders Association (ARTBA) (2018). 2018 Structural Deficient Bridge Report. www.artbabridgereport.org/; www.artba.org/wp-content/deficient_bridge/ARTBA_State_Ranking_2018.pdf
- American Society of Civil Engineers (ASCE) (2017). 2017 Infrastructure Report Card. www.infrastructurereportcard.org
- Center for Automotive Research (CAR), Michigan Tech Transportation Institute (MTTI), Michigan Tech Research Institute (MTRI) and Michigan Department of Transportation (MDOT) (2012). Economic evaluation of commercial remote sensors for bridge health monitoring.
- Chen, Y., Corr, D.J. and Durango-Cohen, P.L. (2014). Analysis of common-cause and special-cause variation in the deterioration of transportation infrastructure: A field application of statistical process control for structural health monitoring. *Transportation Research Part B: Methodological*, 59, 96–116.
- Chen, Y., and Durango-Cohen, P.L. (2015). Development and field application of a multivariate statistical process control framework for health-monitoring of transportation infrastructure. *Transportation Research Part B: Methodological*, 81, 78-102.

- Cusson, D., Ghuman, P., Gara, M. and McCardle, A. (2012). Remote monitoring of bridges from space. 54^o Congresso Brasileiro do Concreto.
- DeWolf, J.T., Lauzon, R.G., and Culmo, M.P. (2002). Monitoring bridge performance. *Structural Health Monitoring*, 1(2), 129–138. <https://doi.org/10.1177/1475921702001002001>
- Egan, M.J. (2007). Anticipating future vulnerability: defining characteristics of increasingly critical infrastructure-like systems. *Journal of Contingencies Crisis Management*, 15(1), 4– 17.
- European Space Agency (ESA) (2016). Monitoring bridge from space. <https://business.esa.int/projects/showcases/monitoring-bridges-space>
- Greenfield, B. (2014). Saving \$5M by monitoring bridges. <https://efficientgov.com/blog/2014/09/02/saving-5m-monitoring-bridges/>
- Halsey III, A. (2016). Nearly 59,000 bridges in U.S. are structurally deficient. *Washington Post*. www.washingtonpost.com/news/dr-gridlock/wp/2016/02/18/nearly-59000-bridges-in-u-s-are-structurally-deficient/?utm_term=.42ea4ecdce5e
- Jaffe, E. (2015). America's infrastructure crisis is really a maintenance crisis. <https://www.citylab.com/solutions/2015/02/americas-infrastructure-crisis-is-really-a-maintenance-crisis/385452/>
- Kahn, M.E. and Levinson, D.M. (2011). Fix it first, expand it second, reward it third: a new strategy for America's highways. *The Hamilton Project*.
- Kosnik, D.E. (2012). A combined weigh-in-motion and structural health monitoring system on a Wisconsin-Michigan border bridge. In *International Conference on Weigh-in-Motion: ICWIM 6*. London, ISTE; Hoboken, 440–449.

- Monga, V. (2015). Neglected roads and bridges take toll on us companies. The Wall Street Journal. www.wsj.com/articles/neglected-roads-and-bridges-take-toll-on-u-s-companies-1424740746
- U.S. Department of Treasury with the Council of Economic Advisers (2012). A new economic analysis of infrastructure investment. p. 3 www.treasury.gov/resource-center/economic-policy/Documents/20120323InfrastructureReport.pdf
- Westgate, R.J., Koo K.Y. and Brownjohn, J.M.W. (2011). Environmental Effects on a Suspension Bridge's Dynamic Response. Proceedings of the 8th Intl. Conference on Structural Dynamics, Leuven, Belgium.

CHAPTER 2

- AASHTO LRFD Bridge Design Specifications. Customary U.S. Units/4th Edition; 2007. American Association of State Highway and Transportation Officials.
- AASHTO Standard Specifications for Highway Bridges, Bridge Design Specifications. AASHTO, Washington, D.C., (1996).
- Bae, H-U. and Oliva, M. (2009). User's guide: bridge analysis and evaluation of effects under overload vehicles. National Center for Freight and Infrastructure Research and Education CFIRE 02-03.
- Barr, P.J., Eberhard, M.O. and Stanton, J.F., (2001). Live-load distribution factors in prestressed concrete girder bridges. Journal of Bridge Engineering ASCE, 6(5), 298-306.
- Brendler, S.A. and Khodair, Y.A. (2015). A parametric study of live load distribution factors for steel girder integral abutment bridge: camera-ready guidelines. 5th International Workshop on Performance, Protection, and Strengthening of Structures under Extreme Loading,

413– 420.

- Cardini, A.J. and DeWolf, J.T. (2008). Long-term structural health monitoring of a multi-girder steel composite bridge using strain data. *Structural Health Monitoring*, 8(1).
- Dai, C. (2013). Exploring Data Quality of Weigh-in-Motion Systems. Thesis.
- Eom, J. and Nowak, A.S. (2001). Live load distribution for steel girder bridges. *Journal of Bridge Engineering ASCE*, 6(6).
- Ghosn, M., Moses, F. and Gobieski, J. (1986). Evaluation of steel bridges using in-service testing. *Transportation Research Record*, (1072), 71–78.
- Kennedy, J.B. and Grace, N.F. (1983). Load distribution in continuous composite bridges. *Canadian Journal of Civil Engineering*, 10, 384-395.
- Kim, S. and Nowak, A.S. (1997). Load distribution and impact factors for I-Girders bridges. *Journal of Bridge Engineering ASCE*, 2(3), 97–104.
- Neely Jr. W.D. (2001). Evaluation of the in-service performance of the Tom’s Creek bridge”. Master Thesis.
- Nowak, A.S., Eom, J., Sanli, A. and Till, R. (2003). Verification of girder distribution factors for continuous steel girder bridges by field testing. *Transportation Research Record*, (99), 62– 67.
- Olund, J. and DeWolf, J. (2007). Passive structural health monitoring of Connecticut’s bridge infrastructure. *Journal of Infrastructure Systems ASCE*, 13(4), 330–339.
- Owusu-Ababio, S. and Schmitt, R. (2014). Evaluation of impacts of allowing heavier log loads in northern Wisconsin during spring thaw. Technical Report, 2011-2013.
- Reiff, A.J., Sanayei, M. and Vogel, R.M. (2015). Statistical bridge damage detection using girder distribution factors. *Engineering Structures*, 109, 139-151.

- Shenton III, H.W., and Hu, X. (2006). Damage identification based on dead load redistribution: methodology. *Journal of Structural Engineering, ASCE*, 132(8)
- Stallings, J. and Yoo, C. (1993). Tests and ratings of short-span steel bridges. *Journal of Bridge Engineering ASCE*, 119(7), 2150–2168.
- Tabsh, S.W. and Tabatabai, M. (2001). Live load distribution in girder bridges subject to oversized trucks. *Journal of Bridge Engineering ASCE*, 6(1).
- Tennyson, R.C., Mufti, A.A., Rizkalla, S., Tadros, G. and Benmokrane, B. (2001). Structural health monitoring of innovative bridges in Canada with fiber optic sensors. *Smart Material and Structures*, 10(3).
- Vincent, G.S., (1969). Tentative criteria for load factor design of steel highway bridges. American Iron and Steel Institute (Bulletin No. 15 Washington, D.C.).
- Zokaie, T., Imbsen, R.A. and Osterkamp, T.A. (1991). Distribution of wheel loads on highway bridges. Final Report to NCHRP 12-26.

CHAPTER 3

- Bergmeister, K. and Santa, U. (2001). Global monitoring concepts for bridges. *Structural Concrete*, 2(1), 29– 39.
- Buckley, E. (1998). Basic influence line equations for continuous beams. *Journal of Structural Engineering ASCE*, 123(10), 1416–1420.
- Cardini, A.J. and DeWolf, J.T. (2008). Long-term structural health monitoring of a multi-girder steel composite bridge using strain data. *Structural Health Monitoring*, 8(1), 47–58.
- Chen, Y., Corr, D.J. and Durango-Cohen, P.L. (2014). Analysis of common-cause and special-

- cause variation in the deterioration of transportation infrastructure: A field application of statistical process control for structural health monitoring. *Transportation Research Part B: Methodological*, 59, 96–116. <http://dx.doi.org/10.1016/j.trb.2013.11.002>.
- DeWolf, J.T., Lauzon, R.G. and Culmo, M.P. (2002). Monitoring bridge performance. *Structural Health Monitoring*, 1(2), 129–138.
- Feng, M.Q., Kim, D.K., Yi, J. and Chen, Y. (2004). Baseline models for bridge performance monitoring. *Journal of Engineering Mechanics ASCE*, 130(5), 562–569.
- Frangopol, D.M., Strauss, A. and Kim, S. (2008). Bridge reliability assessment based on monitoring. *Journal of Bridge Engineering ASCE*, 13(3), 258–270, [http://ascelibrary.org/doi/abs/10.1061/\(ASCE\)1084-0702\(2008\)13:3\(258\)](http://ascelibrary.org/doi/abs/10.1061/(ASCE)1084-0702(2008)13:3(258)).
- Gagarin, N., Flood, I. and Albrecht, P. (1994) Computing truck attributes with artificial neural networks. *Journal of Computing in Civil Engineering ASCE*, 8(2), 179–200.
- Hirachan, J. and Chajes, M. (2005). Experimental influence lines for bridge evaluation. *Bridge Structures*, 1(4), 405–412, <http://content.iospress.com/doi/10.1080/15732480600578485>.
- Ieng, S.S. (2014). Bridge influence line estimation for bridge weigh-in-motion system. *Journal of Computing in Civil Engineering ASCE*, 29(1), 1–4.
- Lansdell, A., Song, W. and Dixon, B. (2017). Development and testing of a bridge weigh-in-motion method considering nonconstant vehicle speed. *Engineering Structures*, 152, 709–726, <https://doi.org/10.1016/j.engstruct.2017.09.044>
- Liu, M. and Frangopol, D.M. (2006). Probability-based bridge network performance evaluation.

- Journal of Bridge Engineering ASCE, 11(5), 633–641.
- Liu, M., Frangopol, D.M. and Kim, S. (2009b). Bridge safety evaluation based on monitored live load effects. *Journal of Bridge Engineering ASCE*, 14(4), 257–269.
- Lydon, M., Taylor, S.E., Robinson, D., Mufti, A., O'Brien, E.J. (2016). Recent developments in bridge weigh in motion (B-WIM). *Journal of Civil Structural Health Monitoring*, 6(1), 69– 81.
- McNulty, P. and O'Brien, E.J. (2003). Testing of bridge weigh-in-motion system in sub-arctic climate. *Journal of Testing and Evaluation*, 31(6), 497–506.
- Moses, F. (1979). Weigh-in-motion system using instrumented bridges. *Journal of Transportation Engineering ASCE*, 105(3), 233–49.
- O'Brien, E.J., Quilligan, M.J. and Karoumi, R. (2006). Calculating an influence line from direct measurements. *Proceedings of the Institution of Civil Engineers - Bridge Engineering*. 159(1), 31–34. <http://www.icevirtuallibrary.com/doi/10.1680/bren.2006.159.1.31>.
- Ojio, T., Carey, C.H., O'Brien, E.J., Doherty, C. and Taylor, S.E. (2016). Contactless bridge weigh-in-motion. *Journal of Bridge Engineering ASCE*, 21(7), [http://dx.doi.org/10.1061/\(ASCE\)BE.1943-5592.0000776](http://dx.doi.org/10.1061/(ASCE)BE.1943-5592.0000776)<http://ascelibrary.org/doi/10.1061/%28ASCE%29BE.1943-5592.0000776>.
- Štimac-Grandić, I., Grandić, D. and Bjelanović, A. (2011). Comparison of techniques for damage identification based on influence line approach. *Machines, Technologies, Materials*, 2(2).

Strauss, A., Wendner, R., Frangopol, D.M. and Bergmeister, K. (2012). Influence line-model correction approach for the assessment of engineering structures using novel monitoring techniques. *Smart Structures and Systems*, 9(1), 1–20, <http://koreascience.or.kr/journal/view.jsp?kj=KJKHFZ&andpy=2012&vnc=v9n1&andsp=1>.

Wall, C.J., Christenson, R.E., McDonnell, A.M.H. and Jamalipour, A. (2009). A non-intrusive bridge weigh-in-motion system for a single span steel girder bridge using only strain measurements. University of Connecticut, Connecticut Transportation Institute, Storrs, CT, USA, Report CT-2251-3-09-5, 57.

Wang, N.B., He, L.X., Ren, W.X. and Huang, T.L. (2017). Extraction of influence line through a fitting method from bridge dynamic response induced by a passing vehicle. *Engineering Structures*, 151, 648–664. <http://dx.doi.org/10.1016/j.engstruct.2017.06.067>.

Zhao, H., Uddin, N., O'Brien, E.J., Shao, X. and Zhu, P. (2014). Identification of vehicular axle weights with a bridge weigh-in-motion system considering transverse distribution of wheel loads. *Journal of Bridge Engineering ASCE*, 19(3). <http://ascelibrary.org/doi/10.1061/%28ASCE%29BE.1943-5592.0000533>.

CHAPTER 4

Adeli, H. (2001). Neural networks in civil engineering 1989-2000. *Computer-Aided Civil and Infrastructure Engineering*, 17, 105–112. doi: 10.1111/0885-9507.00219

Adeli, H. and Panakktat, A. (2009). A probabilistic neural network for earthquake magnitude prediction, *Neural Networks*, 22(7), 1018–1024. <https://doi.org/10.1016/j.neunet.2009.05.003>

- Adeli, H. and Yeh, C. (1989). Perceptron learning in engineering design. *Microcomputers in Civil Engineering*, 247–256. doi: 10.1111/j.1467-8667.1989.tb00026.x
- Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3), 370–374. doi: 10.1002/wics.84
- Bollen, K.A. and Jackman, R.W. (1985). Regression diagnostics: an expository treatment of outliers and influential cases, *Sociological Methods and Research*, 13(4), 510-542. <http://journals.sagepub.com/doi/abs/10.1177/0049124185013004004>
- Browne, M.W. (2000). Cross-Validation methods. *Journal of Mathematical Psychology*, 44(1), 108-132. <https://doi.org/10.1006/jmps.1999.1279>
- Burke, L.I. (1991). Introduction to artificial neural systems for pattern recognition. *Computers and Operations Research*, 18(2), 211–220. [https://doi.org/10.1016/0305-0548\(91\)90091-5](https://doi.org/10.1016/0305-0548(91)90091-5)
- Cattan, J. and Mohammadi, J. (1997). Analysis of bridge condition rating data using neural networks. *Microcomputers in Civil Engineering*, 12(6), 419–429. doi:10.1111/0885-9507.00074
- Cook, R.D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19(1), 15–18. <https://doi.org/10.1080/00401706.1977.10489493>
- Cook, R.D. (1979). Influential observations in linear regression influential observations in linear regression. *Journal of the American Statistical Association*, 74(365), 169–174. <http://www.tandfonline.com/doi/abs/10.1080/01621459.1979.10481634>
- Dunlop, P. and Smith, S. (2003). Estimating key characteristics of the concrete delivery and placement process using linear regression analysis. *Civil Engineering and Environmental Systems*, 20(4), 273-290. <https://doi.org/10.1080/1028660031000091599>
- Elhag, T.M.S. and Boussabaine, A. H. (2002). Tender price estimation using artificial neural

- networks II: modelling. *Journal of Financial Management Property and Construction*, 7(1), 49–64. Retrieved from: <http://discovery.ucl.ac.uk/1319335/>
- Elhag, T.M.S. and Wang, Y. M. (2007). Risk assessment for bridge maintenance projects : neural networks versus regression techniques. *Journal of Computing in Civil Engineering ASCE*, 21(6), 402–409. [https://doi.org/10.1061/\(ASCE\)0887-3801\(2007\)21:6\(402\)](https://doi.org/10.1061/(ASCE)0887-3801(2007)21:6(402))
- Elkordy, M.F. and Chang K. C. (1993). Neural network trained by analytically simulated damage states. *Journal of Computing in Civil Engineering ASCE*, 7(2), 130–145. [https://doi.org/10.1061/\(ASCE\)0887-3801\(1993\)7:2\(130\)](https://doi.org/10.1061/(ASCE)0887-3801(1993)7:2(130))
- Gagarin, N., Flood, I. and Albrecht, P. (1994). Computing truck attributes with artificial neural networks. *Journal of Computing in Civil Engineering ASCE*, 8(2), 179–200. [https://doi.org/10.1061/\(ASCE\)0887-3801\(1994\)8:2\(179\)](https://doi.org/10.1061/(ASCE)0887-3801(1994)8:2(179))
- Kerh, T., Huang, C. and Gunaratnam, D. (2011). Neural network approach for analyzing seismic data to identify potentially hazardous bridges. *Mathematical Problems in Engineering*, 1-15. <http://dx.doi.org/10.1155/2011/464353>
- King, S. (1999). Neural networks vs. multiple linear regression. *Proceedings 12th Central Hardwood Forest Conference*, 159–166. Retrieved from: <https://www.ncrs.fs.fed.us/pubs/ch/ch12/CHvolume12page159.pdf>
- Kumar, U.A. (2005). Comparison of neural networks and regression analysis: a new insight. *Expert Systems with Applications*, 29(2), 424–430. <https://doi.org/10.1016/j.eswa.2005.04.034>
- Kutner, M.H., Nachtsheim, C.J., Neter, J. and Li, W. (2004). *Applied linear statistical models* (3th ed.). New York, NY: McGraw-Hill
- Li, J., Dackermann, U., Xu, Y.L. and Samali, B. (2011). Damage identification in civil engineering

- structures utilizing PCA-compressed residual frequency response functions and neural network ensembles. *Structural Control and Health Monitoring*, 18(2), 207–226. DOI: 10.1002/stc.369
- Li, Z.B., Shi, Z. Q. and Ososanya, E.T. (1996). Evaluation of bridge conditions using artificial neural networks. *Proceedings, IEEE Southeastcon' 96, Bringing Together Education, Science, and Technology*, 366–369. DOI: 10.1109/SECON.1996.510091
- Masri, S.F., Chassiakos, A.G. and Caughey, T.K. (1993). Identification of nonlinear dynamic systems using neural networks. *ASME Journal of Applied Mechanics*, 60(1), 123–133. doi: 10.1115/1.2900734
- Masri, S.F., Nakamura, M., Chassiakos, A.G. and Caughey, T.K. (1996). Neural network approach to detection of changes in structural parameters. *Journal of Engineering Mechanics ASCE*, 122(4), 350–360. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1996\)122:4\(350\)](https://doi.org/10.1061/(ASCE)0733-9399(1996)122:4(350))
- Mata, J. (2011). Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Engineering Structures*, 33(3), 903–910. <https://doi.org/10.1016/j.engstruct.2010.12.011>
- Nguyen, N. and Cripps, A. (2001). Predicting housing value: a comparison of multiple regression analysis and artificial neural networks. *Journal of Real Estate Research*. 22(3), 313–336. <http://www.aresjournals.org/doi/abs/10.5555/rees.22.3.r47443706m17v575?code=ares-site>
- Ok, S., Son, W. and Lim, Y. M. (2012). A study of the use of artificial neural networks to estimate dynamic displacements due to dynamic loads in bridges. *Journal of Physics: Conference Series*, 382(1). <http://iopscience.iop.org/article/10.1088/1742-6596/382/1/012032/meta>
- Reich, Y. (1997). Machine learning techniques for civil engineering problems. *Computer-Aided Civil and Infrastructure Engineering*, 12(4), 295–310. doi: 10.1111/0885-9507.00065

- Suryanita, R. and Adnan, A. (2013). Application of neural networks in bridge health prediction based on acceleration and displacement data domain. Proceedings of the International MultiConference of Engineers and Computer Scientists. 2202(1), 42-47. Retrieved from: http://www.iaeng.org/publication/IMECS2013/IMECS2013_pp42-47.pdf
- Stevens, J.P. (1984). Outliers and influential data points in regression analysis. Psychological Bulletin, 95(2), 334–344. <http://dx.doi.org/10.1037/0033-2909.95.2.334>
- Sutton, C.D. (2005). Classification and regression trees, bagging, and boosting. Handbook of Statistics, 24(4), 303–329. [https://doi.org/10.1016/S0169-7161\(04\)24011-1](https://doi.org/10.1016/S0169-7161(04)24011-1)
- Tokdemir, O.B., Ayvalik, C. and Mohammadi, J. (2000). Prediction of highway bridge performance by artificial neural networks and genetic algorithms. Proceedings, 17th International Symposium on Automation and Robotics in Construction (ISARC), 1091-1098. Retrieved from: <https://pdfs.semanticscholar.org/3340/4f4190d7d1b4a9edb976b7b6d265627ed533.pdf>
- Wang, Y.M. and Elhag, T.M.S. (2007). A Comparison of neural network, evidential reasoning and multiple regression analysis in modelling bridge risks. Expert Systems with Application, 32(2), 336–348. <https://doi.org/10.1016/j.eswa.2005.11.029>
- Williams, T.P. (1994). Predicting changes in construction indexes using neural networks. Journal of Construction Engineering and Management, 120(2), 306–320. [https://doi.org/10.1061/\(ASCE\)0733-9364\(1994\)120:2\(306\)](https://doi.org/10.1061/(ASCE)0733-9364(1994)120:2(306))
- Wu, X., Ghaboussi, J. and Garrett Jr, J.H. (1992). Use of neural networks in detection of structural damage. Computers and Structures, 42(4), 649–659. [https://doi.org/10.1016/0045-7949\(92\)90132-J](https://doi.org/10.1016/0045-7949(92)90132-J)
- Youd, T.L., Hansen, C.M. and Bartlett, S.F. (2002). Revised multilinear regression equations for

prediction of lateral spread displacement. *Journal of Geotechnical and Geoenvironmental Engineering ASCE*, 128(12), 1007–1017. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2002\)128:12\(1007\)](https://doi.org/10.1061/(ASCE)1090-0241(2002)128:12(1007))

CHAPTER 5

- Abdenmour, M., Reggio, S., Le Naour, G., Liu, Y., Poitou, C., Aron-Wisnewsky, J., Charlotte, F., Bouillot, J-L., Torcivia, A., Sasso, M., Miette, V., Zucker J-D., Bedossa, P., Tordjman, J. and Clement, K. (2014). Association of adipose tissue and liver fibrosis with tissue stiffness in morbid obesity: Links with diabetes and BMI loss after gastric bypass. *Journal of Clinical Endocrinology and Metabolism*, 99(3), 898–907. <https://doi.org/10.1210/jc.2013-3253>
- Calinski, T. and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communication in Statistics* 3, 1-27.
- Curman, A.S.N., Andresen, M.A. and Brantingham, P. J. (2015). Crime and place: a longitudinal examination of street segment patterns in Vancouver, BC. *Journal of Quantitative Criminology*, 31(1), 127–147. <https://doi.org/10.1007/s10940-014-9228-3>
- Erman, J. and Arlitt, M. (2006). Traffic classification using clustering algorithms. *SIGCOMM Workshop*, 2006, 281–286. <https://doi.org/http://doi.acm.org/10.1145/1162678.1162679>
- Fiorillo, G. and Ghosn, M. (2014). Procedure for statistical categorization of overweight vehicles in a WIM database. *Journal of Transportation Engineering ASCE*, 140(5). [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000655](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000655)
- Florida Department of Transportation (FDOT) (2000). " Influence of heavy trucks on highway bridges". Summary of Final Report, BC-379.

- Gagarin, N., Flood, I. and Albrecht, P. (1994) Computing truck attributes with artificial neural networks. *Journal of Computing in Civil Engineering ASCE*, 8(2), 179–200.
- Genolini, C., Alacoque, X., Sentenac, M. and Arnaud, C. (2015). kml and kml3d: R packages to cluster longitudinal data. *Journal of Statistical Software*, 65(4), 1–34. <https://doi.org/10.1359/JBMR.0301229>
- Genolini, C. and Falissard, B. (2011). Kml: A package to cluster longitudinal data. *Computer Methods and Programs in Biomedicine*, 104(3), 112–121. <https://doi.org/10.1016/j.cmpb.2011.05.008>
- Herba, C.M., Tremblay, R.E., Boivin, M., Liu, X., Mongeau, C., Séguin, J.R., Côté, S.M. (2013). Maternal Depressive Symptoms and Children’s Emotional Problems. Can Early Child Care Help Children of Depressed Mothers? *JAMA Psychiatry*, 70(8), 830-838. [doi:10.1001/jamapsychiatry.2013.1361](https://doi.org/10.1001/jamapsychiatry.2013.1361)
- Hurault-Delarue, C., Chouquet, C., Savy, N., Lacroix, I., Beau, A-B., Montastruc, J-L. and Damase-Michel, C. (2016). How to take into account exposure to drugs over time in pharmacoepidemiology studies of pregnant women? *Pharmacoepidemiology and Drug Safety*, 25, 770–777
- Laurin, J.C., Geoffroy, M-C., Boivin, M., Japel, C., Raynault, M-F., Tremblay, R.E., and Cote, S.M. (2015). Child care services, socioeconomic inequalities, and academic performance. *pediatrics*, 136(6), 1112–1124. <https://doi.org/10.1542/peds.2015-0419>
- MacKelprang, R.D., Baeten, J.M., Donnell, D., Celum, C., Farquhar, C., De Bruyn, G., Essex, M., McElrath M.J., Nakku-Joloba, E. and Lingappa, J.R. (2012). Quantifying ongoing HIV-1 exposure in HIV-1-serodiscordant couples to identify individuals with potential host

- resistance to HIV-1. *Journal of Infectious Diseases*, 206(8), 1299–1308.
<https://doi.org/10.1093/infdis/jis480>
- Owusu-Ababio, S. and Schmitt, R. (2014). Evaluation of impacts of allowing heavier log loads in northern Wisconsin during spring thaw. Technical Report, 2011-2013.
- Pingault, J-B., Côté, S.M., Vitaro, F., Falissard, B., Genolini, C. and Tremblay, R.E. (2013). The developmental course of childhood inattention symptoms uniquely predicts educational attainment: A 16-year longitudinal study. *Psychiatry Research*, 219(3), 707–709.
<https://doi.org/10.1016/j.psychres.2014.06.022>
- R Package (2016). User manual: kml.
- Schels, M., Scherer, S., Glodek, M., Kestler, H.A., Palm, G. and Schwenker, F. (2013). On the discovery of events in EEG data utilizing information fusion. *Computational Statistics*, 28(1), 5–18. <https://doi.org/10.1007/s00180-011-0292-y>
- Shim, Y., Chung, J. and Choi, I. (2005). A comparison study of cluster validity indices using nonhierarchical clustering algorithm. *Proceedings of CIMCA-IAWTIC'05-Volume 01*, IEEE Computer Society, Washington, DC, 199-204.
- Shin, P., Jasso, H., Tilak, S., Cotofana, N., Fountain, T., Yan, L., Faser, M. and Elgamal, A. (2007). Automatic vehicle type classification using strain gauge sensors. *Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PerComW'07)*, 425–428. <https://doi.org/10.1109/PERCOMW.2007.25>
- Sun, C., Ritchie, S.G. and Oh, S. (2003). Inductive classifying artificial network for vehicle type categorization. *Computer-Aided Civil and Infrastructure Engineering ASCE*, 18(3), 161–172.

<https://doi.org/10.1111/1467-8667.00307>

- Tanaka, H., Nakatani, E., Fukutomi, Y., Sekiya, K., Kaneda, H., Iikura, M., Yoshida, M., Takahashi, K., Tomii, K., Nishikawa, M., Kaneko, N., Sugino, Y., Shinkai, M., Ueda, T., Tanikawa, Y., Shirai, T., Hirabayashi, M., Aoki, T., Kato, T., Iizuka, K., Fujii, M. and Taniguchi, M. (2017). Identification of patterns of factors preceding severe or life-threatening asthma exacerbations in a nationwide study. *European Journal of Allergy and Clinical Immunology*, 1–9. <https://doi.org/10.1111/all.13374>
- Walton, E., Pingault, J-B., Cecil, C.A.M, Gaunt, T.R., Relton, C.L., Mill, J. and Barker, E.D. (2017). Epigenetic profiling of ADHD symptoms trajectories: a prospective, methylome-wide study. *Molecular Psychiatry*, 22, 250–256
- Wheeler, A.P., Worden, R.E. and McLean, S.J. (2016). Replicating group-based trajectory models of crime at micro-places in Albany, NY. *Journal of Quantitative Criminology*, 32(4), 589– 612. <https://doi.org/10.1007/s10940-015-9268-3>
- Yan, L., Fraser, M., Elgamal, A., Fountain, T. and Oliver, K. (2008). Neural networks and principal components analysis for strain-based vehicle classification. *Journal of Computing in Civil Engineering ASCE*, 22(2), 123–132. [https://doi.org/10.1061/\(ASCE\)0887-3801\(2008\)22:2\(123\)](https://doi.org/10.1061/(ASCE)0887-3801(2008)22:2(123))