

Characterization of Stable-Health Older Drivers  
Using Low-Speed Driving Maneuvers From In-  
Vehicle Sensor Data

by

**Nathanael C. Fung, B.Eng.**

A thesis submitted to the Faculty of Graduate and Postdoctoral  
Affairs in partial fulfillment of the requirements for the degree of

**Master of Applied Science**  
in

**Biomedical Engineering**

Carleton University  
Ottawa, Ontario

© 2018, Nathanael C. Fung

## **Abstract**

The Candrive study aims to improve the current practices of screening elderly drivers in Canada by identifying predictors of motor vehicle collisions from monitoring their daily driving behaviours using in-vehicle sensors. The thesis objective was to characterize the baseline behaviour of stable-health older drivers by proposing parameters of interest for detecting changes in behaviour and methods to differentiate drivers using their maneuvers. The in-vehicle sensor data from 12 stable-health drivers were processed, and a turn-identification algorithm with 97.7% accuracy was created for extracting four maneuvers: accelerating from stop, decelerating to stop, right turns, and left turns on 40 to 60 km/h roadways. Most of the drivers exhibited relatively steady month-to-month acceleration behaviours and lower accelerations in adverse driving conditions, which represented their typical driving behaviours. Drivers can be differentiated by the driving patterns from their maneuvers using a multi-expert classifier, which may be applicable for detecting changes in driving behaviour.

## **Acknowledgements**

I would like to first thank my advisors, Dr. Adrian D.C. Chan and Dr. Rafik A. Goubran of the Department of Systems and Computer Engineering, for providing guidance, advice, and encouragement during my research and for providing the opportunity to expand my knowledge and explore my ideas. I also greatly appreciate the support and guidance from my medical advisor, Dr. Frank Knoefel of the Bruyère Research Institute.

I would like to give special thanks to Dr. Bruce Wallace for helping me get started on my research and taking the time to answer my many questions about the Candrive research study.

I would like to thank the Candrive personnel including Lynn Macleay of OHRI for providing access to the Candrive's in-vehicle sensor data. Without their support, this research could not have been conducted.

I would like to thank my colleagues in the graduate program for the fun and intriguing conversations. I greatly appreciate my friends for their constant encouragement and continual friendship.

I would like to express my sincere gratitude to my parents for their unconditional moral support. To my brothers, Barnabas and Solomon, I thank you for your kindness and inspiring me to continually improve myself.

Yours Sincerely,

Nathanael C. Fung

## Statement of Originality

This thesis presents the work of the author, under the supervision of Dr. Adrian D.C. Chan (Carleton University), Dr. Rafik A. Goubran (Carleton University), and Dr. Frank Knoefel (Bruyère Research Institute), completed at Carleton University for the degree of Master of Applied Science in Biomedical Engineering. The work from Chapter 3, 4, 5, and 6 was built on previous research by Dr. Bruce Wallace (Carleton University), and the author's work was an extension designed to focus on the analysis of vehicle maneuvers. The development of the in-vehicle sensor processing pipeline, which focused on preparing trips; the testing of the turn-identification algorithms; the definitions and analysis of the four low-speed maneuvers including environmental factors; and the multi-expert, multiclass classification using only vehicle maneuvers were performed by the author. The work in the thesis was reviewed by Dr. Adrian D.C. Chan, Dr. Rafik A. Goubran, and Dr. Frank Knoefel, and a portion of this thesis was presented in a conference paper as summarized below:

**Nathanael C. Fung**, Bruce Wallace, Adrian D.C. Chan, Rafik A. Goubran, Michelle M. Porter, Shawn Marshall, and Frank Knoefel, "Driver Identification Using Vehicle Acceleration and Deceleration Events from Naturalistic Driving of Older Drivers," accepted in *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, May 2017.

The conference paper described the multiclass classification approach for differentiating between drivers for a group of stable-health older drivers by using a multi-expert

classifier trained on the acceleration and deceleration events previously proposed by Dr. Bruce Wallace. The work from this conference paper was extended by the author to include more specific maneuvers and was presented in Chapter 6. The manuscript was written by the author, under the guidance of Dr. Adrian D.C. Chan, Dr. Rafik A. Goubran, Dr. Frank Knoefel, Dr. Bruce Wallace, and Dr. Michelle M. Porter (University of Manitoba). Dr. Shawn Marshall (OHRI) provided guidance on the initial direction of the research topic. The manuscript was revised following the feedback from the co-authors and was prepared for publication, which was accepted in 2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA) following the peer-review process during February 2017.

The in-vehicle sensor data from the participants is the property of the Candrive research study and Carleton University. The author received written permission and ethics approval to perform research on the group of participants in this thesis.

## Table of Contents

<b>Abstract.....</b>	<b>ii</b>
<b>Acknowledgements .....</b>	<b>iii</b>
<b>Statement of Originality .....</b>	<b>iv</b>
<b>Table of Contents .....</b>	<b>vi</b>
<b>List of Tables .....</b>	<b>ix</b>
<b>List of Figures.....</b>	<b>xi</b>
<b>List of Appendices.....</b>	<b>xvi</b>
<b>Glossary .....</b>	<b>xvii</b>
<b>List of Abbreviations .....</b>	<b>xviii</b>
1 Chapter: Introduction.....	1
1.1 Background.....	1
1.1.1 Older Drivers .....	1
1.1.2 Candrive Research Study.....	2
1.1.3 Intelligent Vehicle Systems and Vehicular Safety.....	3
1.2 Organization of Research Topics.....	3
1.3 Research Objectives.....	9
1.4 Research Contributions.....	9
1.5 Significance of the Study.....	10
1.6 Organization of Chapters .....	11
2 Chapter: Review of Literature .....	13
2.1 Introduction.....	13
2.2 On-the-Road Testing.....	13
2.3 Driving Simulators.....	14
2.4 Questionnaires and Surveys.....	15
2.5 In-Office Medical Assessments .....	16
2.6 In-Vehicle Sensors .....	17
2.6.1 Naturalistic Driving Studies.....	18
2.7 Identifying Patterns in the Driving Data.....	20
2.7.1 Driver Identification.....	21
2.7.2 Motor Vehicle Operations.....	22
2.8 Turn Identification .....	23
2.9 Driving Conditions.....	24
2.10 Candrive Research Studies .....	24
2.10.1 Differentiating Between Older Drivers.....	27
2.11 Summary.....	33
3 Chapter: Methodology .....	35
3.1 Introduction.....	35

3.2	Candrive Research Design.....	35
3.2.1	Candrive Instrumentation.....	36
3.3	Participants.....	38
3.4	Raw Data Processing .....	39
3.5	Data Processing.....	40
3.5.1	Historical Weather Data.....	47
3.5.2	Sunset and Sunrise Data.....	47
3.6	Derived Variables .....	47
3.7	Map Projection.....	49
3.8	Summary.....	49
4	Chapter: Identifying Turns in the GPS Data.....	51
4.1	Introduction.....	51
4.2	Methodology.....	52
4.2.1	Definition of a Turn .....	52
4.2.2	Training and Testing.....	55
4.2.3	Turn-Identification Algorithms.....	55
4.2.3.1	Algorithm #1: Threshold Heading .....	56
4.2.3.2	Algorithm #2: GPS Trace Reduction.....	57
4.2.3.3	Algorithm #3: Radius of Curvature .....	58
4.2.4	Performance Evaluation.....	59
4.3	Results.....	61
4.3.1	Algorithm #1: Threshold Heading.....	61
4.3.2	Algorithm #2: GPS Trace Reduction.....	62
4.3.3	Algorithm #3: Radius of Curvature .....	63
4.4	Discussion.....	64
4.4.1	Misclassification Errors .....	65
4.5	Conclusion .....	68
5	Chapter: Similarities and Differences in Maneuvering Behaviour.....	69
5.1	Introduction.....	69
5.2	Methodology.....	71
5.2.1	Driving Maneuvers .....	71
5.2.2	Definition of the Four Driving Maneuvers .....	72
5.2.3	Derived Variables .....	76
5.2.4	Categories of Driving Events.....	78
5.2.5	Definition of the Driving Conditions.....	81
5.2.6	Analysis of the Events .....	81
5.3	Results and Discussion .....	83
5.3.1	Acceleration and Deceleration Events .....	83
5.3.2	Right Turns Events .....	90
5.3.3	Left Turns Events.....	95
5.4	Summary.....	100
6	Chapter: Driver Identification Using Vehicle Maneuvers .....	102
6.1	Introduction.....	102
6.2	Methodology.....	104
6.2.1	Training and Testing Scheme .....	104
6.2.2	Feature Selection.....	108

6.2.3	Performance Evaluation.....	111
6.3	Results.....	112
6.3.1	Data Visualization.....	112
6.3.2	Event Classifiers .....	114
6.3.3	Expert Trip Classifiers .....	115
6.3.4	Multi-Expert Trip Classifier .....	116
6.4	Discussion .....	117
6.5	Conclusion .....	122
7	Chapter: Summary .....	124
7.1	Overview.....	124
7.2	Summary of Analysis.....	124
7.3	Challenges and Limitations.....	125
7.4	Future Work .....	126
7.4.1	Characterization of Baseline Year .....	126
7.4.2	Analysis of Subsequent Years .....	127
7.5	Conclusion .....	127
	<b>Appendices.....</b>	<b>128</b>
	<b>References.....</b>	<b>149</b>

## List of Tables

Table 3-1. Summary of first year data from the 12 triple-stable older drivers who resided in Ottawa and were the sole driver of their vehicle .....	39
Table 3-2. The data variables that were selected from the raw sensor data for the analysis of driving behaviour.....	41
Table 3-3. Derived variables for each trip from the processed sensor data.....	48
Table 4-1. The criteria for a turn in the GPS trace.....	53
Table 4-2. Description of the six test cases for the turn-identification performance evaluation.....	55
Table 4-3. Comparison of the three turn-identification algorithms using the optimized parameters that yielded the largest F-score from training.....	64
Table 5-1. Definitions of the four driving maneuvers .....	75
Table 5-2. Summary of the categories of driving events .....	79
Table 5-3. The criteria of the favourable and adverse driving conditions for each environmental factor .....	81
Table 5-4. Summary of the mean $\bar{a}_t$ for ACC events per driver over the entire baseline year including the group averages.....	84
Table 5-5. Summary of the mean $\bar{a}_t$ for DEC events per driver over the entire baseline year including the group averages.....	85
Table 5-6. Summary of the comparisons between the mean $\bar{a}_t$ for ACC and DEC during favourable and adverse conditions per driver. From the group difference between the means ( $\Delta = \mu_{adverse} - \mu_{favourable}$ ), a negative value for ACC events indicated that the mean acceleration was lower during adverse condition, whereas a positive value for DEC events indicated a lower deceleration during adverse conditions.....	90
Table 5-7. Summary of the mean $\bar{a}_N$ for acute RT events per driver over the entire baseline year including the group averages. ....	91
Table 5-8. Summary of the mean $\bar{a}_N$ for obtuse RT events per driver over the entire baseline year including the group averages. ....	92
Table 5-9. Summary of the comparisons between the mean $\bar{a}_N$ for acute and obtuse RT during favourable and adverse conditions per driver. From the group difference between the means ( $\Delta = \mu_{adverse} - \mu_{favourable}$ ), a negative value indicated that the mean acceleration was lower during adverse condition.....	95
Table 5-10. Summary of the mean $\bar{a}_N$ for acute LT events per driver over the entire baseline year including the group averages. ....	96
Table 5-11. Summary of the mean $\bar{a}_N$ for obtuse LT events per driver over the entire baseline year including the group average. ....	97
Table 5-12. Summary of the comparisons between the mean $\bar{a}_N$ for acute and obtuse LT during favourable and adverse conditions per driver. From the group difference between the means ( $\Delta = \mu_{adverse} - \mu_{favourable}$ ), a negative value indicated that the mean acceleration was lower during adverse condition.....	100
Table 6-1. List of the features extracted from the driving events and the discrete-time signals $s[n]$ .....	109
Table 6-2. (Left) An example of the class separation of two high mileage drivers, #2 (red) and #10 (blue), using the first two principal components to visualize the events. The green x marker represents the mean of Driver 2 and the yellow x	

marker represents the mean of Driver 10. (Right) The group of 12 drivers showing 95% confidence ellipse from a Gaussian distribution of the events using the first two principal components from PCA..... 113

Table 6-3. Top 10 features from feature selection prior to training the LR event classifier ..... 114

Table 6-4. Summary of the p-values from significance testing between classification scores of LR and RF for each type of maneuver..... 118

Table 6-5. Summary of the p-values from significance testing between the classification scores of the expert trip classifiers and ME trip classifier for each type of maneuver..... 119

## List of Figures

Figure 1-1. Hierarchical organizational chart for Candrive showing the five main categories of research topics. Each level is supported by the results produced by the lower levels. The shaded boxes represent research topics included in this thesis research and the unshaded boxes represent research topics outside the scope of this thesis research. ....	8
Figure 2-1. The proposed two-phase relationship between max acceleration and change in speed (top group) and mean acceleration and change in speed (bottom group) for 4 example drivers in their baseline year. The Gaussian mean represented the mean of the best fit Gaussian distribution that was fitted on the acceleration data points associated with each speed bin. The piecewise linear regression was shown to be a good fit for representing the general trend of the underlying acceleration events from the in-vehicle sensor data [68]. ....	30
Figure 2-2. The proposed two-phase relationship between mean deceleration and change in speed (top group) and max deceleration and change in speed (bottom group) for 4 example drivers in their baseline year. The Gaussian mean represented the mean of the best fit Gaussian distribution that was fitted on the acceleration data points associated with each speed bin. The piecewise linear regression was shown to be a good fit for representing the general trend of the underlying deceleration events from the in-vehicle sensor data [9]. ....	32
Figure 3-1. The data logging device, showing the five components, used by Candrive research study for data collection: (a) GPS antenna for recording longitude and latitude coordinates; (b) the central unit, which is similar to the size of the palm of the hand, for recording and storing the sensor data; (c) the connector from the central unit to the vehicle engine diagnostic sensors via OBDII port; (d) the RDIF antenna for detecting the nearby key fob to track the identity of the driver; and (e) the key fob that is assigned to the driver and attached to the car keys [70]. ....	37
Figure 3-2. Processing pipeline for the data processing steps of the sensor data .....	42
Figure 3-3. (a) Invalid GPS coordinates were detected in the latitude channel. (b) The same latitude data after cleaning had no extreme values. ....	44
Figure 3-4. An example of one entire trip (top) showing the GPS drift, which commonly occurred when the vehicle was stationary, marked inside the red circle. The magnified view of the GPS drift (bottom left) showing the impossible vehicle movements corrected into a straight line (bottom right), which more accurately approximated the actual route of the vehicle. ....	46
Figure 4-1. (a) Example of a turn showing the recorded GPS coordinates and (b) the same turn projected on to a map with a green marker to indicate the center of the turn. The turning angle in the figures are not to scale. ....	53
Figure 4-2. The entire route for trip #5 showing a variety of turns and paths taken by the vehicle. ....	55
Figure 4-3. The performance of algorithm #1 on the 5 training trips showing the F-score for a parameter sweep of the threshold angle $\theta$ between $0^\circ$ to $90^\circ$ in steps of $1^\circ$ . The threshold angle $\theta$ that maximized the F-score was $17^\circ$ , as shown by the square marker. ....	61

Figure 4-4. The performance of algorithm #2 on the 5 training trips showing the F-score for a parameter sweep of $\varphi$ between $0^\circ$ to $90^\circ$ and 1 m to 30 m. The angle $\varphi$ and tolerance $\tau$ that yielded the best F-score was $40^\circ$ and 6 m respectively, as shown by square marker.....	62
Figure 4-5. The performance of algorithm #3 on the 5 training trips showing the highest F-score as indicated by the square marker (a) for a parameter sweep of $\rho_u$ between 0 to 50 m to determine the upper threshold, which was 16 m, and (b) the F-score for the lower threshold $\rho_l$ was 2 m.....	63
Figure 4-6. An example of FP in the GPS trace by algorithm #1 as indicated by the red marker on a map display with the trace of the route.....	65
Figure 4-7. An example of a FP for algorithm #2 which was caused by a short curve in the road (red star) that was emphasized when reducing the GPS trace. (a) The map display shows the turns from the gold standard (green marker) and the identified turns by the algorithm (red marker). (b) The coordinates of the path. (c) The reduced GPS data of the same route.....	66
Figure 4-8. An example of a FN for algorithm #2 (red star). The turn (red star) was more gradual than the other turn (red marker), and the reduction of the GPS trace did not emphasize the corner. Since the two turns were in close vicinity, the turns were merged into one larger and sharper turn. (a) The map display shows the turns from the gold standard (green marker) and the identified turns by the algorithm (red marker). (b) The coordinates of the path. (c) The reduced GPS data of the same route. ....	67
Figure 4-9. An example of a FP for algorithm #3 (red star), where a gentle curve was mistaken as a turn.....	67
Figure 5-1. An example of a trip taken by an older driver showing the occurrence of common driving maneuvers during the trip, namely acceleration, deceleration, and turns. The acceleration and deceleration events were defined in [9] had changes in speed of $\geq 4$ km/h. The vehicle was at steady speeds (blue); the vehicle was accelerating (green); the vehicle was decelerating (red); and the vehicle was turning (magenta). (a) It was evident that there were two expressway ramps (loops), which indicate that the long, straight path in the middle of the trip was the expressway and the other sections were lower speed roadways, which could be an example of the driver's trip between home and place of employment. (b) The magnified view of the lower speed roadways, which was characterized by more corners and bends than the expressway. The start and end destinations of this trip was excluded to preserve anonymity. ...	72
Figure 5-2. An example of the smaller subset of driving events shown in Figure 5-1b. The black X marked the coordinates when the vehicle was stopped. The green highlight indicated the portion of the trip when the ACC event occurred, the red highlight showed the DEC events, and magenta showed the RT events. Most ACC and DEC events occurred in pairs, but there were cases when the RT or LT occurred before or after the ACC and DEC events, which contributed to an unequal amount of ACC and DEC events. ....	74
Figure 5-3. (a) An example of ACC event and (b) RT event showing the discrete-time signal of the speed, acceleration, and jerk values with linear interpolation for visualization. ....	76

- Figure 5-4. Example of three different turns identified from the GPS data. The path of the vehicle was shown by the blue GPS trace and the red circle approximately fitted the arc of the turn. The yellow vector A and purple vector B extended from the center to the edge of the circle and represented the start and ends of the turn respectively. (a) An example of a sharp turn from performing a u-turn with  $r_T = 4$  m and  $\theta_T > 180^\circ$ . (b) A large, gradual turn with  $r_T = 92$  m and  $\theta_T = 53^\circ$ . (c) An example of an on-ramp from the expressway with  $r_T = 62$  m and  $\theta_T = 217^\circ$ . ..... 78
- Figure 5-5. The distribution of the total number of each category of events for all drivers. The red lines represented the median value, the boxes marked the width of the IQR, and the + symbols represented the drivers who had a much larger number of events. The events that occurred on the lower speed roadways were much more abundant than the other events. .... 80
- Figure 5-6. Comparison of the distributions of mean tangential acceleration ( $\bar{a}_t$ ) for ACC events (a) and DEC events (b) per driver across the entire baseline year. The red lines represented the median value and the boxes marked the width of the IQR. The outliers were removed by setting the lower limit to  $Q1 - 1.5(IQR)$  and upper limit to  $Q3 + 1.5(IQR)$ . By removing the outliers, the distribution was approximated by a normal distribution  $N(\mu, \sigma^2)$ . ..... 84
- Figure 5-7. Comparison of the typical monthly behaviours from ACC (a) and DEC (b) events for each driver during their baseline year including the yearlong linear trend. Each month was represented by the median of the mean tangential accelerations ( $\bar{a}_t$ ) across all the events during each month and the linear interpolation imputed the missing in-between months. .... 86
- Figure 5-8. (a) The relationship between the IQR of the median  $\bar{a}_t$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_t$  across the entire year,  $R = 0.897$  ( $p < .001$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_t$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_t$ ,  $R = 0.603$  ( $p = .038$ ), for ACC events. .... 87
- Figure 5-9. (a) The relationship between the IQR of the median  $\bar{a}_t$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_t$  across the entire year,  $R = -0.944$  ( $p < .001$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_t$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_t$ ,  $R = 0.654$  ( $p = .021$ ), for DEC events. .... 88
- Figure 5-10. (a) The relationship between the average of the median  $\bar{a}_t$  (i.e., typical monthly accelerations) across the entire year and the annual mileage,  $R = 0.017$  ( $p = .956$ ), and (b) the relationship between the IQR of the median  $\bar{a}_t$  across the entire year and the annual mileage,  $R = 0.230$  ( $p = .472$ ), for ACC events. .... 89
- Figure 5-11. (a) The relationship between the average of the median  $\bar{a}_t$  (i.e., typical monthly accelerations) across the entire year and the annual mileage,  $R = 0.267$  ( $p = .402$ ), and (b) the relationship between the IQR of the median  $\bar{a}_t$  across the entire year and the annual mileage,  $R = -0.341$  ( $p = .278$ ), for DEC events. .... 89

- Figure 5-12. Comparison of the distributions of mean net acceleration ( $\bar{a}_N$ ) for acute RT events (a) and obtuse RT events (b) per driver for the entire baseline year. The red lines represented the median value and the boxes marked the width of the IQR. The outliers were removed by setting the lower limit to  $Q1 - 1.5(IQR)$  and upper limit to  $Q3 + 1.5(IQR)$ . By removing the outliers, the distribution was approximated by a normal distribution  $N(\mu, \sigma^2)$ . ..... 91
- Figure 5-13. Comparison of the typical monthly behaviours from acute RT (a) and obtuse RT (b) events for each driver during their baseline year including the yearlong linear trend. Each month was represented by the median of the mean net accelerations ( $\bar{a}_N$ ) across all the events during each month and the linear interpolation imputed the missing in-between months. .... 92
- Figure 5-14. (a) The relationship between the IQR of the median  $\bar{a}_N$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_N$  across the entire year,  $R = 0.574$  ( $p = .051$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_N$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_N$ ,  $R = 0.524$  ( $p = .080$ ), for acute RT events..... 93
- Figure 5-15. (a) The relationship between the IQR of the median  $\bar{a}_N$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_N$  across the entire year,  $R = 0.491$  ( $p = .105$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_N$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_N$ ,  $R = 0.400$  ( $p = .198$ ), for obtuse RT events..... 94
- Figure 5-16. Comparison of the distributions of mean net acceleration ( $\bar{a}_N$ ) for acute LT events (a) and obtuse LT events (b) per driver for the entire baseline year. The red lines represented the median value and the boxes marked the width of the IQR. The outliers were removed by setting the lower limit to  $Q1 - 1.5(IQR)$  and upper limit to  $Q3 + 1.5(IQR)$ . By removing the outliers, the distribution was approximated by a normal distribution  $N(\mu, \sigma^2)$ . ..... 96
- Figure 5-17. Comparison of the typical monthly behaviours from acute LT (a) and obtuse LT (b) events for each driver during their baseline year including the yearlong linear trend. Each month was represented by the median of the mean net accelerations ( $\bar{a}_N$ ) across all the events during each month and the linear interpolation imputed the missing in-between months. .... 98
- Figure 5-18. (a) The relationship between the IQR of the median  $\bar{a}_N$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_N$  across the entire year,  $R = 0.561$  ( $p = .058$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_N$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_N$ ,  $R = 0.634$  ( $p = .027$ ), for acute LT events..... 99
- Figure 5-19.(a) The relationship between the IQR of the median  $\bar{a}_N$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_N$  across the entire year,  $R = 0.415$  ( $p = .180$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_N$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_N$ ,  $R = 0.226$  ( $p = .481$ ), for obtuse LT events..... 99

Figure 6-1. The training and testing scheme for the multiclass classification of drivers using maneuvering events. (1) The in-vehicle sensor data was prepared by organizing the dataset by the driver ID and segmented into individual trips. (2) The dataset of trips were split using 5-fold CV for the separate testing and training sets during the evaluation. (3) After extracting all maneuvering events from the trips, (4) the four independent expert classifiers were trained on each type of event, and then the expert classifiers were combined to create a multi-expert classifier to identify the drivers based on their maneuvering events. . 108

Figure 6-2. The average classification accuracies of the ME trip classifier for the group of 12 drivers and the improvement from removing trips with few events. .... 117

Figure 6-3. The average classification accuracies from rank classification of 12 drivers for the ME trip classifier. At rank 1, the average accuracies were the lowest having only one choice for the predicted driver, and at rank 12 the accuracies were 100% since all the drivers were predicted..... 117

Figure 6-4. The average classification accuracies for size 2 to 12 subsets of drivers showing that majority vote (MV) applied to the ACC classifiers consistently improved performance across all subsets of drivers, and the ME classifier consistently had higher performance. The LR and RF identified drivers based on individual driving events, whereas the MV-RF and ME-RF based on trips. .... 119

Figure 6-5. The average classification accuracies for size 2 to 12 subsets of drivers showing that majority vote (MV) applied to the DEC classifiers consistently improved performance across all subsets of drivers, and the ME classifier had higher performance for larger subsets of 6 or more drivers. The LR and RF identified drivers based on individual driving events, whereas the MV-RF and ME-RF based on trips. .... 120

Figure 6-6. The average classification accuracies for size 2 to 12 subsets of drivers showing that majority vote (MV) applied to the RT classifiers consistently improved performance across all subsets of drivers, and the ME classifier consistently had higher performance. The LR and RF identified drivers based on individual driving events, whereas the MV-RF and ME-RF based on trips. .... 120

Figure 6-7. The average classification accuracies for size 2 to 12 subsets of drivers showing that majority vote (MV) applied to the LT classifiers consistently improved performance across all subsets of drivers, and the ME classifier consistently had higher performance. The LR and RF identified drivers based on individual driving events, whereas the MV-RF and ME-RF based on trips. .... 121

## List of Appendices

<b>Appendix A : Data Validation .....</b>	<b>128</b>
A.1 Criteria of Triple-Stable Older Drivers.....	128
A.2 In-Vehicle Sensor Data.....	128
A.3 Weather Data .....	129
<b>Appendix B : Comparison of Favourable and Adverse Driving Conditions.....</b>	<b>132</b>
B.1 Acceleration Events .....	132
B.2 Deceleration Events .....	134
B.3 Acute Right Turn Events .....	136
B.4 Obtuse Right Turn Events.....	138
B.5 Acute Left Turn Events.....	140
B.6 Obtuse Left Turn Events.....	142
<b>Appendix C : Classification Performance .....</b>	<b>144</b>
C.1 Event Classifiers .....	144
C.2 Expert Trip Classifiers .....	145
C.3 Multi-Expert Trip Classifier .....	148

## Glossary

Term	Definition
Candrive	The Canadian driving research initiative for vehicular safety in the elderly study.
CAN-bus	Controller Area Network is a communications bus standard for vehicles.
Cornering	The action of turning the vehicle while driving.
Driving Behaviour	Any actions or decisions by the driver in response to a particular traffic scenario.
Driving Signature	The characteristics of the driving behaviour that uniquely represents the driver.
Event	An instance of a driving maneuver.
F-score	Weighted harmonic mean of the recall and precision evaluation metrics.
GPS Trace	The set of time-series GPS coordinates within a trip or a segment of a trip.
Stable-Health	Synonym of triple-stable.
Straightaways	A portion of the route that is straight.
Trip	Period between the automatic starting of the data logging when the engine turned on until the automatic stopping of the data logging when the engine turned off.
Triple-Stable	Clinical assessment for stability in cognitive, physical, and general health for older adults.
Turning point	A GPS coordinate that lies within a corner or bend in the path that indicates a turning maneuver.

## List of Abbreviations

Abbreviation	Definition
ACC	Acceleration Event
ANN	Artificial Neural Network
CV	Cross-Validation
DEC	Deceleration Event
DOP	Dilution of Precision
DOS	Driver Observation Schedule
EST	Eastern Standard Time
FN	False Negative
FP	False Positive
GPS	Global Positioning System
HMM	Hidden Markov Models
ICC	Intraclass Correlation Coefficient
IQR	Interquartile Range
LDA	Linear Discriminant Analysis Classifier
LR	Logistic Regression Classifier
LT	Left Turn Event
ME	Multi-Expert
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
MSB	Mean Squared Error Between-Group
MSW	Mean Squared Error Within-Group
MV	Majority-vote
MVC	Motor Vehicle Collisions
NDS	Naturalistic Driving Study
OBDII	On-Board Diagnostics 2
PC	Precision
PCA	Principal Component Analysis
Pchip	Piecewise Cubic Hermite Interpolating Polynomial
RC	Recall
RDP	Ramer-Douglas-Peucker Algorithm
RF	Random Forest Classifier
RFID	Radio-Frequency Identification
RT	Right Turn Event
SIMARD-MD	Screen for the Identification of Cognitively Impaired Medically At-Risk Drivers – A Modification of the DemTect
SVM	Support Vector Machine Classifier
TP	True Positive
UTC	Coordinated Universal Time
UTM	Universal Trans Mercator Projection
WGS84	World Geodetic System 1984

# **1 Chapter: Introduction**

## **1.1 Background**

### **1.1.1 Older Drivers**

The Canadian population is projected to have an unprecedented increase in the demographic of older adults (i.e., 65 years and older), which is expected to become 20% of the total Canadian population by 2024 [1]. Correspondingly, there is expected to be an increasing number of older adults driving. As older adults experience cognitive and physical decline in health, their driving ability can be negatively affected, which can result in increased risks of motor vehicle collisions (MVC) [2]. Early detection of cognitive or physical decline in older drivers can improve not only their safety, but also the safety of the general public.

The screening process for medically unfit older drivers is not standardized across Canada. Each province has its own legislation for assessing the driving ability of older adults. For example, older drivers aged 80 years and older are required to attend a renewal session for their driver's license every 2 years in Ontario (2017) [3], whereas medical reports from primary-care physicians are required every 2 years in British Columbia (2017) [4]. However, most physicians are obligated to report their patients who they believe to be medically unfit to drive. Problems may arise during the screening process because it is often difficult to pinpoint when exactly to report the older driver. False negative identification of medically unfit older drivers will put them and others at greater risk while driving, and false positive identification and preventing older adults from driving can have a negative impact on their quality of life, including increased social isolation and depression [5]. Most physicians use in-office cognitive and physical

assessments to determine medical fitness to drive, which often consist of relatively simple tasks. For example, one task is to draw a picture of a clock within a time limit, and another is to recall a pattern of numbers and walk for a specific distance. Although these tests were designed to assess cognitive and physical ability, they may not always be an accurate assessment of their actual driving ability [6][7]. As a result, situations may arise where the assessments indicate the patient is medically unfit to drive, but the patient is a perfectly safe driver. In another scenario, the patient may have spent some time practicing for the in-office medical assessment and obtain a passing score on the tests, but in reality has a low driving ability. On-the-road driving tests can be more accurate in assessing driving ability, but the stress from being observed by an examiner combined with the possibility of imminently losing the driver's license can result in an unusually poor performance. Thus, there is a need for more objective and evidence-based screening tools for identifying medically unfit older drivers.

### **1.1.2 Candrive Research Study**

The Candrive research study<sup>1</sup>, formally known as the *Canadian Driving Research Initiative for Vehicular Safety in the Elderly*, is a longitudinal cohort study in collaboration with several research institutions across Canada. Candrive's aim is to "...improve the health, safety, and quality-of-life of Canada's older drivers". The primary goal is to develop valid tools that assist in the screening of older drivers by accurately detecting "...who is at increased risk of motor vehicle collisions or who is unsafe to drive" [5]. In Candrive, a data logging device was installed in each participant's personal motor vehicle and it was used to record information about the vehicle during ordinary day-to-day operations wherein the participants were driving in their natural environment

---

<sup>1</sup>For more information on Candrive, please visit the website [www.candrive.ca](http://www.candrive.ca).

over a span of 7 years (i.e., 2009 to 2016). The driving data is a rich and detailed historical record of their natural driving behaviour. Candrive is a unique research opportunity to study the natural driving behaviour of older adults and the related effects of aging on driving ability; it could provide supporting evidence for the physicians to make better-informed decisions by reducing false positive and negative identification and inform future government legislation to create a standardized screening process across Canada.

### **1.1.3 Intelligent Vehicle Systems and Vehicular Safety**

Wireless sensor technology and vehicle telematics have enabled organizations, such as Candrive, to monitor their drivers and collect large quantities of data on their driving history. By analyzing the information contained in the driving data, organizations can make better informed decisions based on the real-world driving behaviour. Moreover, gathering detailed data of drivers' behaviours can be beneficial for developing new technologies for improving vehicular safety. Such emerging applications include: insurance telematics where lower premiums are awarded to safe drivers; detecting impersonation (i.e., fraud or insurance claims) or identifying secondary drivers of shared vehicles; detecting vehicle thefts; monitoring a fleet of drivers to improve productivity; designing safer and more personalized advanced driving assistant systems and intelligent autonomous vehicles for driving comfort; and coaching safer driving behaviours and providing alerts to reduce traffic accidents.

## **1.2 Organization of Research Topics**

A hierarchical block diagram in Figure 1-1 depicts the organization of the Candrive research study and the specific research topics included in this thesis. Each

level is supported by the lower levels, and each level is described in the corresponding numbered items below:

1. Candrive's primary research goal is to "create valid, easy-to-use screening tools to identify medically unfit and unsafe older drivers" [5]. Below is a list of stakeholders of this research:
  - a. **Older drivers** gain better physician-patient relationships from the reduction of false positive identification rate of medically unfit drivers and the prolonging of their driving license, which can help provide a high quality of life by maintaining their social independence. Most importantly, the reduction of MVC and traffic incidents can improve their safety.
  - b. **Healthcare providers** can use the screening tools to reduce false positives during the screening process and thus leading to potential benefits such as reduced healthcare costs and reduced time spent per patient. Only those who have increased risk may require further testing with time-consuming and expensive tests, which can improve resource management.
  - c. **The government** can use the research to help create solutions for the aging population such as a reliable and standardized screening process across Canada that will contribute to reduced healthcare costs for older adults.
  - d. **Public** safety will be improved by decreasing the number of MVC and traffic incidents on public roads involving medically unfit older drivers by lowering the false negative identification rate.

2. Several different methods have been used by researchers to study driving behaviour and assess medical fitness to drive including driving simulators that test driving skills in simulated traffic scenarios; questionnaires and surveys to obtain the driver's self-assessments; written tests and road tests with a certified examiner; in-office medical assessments on cognitive and physical function; and vehicle tracking to monitor the real-world driving behaviour in a natural setting.
3. Candrive's primary research method is vehicle tracking, which would be used for long-term monitoring of driving ability and detecting changes in driving behaviour that may indicate a decline in driving ability. The goal is to identify predictors of at-risk behaviour leading to traffic incidents (e.g., MVC) and effects of cognitive and physical impairments in their driving behaviour. These predictors can help create valid detection rules for medically unfit older drivers and performance metrics to measure driving ability, which is analogous to having a threshold for the blood alcohol concentration when testing for alcohol-impaired driving. Together with medical assessments, vehicle tracking can also help identify the cause of MVC and other traffic incidents such as whether it was due to physical or cognitive impairment.
4. To identify relevant parameters in their in-vehicle driving data for developing metrics to assess driving ability, Candrive collected data from hundreds of older drivers over 7 years for analysis. Since the in-vehicle data contains many measured variables based on the driver's input, an important step is to determine what dependant variables and patterns exists in their data and which are important for long-term monitoring. For a single-subject analysis approach, the detailed in-

vehicle data from each subject is studied both quantitatively and qualitatively to understand the individual driving behaviour. As each subject is their own control, the changes in the in-vehicle data that are observed when a traffic incident occurred or when their health condition changed could be the key to identifying relevant measured variables. For example, when a driver is assessed to be healthy in their baseline period and becomes unhealthy later on, the data may show a consistently slower and more erratic deceleration before traffic stops, which may indicate greater difficulty in braking safely. However, the changes in driving behaviour for individual drivers may involve confounding factors, which also influence the changes in behaviour. Therefore, another approach is group analysis to identify common driving behaviours among groups of older drivers with similar traits. For example, a group of older drivers with deteriorating physical health may all have slower vehicle acceleration when compared to healthy drivers, which may indicate difficulty in controlling the gas pedal.

5. Before comparing different subjects to identify predictors of a decline in driving ability, a fundamental step is to characterize the individual driving behaviours by identifying and quantifying the changes in the measured variables of the in-vehicle data to determine what are the expected driving behaviours during their baseline period. Candrive has segmented the longitudinal data into years by which the subjects were enrolled in the study and categorized their respective first years as their baseline period. Since driving behaviour is complex involving static and dynamic characteristics that are influenced by individual preferences as well as external factors (e.g., weather conditions) [8], quantifying individual driving

behaviour using the in-vehicle data from their baseline year can help identify unique behaviours, known as their *driving signature*, and common driving behaviours, which will inform future analyses on developing useful metrics for long-term monitoring of driving ability.

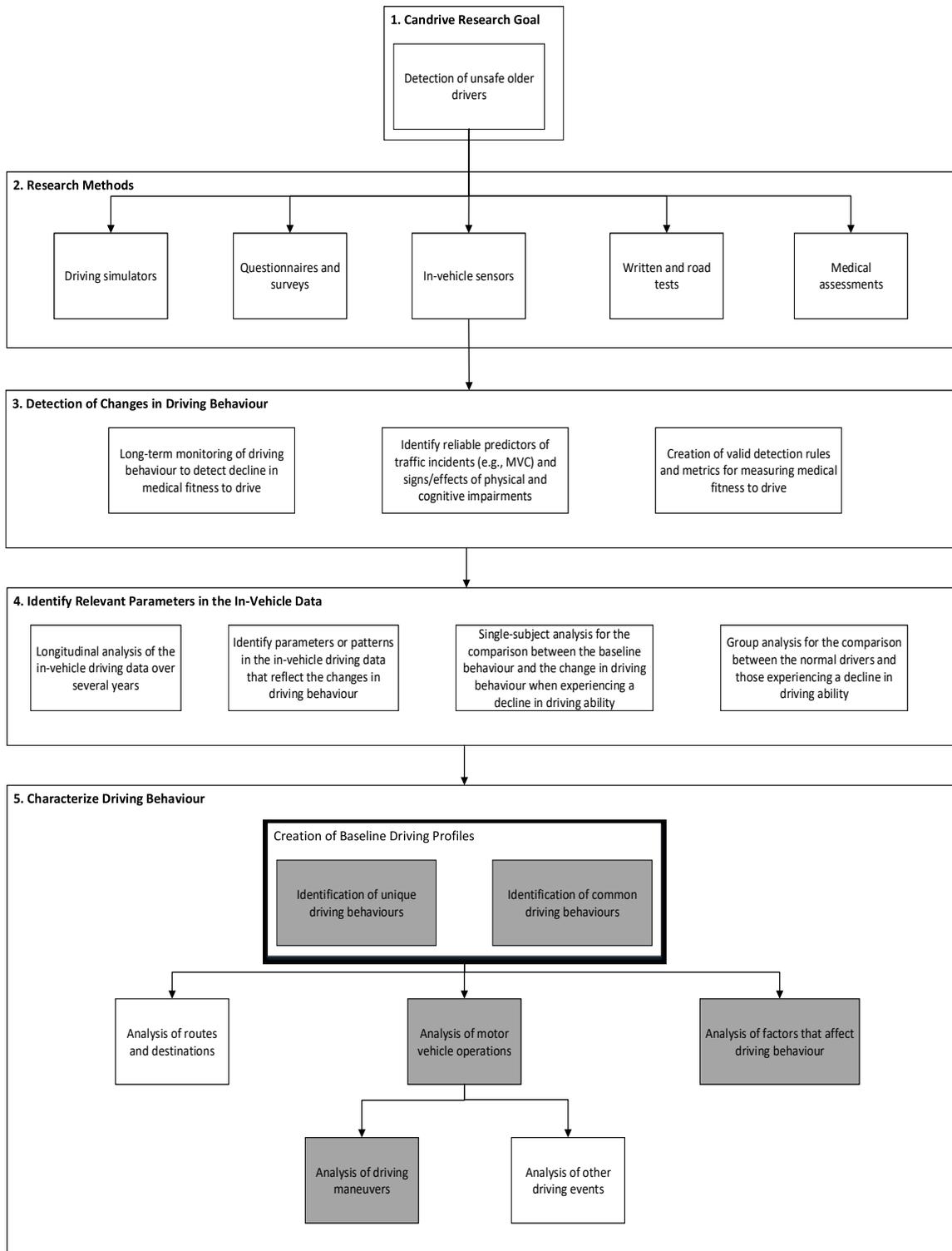


Figure 1-1. Hierarchical organizational chart for Candrive showing the five main categories of research topics. Each level is supported by the results produced by the lower levels. The shaded boxes represent research topics included in this thesis research and the unshaded boxes represent research topics outside the scope of this thesis research.

### 1.3 Research Objectives

As indicated in Figure 1-1, the topics of this research are categorized into level 5, which focuses on characterizing driving behaviour. The objective of this research was to characterize the driving behaviours of a group of stable-health older drivers in their baseline year of the Candrive study by identifying similarities and differences in their motor vehicle operations, as opposed to route choices and destinations, from low-speed driving maneuvers using their in-vehicle sensor data. In addition, another objective was to develop a method to differentiate between drivers using their unique driving patterns from low-speed driving maneuvers. This research is an extension of the previous work by Wallace *et al.* [9] in characterizing older drivers in their baseline year by exploring the unique and common behaviours in their acceleration and deceleration driving events.

### 1.4 Research Contributions

There are three main contributions from this thesis research as listed below:

- 1) Proposed three different algorithms for identifying when the driver was turning in the GPS data and evaluated the performance of these three algorithms to determine which algorithm had the best performance in identifying turns (Chapter 4).
- 2) Analyzed four common driving maneuvers (i.e., acceleration, deceleration, right turns, and left turns) that occur abundantly in low-speed roadways (i.e., 40 km/h to 60 km/h) for a group of stable-health older drivers by identifying similarities and differences in their maneuvering behaviours during their baseline year of the Candrive study. This included the analysis of the effects of four environmental factors (i.e., weather, season, traffic condition, and daylight) on their maneuvering

behaviour and proposing a measure for consistency of their maneuvering behaviour (Chapter 5).

- 3) Proposed a novel multi-expert classification approach for driver identification based on the patterns on the four different driving maneuvers and evaluated the classification performance (Chapter 6).

## **1.5 Significance of the Study**

The unique and common characteristics in their in-vehicle data from their driving maneuvers can help establish the baseline driving behaviour of stable-health older drivers and propose parameters of interest from the in-vehicle sensor data for detecting changes in driving behaviour. Instead of the analysis of routes, which describes “where they drove”, the analysis of driving maneuvers more effectively describes “how they drove”. As older drivers experience a decline in health, the associated changes in their driving patterns can contribute in identifying predictors of at-risk behaviours and ultimately detect a decline in driving ability (level 4 in Figure 1-1). Driver identification demonstrates the capabilities of pattern recognition techniques (i.e., classification methods) for differentiating between individual drivers by their unique driving patterns, which can act as a proxy for differentiating between the same driver during different time periods. The potential application for driver identification in future studies is differentiating between the same driver in different time periods when there is a known decline in health. The unique driving patterns associated with the decline in health can contribute to identifying important parameters in the in-vehicle sensor data that are related to an increased risk of MVC.

This research also provided methods of processing the raw in-vehicle sensor data,

extraction of specific driving maneuvers, and methods for distinguishing drivers, which have many applications including the development of a digital driving signature that uniquely identifies a driver. A driving signature could provide the following benefits:

- Improve insurance telematics programs by adjusting premiums based on individual driving patterns and less on simple detection rules such as detecting aggressive events.
- Differentiate drivers of shared vehicles such as older drivers who shared their vehicles during the Candrive study or for insurance programs for a family that share a vehicle.
- Truck and construction drivers in a telematics program could improve their efficiency based on performance metrics of their driving maneuvers.
- Personalization of intelligent and autonomous vehicles systems to fit the preferences of their customers. Perhaps when autonomous vehicles become mainstream and reliable, older drivers who become medically unfit to drive may be incentivized to use autonomous vehicles to lower risks of MVC. In this case, the vehicles could adopt the safe behaviours of the drivers to provide a seamless transition and high user-acceptance rate.

## **1.6 Organization of Chapters**

After the introduction (Chapter 1), a review of the recent state-of-the-art research and literature pertaining to the driving behaviour of older adults using vehicle tracking is presented. Following the literature review (Chapter 2), the methodology of this study is presented and describes the in-vehicle sensor data from the selected stable-health participants (Chapter 3). Next, the three algorithms for identifying turns in the GPS data

are presented and the best algorithm was selected and implemented for the following chapters (Chapter 4). After selecting the best algorithm, the four driving maneuvers are defined and analyzed to identify similarities and differences in behaviours among the stable-health older drivers, including the analysis of the effects of different driving conditions (Chapter 5). Once the four driving maneuvers have been analyzed, the next chapter (Chapter 6) describes and evaluates a multiclass, multi-expert classification method to identify the drivers based on their unique maneuvering patterns and provides a few methods for improving performance. Lastly, the final chapter (Chapter 7) summarizes the key findings from the previous chapters and provides suggestions for future work.

## **2 Chapter: Review of Literature**

### **2.1 Introduction**

There is a growing amount of research on the driving behaviour of older adults, especially because long-term driving studies are emerging and show promising results for providing real-world evidence of driving ability. In addition, there have been many studies on driving behaviour and driving signature because of the recent technological advances in the automotive industries for applications in autonomous and intelligent vehicle systems. Marshall *et al.* [2] suggested that the medical conditions and cognitive impairments are related to an increased risk of MVC, but drivers with these medical and cognitive conditions may still be fit to drive. There are many challenges in screening older drivers [10] because there is no standardized procedure for quick screening. Several recent studies [6][7] show that current methods of assessing medical fitness to drive (i.e., in-office assessments) may be inadequate. As such, long-term driving studies, also called Naturalistic Driving Studies (NDS), are an emerging research area that utilizes unobtrusive in-vehicle sensors to track the driver's actions and may provide solutions to many problems with current screening methods. This chapter presents a literature review on the current research on the driving behaviour of older drivers including current methods of assessing medical fitness to drive, vehicle tracking studies using vehicles sensor systems, driving patterns for driver identification, factors that affect driving, and the recent Candrive studies.

### **2.2 On-the-Road Testing**

Assessing medical fitness to drive using on-the-road testing is often considered as the gold-standard because it directly tests the driver's functional ability and decision-

making in real-life traffic scenarios [11][12]. However, on-the-road tests are often expensive (e.g., \$300 to \$600 CAD per assessment) and have limited availability. Since they are often time-consuming and expensive, it is unfeasible to use on-the-road tests for screening all older drivers every year. To better replicate the driver's natural settings, Vlahodimitrakou *et al.* [12] described the development and evaluation of a new protocol called the Driving Observation Schedule for monitoring and characterizing the driving ability of older drivers using several trained observers. For a group of 33 older drivers, each driver was given a personalized testing route, called the DOS route, which started and ended at their homes. Although the routes were not the same, each driver encountered common traffic scenarios while being observed to assess basic safety skills such as checking mirrors when merging lanes, executing low-speed maneuvers, handling different traffic densities, turn signaling, and speeding. On-the-road testing can contribute to creating a baseline driving profile by using personalized testing routes, which could also be performed and monitored over time.

### **2.3 Driving Simulators**

Driving simulators have been widely used to study driving behaviours because of its advantages including a safer environment to test drivers where the risk to the driver is removed; greater ease to recreate specific traffic scenarios with greater precision, and the ability to test many participants under the same conditions while being able to measure many parameters. The disadvantages include researchers have not agreed upon the ecological validity of simulators and there has not been enough agreement between simulators systems to have recommended pass-fail thresholds unlike the in-office assessments [11]. Imamura *et al.* [13] proposed a warning system based on data collected

from simulators of 11 participants to reduce MVC by detecting abnormal vehicle steering wheel operations. Lee *et al.* [14] explored the use of driving simulators to identify older drivers at higher risk of MVC for 129 participants. Results showed that the cognitive functions associated with crash risk were working memory, decision-making under pressure, and confidence in driving at high speeds. There was some evidence that the simulator could identify older drivers at inflated risk but it required further research on the predictive validity as a screening tool. From examining the validity and reproducibility of evaluations of older drivers using driving simulators, Bedard *et al.* [15] and Lee *et al.* [16] suggested that simulators can be used to facilitate the evaluation of medical fitness to drive.

## **2.4 Questionnaires and Surveys**

Questionnaires and surveys are widely used in driving behaviour studies because of their low-cost and simplicity in administering and collecting data from participants. When asked to self-report and self-assess driving ability, older drivers tend to be biased and inaccurate because they are prone to memory bias, tend to have difficulty in recalling events, and may have limited cognitive capacity due to aging and cognitive impairments [11][12][17]. It is more often that family members raise concerns about the older drivers than themselves, since it is expected that older drivers are reluctant to give up their driving as it is tied to their social independence. Crizzle *et al.* [18] studied the self-regulatory practices of drivers with Parkinson's disease to assess the accuracy of the patients' self-reporting using in-vehicle sensors to track 26 participants. Results showed that self-reported metrics, such as distances driven and time of day driven, should not be taken at face value by researchers because they were prone to recall errors and have less

awareness of their driving ability. Porter *et al.* [19] examined the older drivers from the Candrive study and their accuracy in self-reported travel distances. Their results showed that self-assessments can be useful for broad group comparisons, but not for individual assessment.

## **2.5 In-Office Medical Assessments**

Many older drivers are assessed by in-office medical assessments, which are used as a proxy for testing driving ability. The medical assessments are commonly used by the primary-care physicians because of its wide availability, easy administration, and the low cost. There exist many in-office medical assessments that the primary-care physician can use to assess physical and cognitive ability of older drivers. Among the most common assessments include the Montreal Cognitive Assessment (MoCA) [20], Mini-Mental State Examination (MMSE) [21], Trail-Making Test A and B [22], Screen for the Identification of Cognitively Impaired Medically At-Risk Drivers – A Modification of the DemTect (SIMARD-MD) [23], Clinical Dementia Rating scale, Clock-Drawing test, Useful Field of View [24], and mobility tests [25]. Each assessment tests a particular function, such as visual acuity or decision-making, to identify deficits in normal functioning to warrant interventions, which include driving cessation programs, usage of assistive devices, and driving restrictions. In these assessments, there is a recommended pass-fail threshold to detect unsafe older drivers, but the thresholds may be too broad. The scores from several assessments are often cross-referenced to make a decision, which is ultimately based on the professional judgment of the primary-care physician [26]. This screening process is often considered to lack evidence-based information [11]. Bédard *et al.* [27][28] studied the validity of SIMARD-MD and provided evidence contrary to the

original claim by the founding researchers of SIMARD-MD that "...a high degree of accuracy that can be used for immediate decisions in clinical setting". Results showed that the scores from SIMARD-MD were not significantly different in performance than the other established assessments (e.g., MMSE), and should not be used as a standalone tool for making decisions. Kwok *et al.* [7] studied the predictive validity of MoCA as a screening tool for driving ability. From retrospective analysis of 154 older drivers who completed the MoCA, the predictive performance had 84.5% sensitivity and 50% specificity and concluded that although MoCA can be a valuable screening tool, its predictive validity was not strong enough to be used as the sole tool for identifying medically unfit older drivers. Hollis *et al.* [6] studied the validity of MMSE and MoCA in the prediction of driving ability for drivers with cognitive impairments. From a retrospective analysis of 92 older drivers, their medical assessment scores were compared to their performances on a standardized on-the-road test. For drivers with pre-established diagnosis of cognitive impairments, there was a significant relationship between MoCA and on-the-road test scores, but no relationship in the drivers that were healthy. They concluded that MMSE and MoCA scores were not reliable for those who have not been previously diagnosed with cognitive impairments.

## **2.6 In-Vehicle Sensors**

To provide a more evidence-based assessment of driving ability, unobtrusive sensor systems, including smartphones [29], have been proposed for the continuous monitoring of daily driving behaviour to monitor driving performance and detect impaired behaviour. With the advances in sensor technology and vehicle technology, researchers can easily track their subjects while driving to build a detailed historical

record of actions during important traffic events to understand the exact factors and decision-making that influences driving safety. An emerging research method is the large-scale driving study that typically involves outfitting the participants' vehicle with sensor systems and data logging devices to track the participants over many months and years. In NDS, the participants would drive in their natural settings and important traffic events, such as MVC, may occur. The researchers can examine the pre-crash and crash data for factors and predictors of MVC. This method can provide greater external validity of driving behaviours of older drivers than other empirical methods like simulators and on-the-road testing.

### **2.6.1 Naturalistic Driving Studies**

The first NDS in the U.S. was the 100-car naturalistic driving study [30] (ended in 2006) and it built a large database of examples of impaired driving, high risk maneuvers, and traffic violations from ordinary drivers. The data included video recordings and kinematic sensors for 241 drivers over 1 year (i.e., over 2 million miles of driving). Among the list of objectives include analysis of rear-end events, period of inattention, lane changing, and monitoring driving behaviour over time. One of the findings was that 80% of all crashes involved the driver looking away from the forward direction. SHRP2 (2006 to 2015) [31] is another NDS study in U.S. that examined the role of driving behaviour and performance in traffic safety including the factors that affect collision risks. For 2,360 participants, video cameras and sensors were installed in their vehicles and baseline tests were conducted to score executive function and cognition. UDrive (2012 to 2017) [32] is the first NDS in Europe and includes all types of vehicles such as automobiles, trucks, and scooters. Its goal was to gain a better understanding of driving

behaviour in everyday traffic scenarios including important factors by using a combination of video cameras, external sensors, and the CAN-bus sensors in the vehicles. Ozcandrive (2008 onwards) [5], which is affiliated with Candrive and the Australian NDS (2015 onwards) [33] are large scale driving studies in Australia. Candrive [5] (2008 onwards) is a one-of-a-kind NDS for older drivers.

Eby *et al.* [34] explored the use of various in-vehicle sensors to describe driving behaviours that may be common among older individuals with early stages of dementia and compared the same metrics with those without cognitive impairments. For 17 drivers who had early stage of dementia and were medically fit to drive, they had their vehicles instrumented with a sensor system and drove 1 to 2 months in their natural setting. The in-vehicle sensor system included video cameras, forward radar sensors, accelerometers, gyroscopes, and GPS antennae. The 17 drivers were compared to 26 normal drivers without dementia, and those with dementia had significantly restricted driving space compared to the normal drivers. Both groups drove safely, but the drivers with dementia were more likely to lose their directions and get lost. Their future work included increasing the number of participants to reduce self-selection bias and to provide a more comprehensive characterization of the cognitive ability of the drivers with dementia since their cognitive ability could have had a large variability. Babulal *et al.* [35] conducted a small scale study on 5 older drivers by using in-vehicle sensor data to understand the driving behaviours of older drivers. They suggested analysis methods for characterizing the driving behaviour and metrics including detecting hard braking, sudden acceleration, speeding, number of trips taken, and their unique trip destinations (using unanonymized data).

## 2.7 Identifying Patterns in the Driving Data

By collecting large amounts of data using in-vehicle sensors, researchers can identify patterns and create models of driving behaviour to further understand and explore the complexities of driving behaviour. There are many applications of modeling individual driving behaviours including identification of driving styles and preferences to enhance human-computer interaction and improve driving comfort; detection of abnormal behaviour for warning systems; detect the current state of the driver (e.g., drunk, drowsy, aggressive, or calm); and driver identification for vehicle theft and insurance claims [36]. Many current approaches use traditional statistical methods and pattern recognition techniques including machine learning models to actively and automatically learn the patterns in the driving data and model driving behaviour [37]. Hidden Markov Models (HMM) and Artificial Neural Networks (ANN) have been used [36] to model the short-term sequential behaviours (e.g., turned the steering wheel and then pressed gas pedal) and long-term behaviours (e.g., accelerated, then turned, and then left lane change) to recognize and predict the driver's intentions. Al-Sultan *et al.* [38] proposed a context aware system to detect abnormal behaviours from a driver to warn other vehicles in the vicinity of the high-risk driver. Their system used a Bayesian network to classify whether the driver was normal, drunk, reckless, or fatigued by combining information about the environment, the driver, and vehicle. Although the design was based on simulated data, it showed that contextual information was important for modeling driving behaviour.

### 2.7.1 Driver Identification

The recent methods [39] of using machine learning algorithms to analyze driving styles construct driving profiles that can uniquely represent and identify one driver from another. Most of the proposed systems rely on detection of specific driving events and how these events were executed to characterize the driver including aggressive driving such as speeding, harsh acceleration and deceleration, tailgating, and improper lane changing. Many common classification algorithms have been studied for use in characterizing driving behaviour. Enev *et al.* [40] proposed a system that can accurately differentiate drivers based on their “driving fingerprint” (i.e., driving signature). Using in-vehicle sensor data extracted from the CAN-bus including brake pedal position, each of the 15 participants performed a series of maneuvers in a parking lot and drove a designated open-road course for 50 miles. Their approach used overlapping sliding windows to extract features to capture the statistical and morphological characteristics of the entire time-series data. Using one-vs-one multiclass classifiers and a 90/10 cross-validation (CV) testing, they claimed to have 100% accuracy in differentiating 15 drivers only needing 90% of the in-vehicle sensor data and ensuring none of the sliding windows overlapped in either testing or training sets. Hallac *et al.* [41] showed some evidence that drivers can be identified by how they turn from the in-vehicle sensor data. For the same turns, several CAN-bus signals, including steering wheel angle, could distinguish between drivers with a random-forest classifier achieving an average of 76.9% accuracy for a group of 2 drivers and 50.1% for 5 drivers across all 12 different types of turns. There was no single best feature that distinguished between drivers, and some drivers were more easily identified than others. Meng *et al.* [42] used HMM to distinguish

between 7 drivers in a simulated driving course with 80% accuracy. Instead of manually tuning parameters or manually identifying patterns, Dong *et al.* [43] proposed the use of deep learning to learn from the GPS data of drivers' trips and automatically extract features and identify complex patterns. They showed some evidence that it can outperform traditional machine learning techniques in differentiating drivers. Similarly, Kuderer *et al.* [44] proposed a reinforcement learning system to automatically model the parameters that fit the driver's particular driving style and can predict trajectories of vehicles in specific driving scenarios.

### **2.7.2 Motor Vehicle Operations**

A common theme among creating driving profiles is identifying important driving events (e.g., traffic incidents) and how drivers controlled the vehicle (i.e., braking). On one hand, analysis of the drivers' routes and destinations gain understanding of their overall trip planning and preferences including trends on a macroscopic level [45]. On the other hand, details within the trips such as traffic incidents, driving events, and evasive maneuvering [46] explore the behaviours on a microscopic level to gain further insight into their overall driving behaviour. Many industries that use in-vehicle sensors and vehicle telematics incorporate both the trends in the routes and the details in the motor vehicle operations to create the driving profile for their drivers. For example, automotive insurance companies that employ insurance telematics to quantify driver's risk measure the driving ability of their customers to adjust premiums, instead of basing the pricing on vehicle characteristics and driving record. The metrics include travel distances, time of day travelled, hard acceleration, hard braking, cornering speed, idle time, and speeds relative to the speed limit [47][48]. These programs can lead to cost-savings for safer

or infrequent drivers and promote and encourage safer driving behaviours leading to reduction in MVC. Guo *et al.* [49] used in-vehicle sensor data to identify factors that were associated with individual driving risk and predicted the high-risk drivers from their driving data. Using data from the 100-car NDS, the critical incident events (e.g., high acceleration or deceleration) showed some evidence of reliably predicting high-risk drivers. Ellison *et al.* [50] proposed a driving behaviour profile framework for road safety analysis to score the driving behaviour by assessing aggressive acceleration and braking among other maneuvers.

## **2.8 Turn Identification**

The way drivers negotiate around corners and bends in their trips can reveal individual characteristics of their driving behaviour. Ellison *et al.* [50] and Alhajyaseen *et al.* [51] showed that cornering is an important maneuver when assessing road safety and that cornering is actively monitored by insurance telematics to assess individual risks. Johnson *et al.* [52] and Hong *et al.* [53] used a combination of sensors (e.g., GPS, accelerometer, gyroscope, and magnetometer) to classify vehicle movements including detecting when the vehicle was turning. Hallac *et al.* [41] classified turns as a change in heading of at least  $70^\circ$ , at most 10 seconds in duration, and travelling relatively straight 5 seconds before the turn. Howcroft *et al.* [54] used the Candrive data and categorized a subset of acceleration events as turns, which were originally proposed by Wallace [55], as changes in heading of  $70^\circ$  to  $110^\circ$  across the entire event (i.e., total change in heading from the start to end of the event). Wallace *et al.* [56] identified when the vehicle was turning using only the GPS data from a smartphone GPS sensor and the proposed turn-identification algorithm was based on the change in heading of  $30^\circ$  between sampled GPS

coordinates; their performance was 73% true positive rate, 3% false positive rate, and 14% false negative rate tested on 215 true turns. In the same study, the turn-identification algorithm was compared to using the Google Maps GIS algorithm for identifying turns. The proposed turn-identification algorithm outperformed the Google Maps GIS algorithm, which had 58% true positive rate, 4% false positive rate, and 21% false negative rate tested on the same dataset of 215 manually labeled turns. Most studies developed their own criteria for identifying and classifying turning maneuvers, since they have different types of sensor data (e.g., steering wheel angle or angular velocity from inertial measurement unit) and the definition of a turning maneuver can be subjective (e.g., sharp vs. gradual turns).

## **2.9 Driving Conditions**

Driving conditions (e.g., snowing) have been widely considered to influence driving decisions, and older drivers tend to consider weather as an important factor in trip planning [8][57]. Myers *et al.* [57] used in-vehicle sensor data to examine the driving patterns of older drivers aged 65 to 91 years old from November to March for 46 participants. Their results showed that the daylight was an important factor because travel distance varied significantly at night by month. The results did not show significant changes in behaviour between winter and non-winter driving or clement and inclement weather in contrast to self-assessments because they may have been accustomed to winter driving. Kilpelainen *et al.* [58] provided evidence that driving behaviour is affected more by the prevailing observable driving conditions, rather than the forecasted weather.

## **2.10 Candrive Research Studies**

The complete protocol for Candrive [5] described the motivation, objective, and

experimental design of the Candrive study. About 1,000 participants were recruited from several major cities across Canada and data from in-vehicle sensors, medical assessments, and questionnaires were used to study the driving behaviours. The inclusion criteria for the participating older adults include: active and experienced drivers, age 70 years or older, expected to drive at least 4 times a week, under the care of a family physician, personal vehicle must be newer than 1995, and reside in their home province for at least 10 months of the year. The exclusion criteria for the older adults include: change of residence outside of their home province, medical contraindication to driving within the last 6 months according to the Canadian Medical Association Guidelines (e.g., heart attack), and diagnosis of a progressive condition (e.g., dementia) [5].

Preliminary studies of Candrive mainly focus on the analysis of the data collected in the baseline year. Tuokko *et al.* [59] examined “age and gender differences on several psychosocial measures developed specifically for older drivers, as well as associations with health-related measures”. They found that older drivers felt less comfortable driving than younger drivers and men felt more comfortable driving compared to women, concluding that the age and gender influenced attitudes towards driving. Langford *et al.* [60] studied the associations between the annual distance driven and the medical assessments for fitness to drive. The participants were grouped based on total distance travelled in their first year: low mileage (<5001 km), medium (>5,000km and <15,000 km), high (> 15,000 km). The results provided evidence that high mileage drivers were more likely to be male and low mileage drivers were more likely to be female. The high mileage drivers were more likely to be younger and high mileage drivers were more likely to perform at the highest level. The study provided further support to previous

literature that low mileage drivers have increased risk and tend to have reduced fitness to drive because they already limit themselves and self-regulate since they feel less comfortable driving. Rapoport *et al.* [61] studied the relationship between cognitive performance (e.g., MMSE, MoCA, Trail Making Tests A and B), driving comfort, perceived driving abilities, and self-reported driving restrictions using linear regression analysis. The results suggested that psychomotor speed and executive function tests in Trail Making tests were the best medical assessments for driving ability, although not all studies showed common results. The conclusion was not definitive since there were weak relationships that required further analysis.

A.P. Smith [62] examined the seasonal variations in older drivers' trip distances for participants across Canadian sites using the in-vehicle sensor data. Specifically, the objective was to determine if there were significant changes in trip distances among older drivers between different driving conditions including winter vs. non-winter and clement vs. inclement weather using a subset of the participants from Candrive. From a group of 248 older drivers, there was a 7% decrease in trip distances during winter and a 5% decrease in trip distance during inclement weather. Although the decrease in trip distances was expected, the minimal change in trip distances suggested that older drivers may have not been adjusting their driving behaviours during adverse conditions as much as expected from drivers' self-assessments from literature. This analysis is categorized in level 5 of Figure 1-1 as "Analysis of factors that affect driving".

Smith *et al.* [63] provided evidence that certain common measures of physical ability (e.g., finger-thumb opposition) were stable throughout the baseline year and could be used to reliability monitor driving ability. This study provided support for the

categorization of two groups of older drivers: those clinically assessed to have stable characteristics in health scores and those with unstable. Howcraft *et al.* [54] used in-vehicle sensor data from those two groups of drivers to examine the acceleration driving patterns of older drivers by proposing the use of the coefficient of variation as a metric to distinguish between stable and unstable drivers, which was a new and ongoing study. From comparing 12 stable-health and 16 unstable-health older drivers in their baseline year, the study showed that unstable drivers had lower acceleration coefficient of variation than stable drivers for a few categories of acceleration events including changes in speeds of 15 km/h and 30 km/h. Future work was recommended to measure the accuracy and validity of this metric for identifying cognitive and physical impairments. This analysis is categorized in level 4 of Figure 1-1 as “Group analysis of older drivers”.

### **2.10.1 Differentiating Between Older Drivers**

The research conducted by Wallace [55] was the motivation behind the research of this thesis. The topics that were studied by Wallace included big data processing techniques for the large amount of structured data from the in-vehicle sensors, identifying differences and similarities between older drivers to characterize their baseline behaviour, and assessments of cognitive ability using in-vehicle sensor data.

The fundamental processing framework [64] described the necessary data processing steps for the in-vehicle sensor data including data validation, anonymization of GPS coordinates, and data augmentation to integrate external data (i.e., weather). The significance of the processing framework was to preserve privacy of the participants, especially because using the real GPS coordinates could easily identify the participants. There exist many other in-vehicle sensor signals, which were used by other studies, that

were not collected by Candrive because they were not easily accessible and required modifications to the participant's vehicle. These signals were located in the CAN-bus network in the vehicle which included turn signaling, steering wheel angle, braking pedal position, and the anti-lock brake system. Since the OBDII was designed for specific engine diagnostics for service technicians, the in-vehicle sensor data from Candrive only included a limited portion of all the sensor data that exists in a modern vehicle. For example, developers of autonomous vehicles can reverse engineer the CAN-bus sensor network to autonomously control the vehicle by using radar technology. The improper modifications of the CAN-bus signals could seriously change the vehicle operational settings and jeopardize the safety of the driver, which was among the list of the reasons for not collecting data from the CAN-bus.

Using the anonymized Candrive data, an analysis of 100 trips from 4 older drivers [65] can identify individual differences and similarities between drivers with similar traits. General characteristics of their trips including their chosen routes, total distances travelled, time of day of trips, duration of trips, distribution of vehicle speeds, and tendency to travel above or below the posted speed limit during the trips can be used to distinguish drivers and group drivers based on similar preferences. This analysis is categorized in level 5 of Figure 1-1 "Analysis of routes and destinations".

For the long-term monitoring of driving behaviour to detect cognitive and physical impairments, it is necessary to establish a baseline behaviour. The repetitive nature of trips can be used to compare similar trips taken over several years to assess navigational ability and vehicle operations, specifically turn signal usage. By measuring the differences between the routes driven and the optimal routes [66], a baseline

behaviour for trip planning could be used for assessing a change in cognitive ability. The creation of a baseline profile for the driver's efficiency of the trips driven can be scored and be used as a performance metric to be monitored for less efficient trip planning and taking less complex routes. A proposed scoring system [67][56] for variations in trips in an attempt to detect drivers with lower cognitive ability include scoring the turn signal usage for when the driver properly signaled before turning. Although it was expected that drivers normally miss some proper turn signaling, there should be consistency in performance over long-term and any change could indicate a cognitive impairment. This research was conducted using an iPhone 4S as the data logging device and working with the real GPS coordinates. This analysis is categorized in level 5 of Figure 1-1 as "Analysis of routes and destinations" and "Analysis of motor vehicle operations".

Furthermore, the patterns that exist in the vehicle acceleration and deceleration behaviours could be used to characterize the baseline behaviour and then observed over time for significant changes. The moments of substantial acceleration [68], called acceleration events, were studied to identify the drivers' unique acceleration behaviours and stability over the first two years of Candrive. The acceleration events were extracted for 14 drivers who were clinically assessed to have similar traits, and a two-phase relationship was proposed to model the acceleration vs. the change in speed for events accumulated over the entire baseline year, which revealed individual differences in behaviour (Figure 2-1). The two-phase model was derived from qualitative observations of the bivariate analysis of acceleration vs. magnitude of the change in speed within each acceleration event. It was observed that the acceleration behaviours of the older drivers had two distinct phases. The rate of change in acceleration with respect to change in

speed was greater for the first phase when compared to the second phase, which exhibited much smaller rate of change. The smaller changes in speed may have represented the behaviours that were more “influenced by external factors such as traffic” and the larger changes represented more of the driver’s preferences. When comparing all possible pairs of drivers from the group of 14 selected drivers, there was evidence that suggested at least 80% of the 91 unique pairs could be distinguished ( $p < .05$ ) comparing the distributions of the mean and max accelerations of the events (i.e., two-sample t-test). After characterizing the baseline year, comparing the two-phase relationship with the following year showed a mean correlation of 0.971 to the first year for the 14 drivers, which showed some evidence that this behaviour remained somewhat consistent over time.

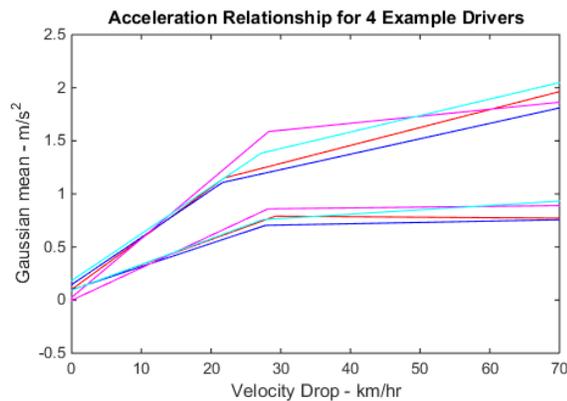


Figure 2-1. The proposed two-phase relationship between max acceleration and change in speed (top group) and mean acceleration and change in speed (bottom group) for 4 example drivers in their baseline year. The Gaussian mean represented the mean of the best fit Gaussian distribution that was fitted on the acceleration data points associated with each speed bin. The piecewise linear regression was shown to be a good fit for representing the general trend of the underlying acceleration events from the in-vehicle sensor data [68].<sup>2</sup>

Similarly, the deceleration events were extracted for the 14 drivers with similar traits [9][69] and the two-phase model was proposed (Figure 2-2) including a list of

<sup>2</sup> Permission was obtained from the author Bruce Wallace (wally@sce.carleton.ca), © 2016 IEEE

distinguishing features that could identify who is driving based on the in-vehicle data from a trip. To determine the transition “elbow” of the two-phase model, the transition point was varied through the change in speed (km/h) and the point that yielded the highest correlation coefficient or the best fit to the underlying data was selected. Each driver had their own unique two-phase model. With the two phase model from the baseline year, the same model was fit to the underlying data of the second year and it was shown that the correlation was also high, which provided evidence that there was also consistent behaviour among the drivers over the years for deceleration events. The correlation of the two-phase model with the underlying data increased by eliminating the trips with few events until having 8 events per trip, which meant that more events in a trip could better distinguish a driver. From Figure 2-1 and Figure 2-2, each of the 4 example drivers had a transition elbow at around 30 km/h, which represented the minimum change in speed for analyzing their individual preferences. This analysis is categorized in level 5 of Figure 1-1 as “Identification of unique behaviours” and “Analysis of motor vehicle operations”.

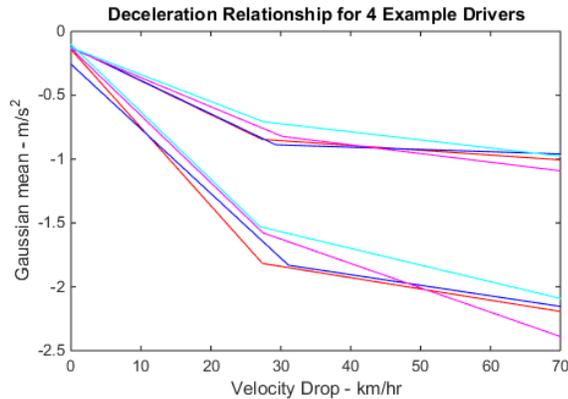


Figure 2-2. The proposed two-phase relationship between mean deceleration and change in speed (top group) and max deceleration and change in speed (bottom group) for 4 example drivers in their baseline year. The Gaussian mean represented the mean of the best fit Gaussian distribution that was fitted on the acceleration data points associated with each speed bin. The piecewise linear regression was shown to be a good fit for representing the general trend of the underlying deceleration events from the in-vehicle sensor data [9].<sup>3</sup>

Along with the acceleration and deceleration events, the list of distinguishing features [9] included trip distances, vehicle speed history, time of trip, the routes, and other trip attributes such as histogram of the acceleration values from the events. The trip attributes showed that it could distinguish at least 90% of the 91 unique pairs when comparing the distributions of the drivers ( $p < .05$  using two sample t-test). The two phase model fit the max deceleration better (mean correlation of 0.74) than the mean deceleration (mean correlation of 0.54). After identifying features from the acceleration events, deceleration events and trip attributes, a list of features that described the driving behaviour of 14 older drivers was defined [55], and pattern classification techniques were used to automatically identify complex patterns in the large data for differentiating drivers. The list of features specifically described how they drove (e.g., vehicle speeds), when they drove (e.g., time of day), and where they drove (e.g., the speed limit of the routes) including the trip distance, trip duration, and a histogram of the acceleration and

<sup>3</sup> Permission was obtained from the author Bruce Wallace (wally@sce.carleton.ca), © 2016 IEEE

speeds during the trips. A two-class classification scheme was proposed to distinguish between all pairs of drivers from a pool of 14 older drivers. For a total of 162 features, a 10-fold CV testing scheme was used to evaluate the performance of identifying who is driving given a set of features from a trip. Among a group of common classifiers, linear discriminant analysis classifier (LDA) achieved a mean error rate of 8.6% (91.1% accuracy) among all 91 pairs of drivers and the best error rate of 1.6 % (98.4% accuracy) for the most distinguishable pair when classifying the trips between two drivers. The performance was improved slightly by 2.7% for the least distinguishable pair by using more complex classifiers (e.g., SVM, ensemble classifiers). The best features were the speed limits of the routes, the histogram of speeds during a trip, and the two-phase acceleration and deceleration models for describing acceleration and deceleration behaviours. The worst features included the overnight driving times and when the vehicles were at stationary speeds. Altogether, the classification approach showed that patterns in the in-vehicle data can differentiate between pairs of drivers. Future work was suggested to explore other driving events to find additional features that improve differentiating drivers and to explore weather patterns and other important driving conditions to account for factors that affect driving. This analysis is categorized in level 5 of Figure 1-1 as “Identification of unique behaviours”, “Analysis of trips and destinations”, and “Analysis of motor vehicle operations”.

## **2.11 Summary**

Current methods of assessing medical fitness to drive of older drivers often lack evidence from real-world driving to support the physician’s decisions for driving interventions. To improve upon existing methods, emerging research on naturalistic

driving behaviours of older driver using large-scale driving studies (i.e., Candrive) could provide solutions to improve existing screening tools by using in-vehicle sensor data to quantify driving behaviour and identify important driving patterns. With a large amount of data, statistical and pattern recognition techniques are important tools for uncovering patterns in the driving data and identifying individual preferences. The patterns in the driving maneuvers of older drivers, such as accelerating and decelerating, have been shown to uniquely identify the driver and can be used to characterize their baseline driving behaviour.

## **3 Chapter: Methodology**

### **3.1 Introduction**

Real-world data from older drivers were obtained from the Candrive research study (see Chapter 1.1.2) to conduct the analysis on driving behaviour of older adults. Research ethics approval was obtained from Carleton University Research Ethics Board-B on May 25, 2016 with the protocol number #100680. The raw data for each participant required preparation to minimize errors and to improve accuracy of the data analysis. Wallace *et al.* [64] have pioneered the analysis of the Candrive in-vehicle sensor data by developing a data processing framework to cover the fundamental processing steps and storage methods of the raw data (e.g., anonymization and data integrity checking) and proposing methods for the analysis of the acceleration and deceleration of the vehicles [69] and per trip metrics [9]. For this study, the data processing steps were developed as an extension to the analysis performed by Wallace *et al.* for the purpose of the analysis of specific driving maneuvers.

This chapter provides a description of the dataset from the Candrive study and the details behind the original data collection procedure. Following the background on the data collection, a detailed description of the processing pipeline is presented to describe the steps for preparing the raw data for analysis: these steps include data verification and identifying invalid data, deriving data variables, and data imputation (i.e., handling missing values).

### **3.2 Candrive Research Design**

In the Candrive study, participants were recruited from several major Canadian cities, and a network of researchers, from the participating research institutions, had

collected data on the older drivers for their respective sites (i.e., home city). All participants underwent a baseline medical assessment at the beginning of their enrollment, which were on different dates. The participant's own personal vehicle was outfitted with the data logging device. Once the data logging device was installed in the participant's vehicle at the beginning of enrollment, it remained in the vehicle for the entire duration of the study.

### **3.2.1 Candrive Instrumentation**

The data logging device, designed and developed by Persentech (Winnipeg, Manitoba, Canada), comprises a central unit named the OttoView-CD, a GPS antenna, a RFID Antenna, and a key fob as shown in Figure 3-1. The device recorded data from the engine diagnostic sensors and the GPS antenna at 1 Hz sampling frequency (i.e., 1 second intervals between samples). Each data sample is time stamped in Coordinated Universal Time (UTC), and data were saved locally on a 4GB SD card. Data from the vehicles were recorded from the beginning until the end of study, and the raw data were retrieved from the SD card in the central unit of the data logging device in 4-month intervals. The data logging device was designed to automatically power on when the engine is turned on and automatically power down when the engine is turned off. When the data logging device was powered on, it automatically began the initialization routine lasting for a few seconds (e.g., obtaining the satellite fix) and then started recording the data from the engine diagnostic sensors via OBDII port and the longitude and latitude coordinates from the GPS signal. The RFID antenna was used to detect the driver's assigned key fob to keep a record of the identity of the driver in the case of a shared vehicle. The entire package is small and low-profile to allow for unobtrusive data logging.

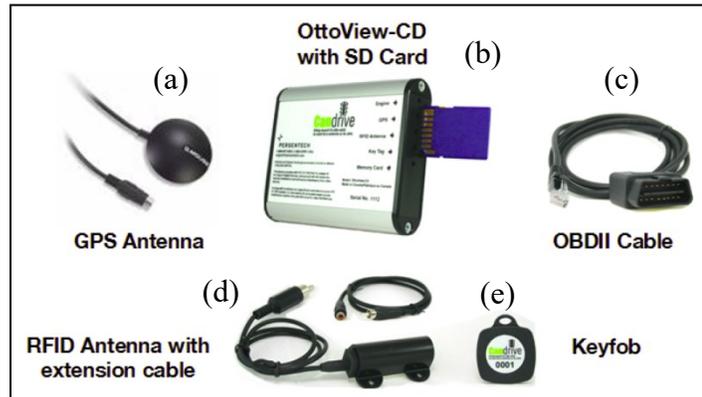


Figure 3-1. The data logging device, showing the five components, used by Candrive research study for data collection: (a) GPS antenna for recording longitude and latitude coordinates; (b) the central unit, which is similar to the size of the palm of the hand, for recording and storing the sensor data; (c) the connector from the central unit to the vehicle engine diagnostic sensors via OBDII port; (d) the RDIF antenna for detecting the nearby key fob to track the identity of the driver; and (e) the key fob that is assigned to the driver and attached to the car keys [70].

When the data were extracted from the data logging device, a proprietary Windows software application (Candrive Driver Tracking System v1.1.4098.15615) developed by Persentech (Winnipeg, Manitoba, Canada) must be used to access the data in the SD card. This application downloaded the latest cities' map data from a proprietary municipal database in a proprietary map format, and the information from the downloaded maps were used to integrate environmental information based on the recorded GPS coordinates such as road sign hazards, traffic alerts, and road speed limits. The GPS positional accuracy was within 10 m and GPS speed accuracy was 1.0 km/h [70], which was typical for civilian GPS systems. Although working at the limits of accuracy can contribute to overestimations or underestimations for the derived variables such as distance travelled or vehicle speed, the potential measurement errors can be mitigated by working with increased granularity of the data by using summary statistics such as means and median values. The dilution of precision (DOP) and the GPS fix status

were two sensor measurements (i.e., data variables) that were recorded with every single sample of the longitude and latitude coordinates. These values indicated the quality of the GPS reception; lower DOP values indicated better quality and the fix status indicated the number of visible satellites. Many other factors including satellite geometry, surrounding buildings, and atmospheric conditions can also affect GPS reception and cause inaccuracies in the GPS data. These factors were uncontrolled and were assumed to be inherent of the data.

### **3.3 Participants**

The participants used for this thesis research satisfied the following three inclusion criteria:

1. Resided in Ottawa, Canada
2. Sole driver of their vehicle
3. Stable in cognitive health, stable in physical health, and stable in general health in their first year of the Candrive study (collectively termed *triple-stable* or *stable-health*)

Only participants from the Ottawa site were selected for this study to control for the regional differences between cities. There were 257 participating older drivers from Ottawa and 28 of those older drivers were the sole drivers of their vehicles (i.e., they did not share their vehicles). This ensured that the driving data under analysis was guaranteed to be from the intended driver, since some participants who shared their vehicles had accidentally shared their key fobs. Out of those 28 sole drivers, 12 drivers were clinically assessed to be *triple-stable* during their respective first year of enrolment; that is, stable in cognitive, physical, and general health (according to definition in Appendix A.1). This

group of participants had a relatively constant health condition throughout their first year, which controlled for the effects of changes in health on driving behaviour. Since these 12 drivers were only categorized as stable-health in their first year, their in-vehicle data collected from their first year was selected for this study. These 12 triple-stable, sole drivers of their vehicles are listed in Table 3-1 including the amount of driving from their first year and summary statistics including standard deviation ( $\sigma$ ).

Table 3-1. Summary of first year data from the 12 triple-stable older drivers who resided in Ottawa and were the sole driver of their vehicle

<b>ID #</b>	<b>Number of Trips*</b>	<b>Total Distance Travelled (km)</b>	<b>Number of Hours</b>
1	1,344	11,686	275
2	2,084	15,858	574
3	982	9,900	268
4	1,482	34,788	515
5	647	15,063	245
6	2,222	16,436	373
7	1,465	11,241	336
8	1,032	14,179	296
9	1,072	15,950	289
10	2,771	35,596	741
11	1,211	15,554	331
12	1,360	13,777	314
Max	2,771	35,596	741
Min	647	9,900	245
Mean	1,473	17,502	380
$\sigma$	603	8,517	151

\*Please see Chapter 3.5 for definition of *Trips*

### 3.4 Raw Data Processing

A complete list of the sensor measurements (i.e., data variables) recorded by the data logging device is shown in Appendix A.2. Below is a list of the preliminary processing steps performed by Wallace *et al.* [64] for the anonymization and preparation of each raw file:

- 1) The first valid GPS longitude and latitude coordinate of the data files were set as the origin for all proceeding coordinates, and all the following coordinates within that file were altered to be relative to the new origin.
- 2) The recorded samples were removed if there were an unexpected number of data variables, contained invalid calendar dates, or were missing sample numbers (< 1 % occurrence).
- 3) The time zone was set to Eastern Standard Time (EST) for Ottawa (i.e., -5:00 from GMT) without daylight savings.

### **3.5 Data Processing**

From the sensor signals that were recorded Appendix A.2, the data variables listed in Table 3-2 were chosen for this study because the other data variables did not contain useful information related to the vehicle movement. For location, the longitude and latitude coordinates were recorded without elevation, which was not considered in this study. For all distance calculations, the earth was assumed to be an ellipsoid according to the World Geodetic System 1984 (WGS84) world standard for GPS reference coordinate systems. Vehicle speed was both measured from the engine computer ( $V_{OBDII}$ ) and from the GPS antenna ( $V_{GPS}$ ). The posted speed limit ( $\bar{V}$ ) was recorded at every location of the vehicle and included the following speeds: 0, 40, 50, 60, 70, 80, and 100 km/h, where 0 km/h indicated an absence of speed limit information for the current roadway. The trip number was an incremental counter for every new trip taken by the driver.

Table 3-2. The data variables that were selected from the raw sensor data for the analysis of driving behaviour.

<b>Data Variables</b>	<b>Units/Format</b>	<b>Symbol</b>	<b>Format</b>	<b>Example</b>
Date & Time	Date String	-	DD Month YYYY hh:mm:ss	27 July 2012 11:09:37
Longitude & Latitude	Decimal Degrees	-	(XX.XXXXXX, XX.XXXXXX)	(23.123456,- 34.654321)*
OBDII Speed	km/h	$V_{OBDII}$	XX	82
GPS Speed	km/h	$V_{GPS}$	XX.X	82.5
Posted Speed Limit	km/h	$\bar{V}$	XX	40
Trip Number	Integer	-	XXX	200

\*Altered for anonymity

The processing pipeline, shown in Figure 3-2, was implemented in MATLAB 2017b using parallel processing on the local machine to process the selected raw sensor from Table 3-2.

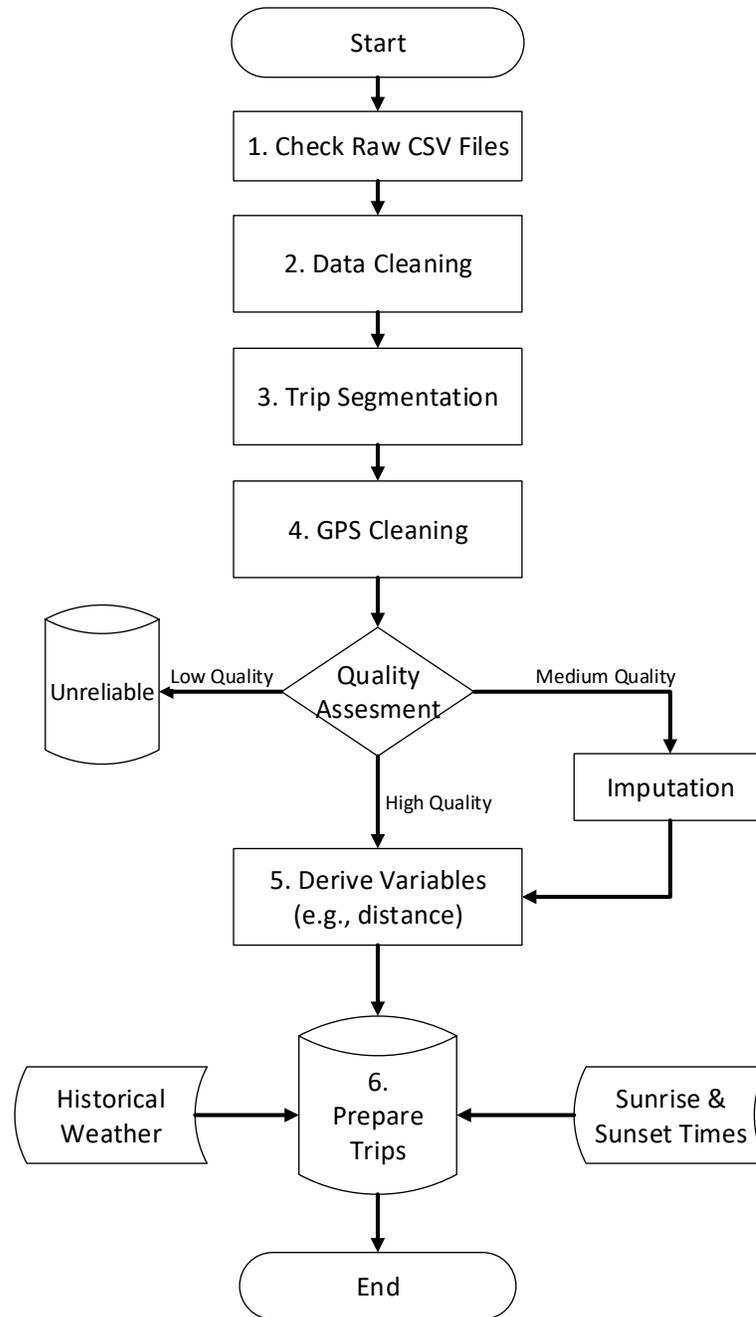


Figure 3-2. Processing pipeline for the data processing steps of the sensor data

For Figure 3-2, each of the numbered steps are described below:

- 1) A unique identification number was assigned to each participant and the CSV files that contained the data were organized by driver. Each measurement was verified to have realistic values (e.g.,  $V_{GPS} < 150$  km/h). If missing or invalid values were

detected at the beginning or end of the data in the file, these values were replaced with the nearest neighbouring valid value, whereas linear interpolation from the adjacent valid values was used for the rest of the data in the file. Appendix A.2 lists the data verification rules applied to each data variable and the corresponding occurrence rates of the invalid values.

- 2) The sensor data was segmented into *trips*, which were the periods between the automatic starting of the data logging when the engine turned on until the automatic stopping of the data logging when the engine turned off. Trips that were less than 2 minutes or shorter than 100 meters were removed (~1% of all trips), which represented events such as accidentally turning on and off the engine or when the vehicle was idling.
- 3) The longitude and latitude coordinates were inspected for invalid values by checking the following: fix status is “No Fix”, contained missing values, contained impossible distances travelled between coordinates (i.e.,  $> 150$  km/h), or DOP  $> 5$ . All the invalid GPS coordinates were identified for each trip (~6% of all samples), and the trips were categorized into three groups based on their invalid GPS coordinates:
  - **High Quality:** absence of any invalid coordinates (~98% of all the trips). From visually inspecting a sample of 30 routes from different drivers, the longitude and latitude did not have any visible GPS jitter that is common in smartphones-based GPS tracking [66]. Smoothing filters were not required for the longitude or latitude signals.
  - **Medium Quality:** contained a few invalid coordinates such as when the vehicle travels momentarily under a bridge or through a tunnel. A few

instances when the sensor data was erratic at the beginning of the data logging were observed and these were imputed by replacing with the nearest valid neighbouring coordinate as shown in Figure 3-3. For invalid values during a trip, it was imputed by linearly interpolation of the longitude and latitude values. The GPS speed was imputed by copying the values from the OBDII speed,  $V_{GPS} = V_{OBDII}$  (~1% of all the trips).

- **Low Quality:** contains many invalid coordinates (i.e., more than three contiguous segments of coordinates lasting 10 s) during the trip. These trips were considered unreliable and were subsequently discarded (<1% of all the trips).

From the quality assessment of all trips, a small portion of the trips (~ 2%) have medium or low quality.

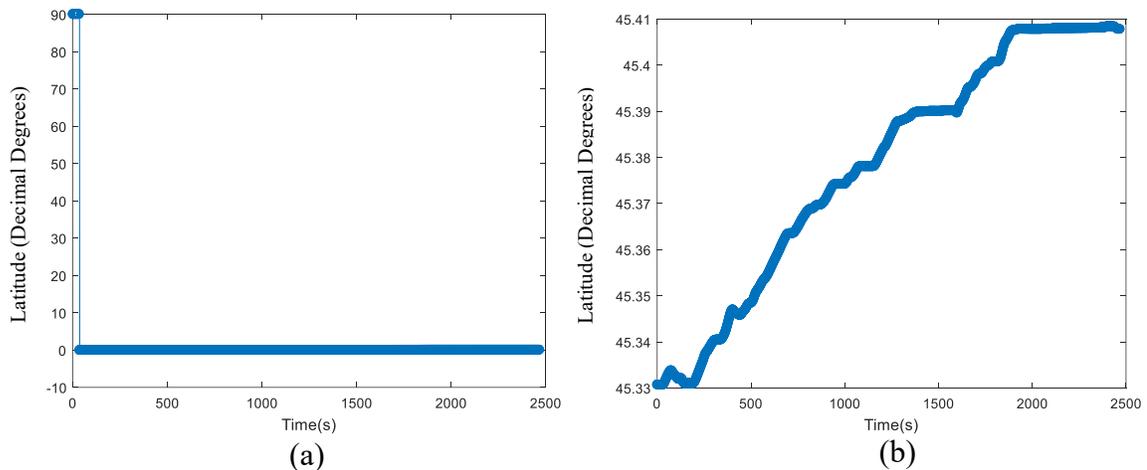


Figure 3-3. (a) Invalid GPS coordinates were detected in the latitude channel. (b) The same latitude data after cleaning had no extreme values.

- 4) To be able to use Universal Trans Mercator (UTM) map projection for deriving variables, the coordinates must be shifted back to the Ottawa region to obtain the proper map zone. Accordingly, the origins were shifted to a location that was

roughly in the middle of downtown Ottawa. The new origin became (+45.416666°,-75.700000 °), such that the GPS data of all the participants were approximately in the Ottawa region.

- 5) The speed measured from the GPS ( $V_{GPS}$ ) was compared to the speed from the engine computer ( $V_{OBDII}$ ) for each trip to remove any values that were significantly different from each other by replacing any  $V_{GPS}$  that were greater than three standard deviations of the relative differences (Eq. 3-1) with the  $V_{OBDII}$  [6, p.50].

$$relDiff[n] = \begin{cases} 0 & V_{GPS}[n] \&\& V_{OBDII}[n] = 0 \\ \frac{V_{GPS}[n] - V_{OBDII}[n]}{mean(V_{GPS}[n], V_{OBDII}[n])} & otherwise \end{cases} \quad \text{Eq. 3-1}$$

Typically, automobile manufacturers design the speedometer to report slightly higher than actual speed of the car to err on the side of caution and improve safety. Moreover, Wallace [55] compared the speed reported from the OBDII ( $V_{OBDII}$ ) and the GPS sensor ( $V_{GPS}$ ) using the Candrive sensor and showed evidence that  $V_{OBDII}$  had an accuracy of 10% (e.g.,  $\pm 4$  km/h at 40 km/h), while the  $V_{GPS}$  had an accuracy of  $\pm 0.4$  km/h. Thus, the  $V_{GPS}$  was more accurate and was considered the primary source of the vehicle's speed, and the  $V_{OBDII}$  was the secondary source. However,  $V_{GPS}$  was prone to errors due to occasional GPS drifts and inaccuracies. The  $V_{OBDII}$  had greater consistency from not being afflicted by GPS errors, but also contained more missing values. Therefore,  $V_{OBDII}$  was a reliable secondary source for correcting the GPS speed when invalid values were detected (~6% of all samples).

- 6) The time-series of GPS coordinates, called the *GPS trace*, contained errors due to the GPS drifts, where there was recorded movement of the vehicle according to the GPS, but the automobile was stationary according to the  $V_{OBDII}$  (e.g., stopped at a red



account for daylight savings time.

### **3.5.1 Historical Weather Data**

The historical hourly and daily weather in Ottawa was integrated into the data. Since the participant's exact location was anonymous, the weather was approximated by using the weather recorded by the main airport in Ottawa, the Macdonald–Cartier International Airport. The files were downloaded from the public website hosted by the Government of Canada website [71]. The downloaded data were formatted into CSV files for each respective year. The weather information (e.g., temperature, outlook, rain level) was integrated into the data by adding the weather to each sample for every trip. The missing values were imputed by using the nearest neighbouring valid samples and empty data variables were removed. The weather was reported in Local Standard Time and thus, daylight savings time was added manually. A summary of the data verification and imputation methods for the hourly and daily reported weather data is presented in Appendix A.3.

### **3.5.2 Sunset and Sunrise Data**

To determine the amount of daylight during the day, the sunset and sunrise times for each day was integrated into the sensor data. The sunset and sunrise times were obtained from the publicly accessible website hosted by the National Research Council of Canada [72]. The data were verified to have no missing values and all realistic times. The times were reported in Local Standard Time (Ottawa) and all times were manually adjusted for daylight savings time.

## **3.6 Derived Variables**

After performing data verification and imputation on the data variables, additional

variables were derived. Since the movement of the vehicle is essential for the analysis of driving behaviour, the acceleration, jerk, and distance travelled were calculated as described in Table 3-3.

Table 3-3. Derived variables for each trip from the processed sensor data.

Variable	Symbol	Units	Description
Instantaneous distance	$d[n]$	m	The distance travelled between the current coordinate and the previous coordinate was estimated by using the geodesic arc length (shortest distance) with an ellipsoid reference shape as define in WGS84. It was beneficial to relocate the anonymous map data to Ottawa for better estimation of the distances.
Total distance	$d_T$	km	The distance travelled during the entire trip.
Change in heading	$\Delta h[n]$	degrees	The change in heading is the change in the direction of travel from the current heading to the next as seen on a map using the UTM map projection (Eq. 3-2).
Direction of the change in heading	$\Delta h_{dir}[n]$	{left, right}	Describes whether the change in heading is leftwards or rightwards relative to the previous heading (Eq. 3-3).
Instantaneous tangential acceleration	$a_t[n]$	m/s <sup>2</sup>	Describes the rate of change in speed over time of the vehicle at each position. It was calculated using a 2-point central difference formula (Eq. 3-4) [73].
Instantaneous jerk	$j[n]$	m/s <sup>3</sup>	Describes the rate of change (i.e., smoothness) in acceleration over time of the vehicle at each position. It was calculated using a 2-point central difference formula (Eq. 3-4).

$$\Delta h[n] = \cos^{-1} \frac{\vec{A} \cdot \vec{B}}{|\vec{A}||\vec{B}|}$$

where,

$$\vec{A} = (x[n+1] - x[n], y[n+1] - y[n])$$

$$\vec{B} = (x[n] - x[n-1], y[n] - y[n-1])$$

Eq. 3-2

$$\Delta h_{dir}[n] = \begin{cases} \text{Left,} & \text{sign } \overrightarrow{(\overline{A} \times \overline{B})}_z = \text{positive} \\ \text{Right,} & \text{sign } \overrightarrow{(\overline{A} \times \overline{B})}_z = \text{negative} \end{cases} \quad \text{Eq. 3-3}$$

$$a_t[n] = \frac{V_{GPS}[n+1] - V_{GPS}[n-1]}{2} \quad \text{Eq. 3-4}$$

$$j[n] = \frac{a_t[n+1] - a_t[n-1]}{2} \quad \text{Eq. 3-5}$$

### 3.7 Map Projection

For the heading calculations ( $\Delta h[n]$  and  $\Delta h_{dir}[n]$ ), the measured units of position were converted from 3D coordinates system (i.e., longitude and latitude in decimal degrees) to 2D Cartesian coordinate system (i.e., x-axis called *eastings* indicate the horizontal position and y-axis called *northings* indicate the vertical position in meters as represented by  $(x[n], y[n])$ ) using UTM map projection [74]. The UTM map projection uses distinct codes for different regions, called *UTM zones*, on the Earth to minimize distortions in map projection and is among the most popular map projections. Since almost all the drivers' trips occurred in Ottawa, which is within a single zone (i.e., zone 18), the UTM map projection was appropriate for the most accurate representation of the 3D coordinates system. A disadvantage is that the UTM map projection is less accurate when the drivers' trips crossed UTM zones (e.g., travelling from Ottawa to Toronto crosses zones 17 and 18), however these long distance trips were uncommon. OpenStreetMap was used for the 2D map display of the GPS data.

### 3.8 Summary

To study the driving behaviour of older adults in their natural environment, this study used the real-world data that were obtained on stable-health older drivers from

participants of the Candrive research study, which used a data logging device to record daily driving activities of older drivers. This chapter described the Candrive study, the processing pipeline, and the data integration with external data sources to prepare the raw sensor data for analysis. The benefit of the data processing was minimizing errors during the analysis of the driving behaviour by cleaning invalid measurements to obtain a more accurate record of their driving activity. From the quality assessment of all the trips, about 98% of the trips did not have any invalid GPS data, which indicated that most of the data were high quality. Around 1% of the trips had minor errors that required data imputation, while another 1% of trips contained major errors and were discarded. Although about 98% of trips were high quality, GPS drifts occurred in each instance the vehicle was stationary, and thus the GPS drifts were identified and corrected for all trips. By completing the data verification and processing steps, the sensor data were prepared for the analysis of driving maneuvers.

## 4 Chapter: Identifying Turns in the GPS Data

### 4.1 Introduction

One of the most important driving maneuvers is maneuvering the vehicle around bends and corners on the road, called *cornering*. Cornering is a basic skill that every driver develops and is an essential skill for driving safely. For the novice drivers, cornering is simply no more than turning the steering wheel while obeying all the traffic signs and rules. For experienced drivers, advance driving techniques improve cornering efficiency and aid in difficult driving conditions. Drivers have their own individual preferences on how fast or slow they drive while cornering. For example, some drivers prefer approaching a bend at slower speeds while another at faster speeds, and some drivers prefer accelerating during the turn or upon exiting the turn.

To analyze the cornering of older drivers, the portions of the trip where the vehicle was turning must be identified. The GPS trace, which is the time-series samples of coordinates that comprise the path of a trip, contained corners and bends that indicate when and where the vehicle was cornering. Since there may be thousands of trips taken by a driver, a computer algorithm must be implemented to automatically and effectively identify the corners and bends in the GPS trace. Wallace *et al.* [56] have studied the turning maneuvers from GPS data and they used a common approach by applying a threshold to the change in heading to filter for *turning points*, which are the coordinates that lie within the turn itself. The proposed algorithm by Wallace *et al.* [56] had a 73% true positive rate, 3% false positive rate, 14% false negative rate for a threshold of 30° in change in heading.

This chapter proposes and evaluates the performance of three different algorithms

for identifying turns in the GPS trace. Firstly, the criteria of a turn are defined and the three turn-identification algorithms are explained. Secondly, GPS data were collected to create six independent trips to create a gold standard for the performance evaluation of the three algorithms. Lastly, the results from testing the algorithms are discussed including the false positives and negatives of each turn-identification algorithm.

## **4.2 Methodology**

### **4.2.1 Definition of a Turn**

A turn is a portion of the GPS trace where the vehicle was in the act of cornering, which occurred when there were substantial corners or bends in the GPS trace and the vehicle was consistently changing directions either leftwards or rightwards while travelling forward. Figure 4-1 shows an example of a typical turn in the GPS trace and the equivalent map display of the same turn. Using the data logging device as described in Chapter 3.2.1, unanonymized GPS data were collected from driving within the city of Ottawa in a 2016 compact sedan with automatic transmission to create test cases (i.e., trips) to evaluate the turn-identification algorithms. Unlike the GPS data from Candrive, this GPS data contained the actual coordinates for all the trips, which aided in accurately labeling the turns. Using the observations made from many manually identified turns, a list of the main characteristics are listed in Table 4-1. Although the on-ramp and off-ramps of an expressway are more like loops instead of corners or bends, they were considered to be turns as well. The turns that occurred while maneuvering in a confined area, such as a parking lot, were not considered because the drivers' behaviours (i.e., individual preferences) were better characterized by the common, reoccurring turns during the trips rather than the unpredictable and irregular maneuvering in response to

other drivers. A turn is represented by  $(x[n], y[n])$  as the eastings and northings coordinates of the turn in a trip for  $n = k, k + 1, \dots, k + N$ , where  $k$  is the starting sample of the turn and  $N$  is the duration of the turn. The turns must satisfy the criteria defined in Table 4-1. These criteria were necessary for achieving a low false positive rate for identifying turns.

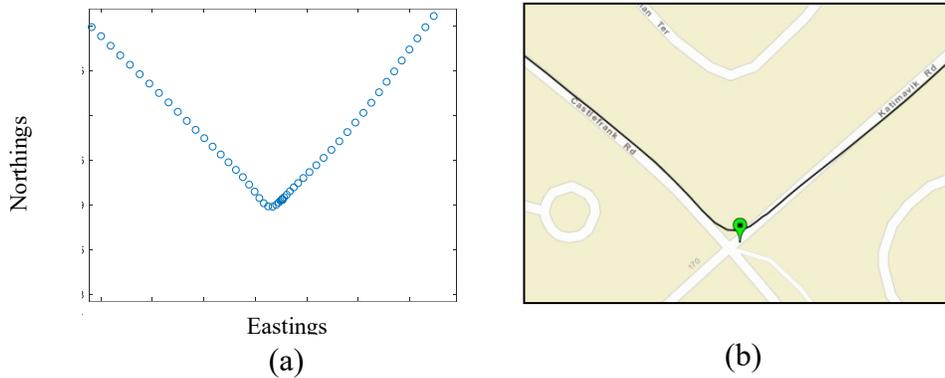


Figure 4-1. (a) Example of a turn showing the recorded GPS coordinates and (b) the same turn projected on a map with a green marker to indicate the center of the turn. The turning angle in the figures are not to scale.

Table 4-1. The criteria for a turn in the GPS trace.

Item #	Rule	Rationale
1	Duration is at least 3 seconds. $N \geq 3$	This is important to remove impossible vehicle movements due to occasional GPS inaccuracies and quick maneuvers such as overtaking a vehicle or changing lanes.
2	Constant heading direction throughout the turn $\Delta h_{dir}[n] = \textit{left or right}$	One continuous turning motion has a constant direction either leftwards or rightwards. A turn with a non-constant heading direction may indicate unusual steering such as when the driver had to maneuver the vehicle to avoid obstacle.
3	All change in headings must be between $6^\circ$ and $90^\circ$ . $6^\circ \leq \Delta h[n] \leq 90^\circ$	It was physically impossible to have a change in heading that exceeded $90^\circ$ while travelling forward (i.e., second to second change in heading), but it was possible from maneuvering in reverse, such as leaving the driveway, which was not counted as a turn.

		After manually labeling and inspecting the start and end of each turn from a variety of trips, the threshold of $6^\circ$ was determined for straight paths by selecting the minimum $\Delta h[n]$ among all the manually identified starts and ends of the turns.
--	--	--

### 4.2.2 Training and Testing

Six trips were created to evaluate the performance of the three turn identification algorithms (Table 4-2) using independently collected GPS data as described in Chapter 4.2.1. These trips were designed to test a wide variety of common turning scenarios encountered in real-world driving to be representative of the turns in the GPS data from Candrive. The start and end of every turn was manually labeled, according to the criteria in Table 4-1, for all test cases to create a gold standard for which to compare with the experimental results from each algorithm. Figure 4-2 shows an example trip #5.

Table 4-2. Description of the six test cases for the turn-identification performance evaluation.

<b>Trip #</b>	<b>Number Of Turns</b>	<b>Description of the Trips</b>
1	13	Expressway and downtown in light traffic
2	32	Expressway and downtown in heavy traffic
3	14	Expressway and residential in moderate traffic
4	68	Residential and arterial roads in light traffic#1
5	44	Residential and arterial roads in light traffic#2
6	97	Residential, arterial roads, and expressway in light traffic
Total:	268	There was minimal overlap in routes between trips

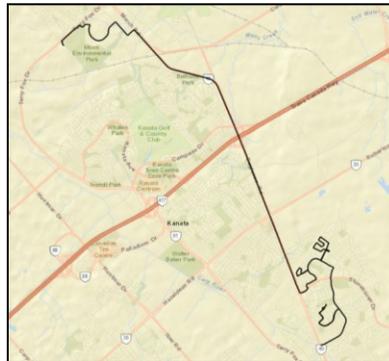


Figure 4-2. The entire route for trip #5 showing a variety of turns and paths taken by the vehicle.

### 4.2.3 Turn-Identification Algorithms

The three turn-identification algorithms involved locating the coordinates that occur within a turn, and then locating the start and the end of the turns. The first

algorithm was based on a threshold on the change in heading to find significant corners and bends, and this algorithm was implemented by Wallace *et al.* [56] for counting the number of turns in a trip. The second algorithm was an extension of the first algorithm by simplifying the GPS trace before applying a threshold on the change in heading to emphasize the corners. The third algorithm was based on the radius of curvature at each sampled coordinate to filter for turns based on its physical size.

#### 4.2.3.1 Algorithm #1: Threshold Heading

The first algorithm was the simplest of the three because it only applied a threshold to  $\Delta h[n]$  to determine the points that lie within the turns. The disadvantage of this approach was that the candidate turning points were affected by the vehicle's speed and not strictly on the geometry of the path. For the same turn, it is possible that travelling slower will change  $\Delta h[n]$  during the turn and making it fall below the chosen threshold. Another disadvantage is that it can be sensitive to gradual curves in the road and sporadic maneuvering, such as overtaking another vehicle, when the threshold is too low. Below is a list of the main steps of the algorithm:

1. Temporarily remove any coordinates that have 0 km/h vehicle speed  $V_{GPS}$  (i.e. when the vehicle is stopped). This ensures that interrupted turns are not counted as two turns.
2. Apply a threshold  $\theta$  to the change in heading ( $\Delta h[n] > \theta$ ) to select the desired candidate turning points (i.e., a coordinate that exists within a turn). A candidate turning point was represented by  $\Delta h_c^i[n]$  for the  $i^{\text{th}}$  candidate turning point for a trip.
3. For each candidate turning point, perform a forward (3a.) and backward search

(3b.) of the neighbouring coordinate samples to determine the start and end of the turn, which were the last neighbouring points that satisfied the criteria specified in Table 4-1.

3a. Forward search until  $\Delta h_c^i[n+k] < 6^\circ$  or  $\Delta h_c^i[n+k] > 90^\circ$

3b. Backward search until  $\Delta h_c^i[n-k] < 6^\circ$  or  $\Delta h_c^i[n-k] > 90^\circ$

4. Add the 0 km/h stationary coordinates back into the GPS data and mark the start and end points of all turns in the unmodified GPS data. The start and end points were the farthest neighbouring points from the candidate turning points  $\Delta h_c^i[n]$  that satisfied the criteria specified in Table 4-1. Also, ensure there were no duplicate turns by removing repeated turns from the list.

#### 4.2.3.2 Algorithm #2: GPS Trace Reduction

To improve upon the first algorithm, the second simplified and reduced the GPS trace before applying a threshold on the change in headings  $\Delta h[n]$ . Since the first algorithm can be sensitive to gradual curves and susceptible to occasional quick maneuvers, reducing these curves in the GPS trace to emphasize the location of the corners could improve performance. Below is a list of the main steps of the algorithm:

1. See Chapter 4.2.3.1 step #1.
2. Simplify and reduce the GPS trace to fewer coordinates by using Ramer-Douglas-Peucker algorithm (RDP) [75] [76], which is a line simplification algorithm with the goal of taking a curve, which was represented by line segments, and finding a similar curve that contained fewer line segments. A single parameter, the tolerance ( $\tau$ ), was the distance in meters of how far the new path (i.e., line segments) can deviate from the original path. For larger  $\tau$ , more of the sampled

coordinates were removed from the GPS trace and the sharper turns were preserved.

3. Apply a threshold  $\varphi$  to the change in heading created from applying RDP ( $\Delta h_{RDP}[n] > \varphi$ ), which was based on the reduced GPS data, to select the candidate turning points.
4. See Chapter 4.2.3.1 step #4.
5. See Chapter 4.2.3.1 step #5.

#### 4.2.3.3 Algorithm #3: Radius of Curvature

The first and second algorithms depended on the change in heading for identifying the candidate turning points. The third algorithm identified the turns based on the geometry of the road, but was also affected by the vehicle speed. To do so, the third algorithm inspected the radius of curvature at each coordinate in the GPS trace to determine the candidate turning points. The advantages include filtering out turns that occurred on much smaller and larger curves. Below is a list of the main steps of the algorithm.

1. See Chapter 4.2.3.1 step #1.
2. Fit a piecewise cubic Hermite interpolating polynomial (pchip) [77] to the longitude samples  $x(t)$  and latitude samples  $y(t)$  to obtain a piecewise continuous function that described the route travelled by the vehicle. Pchip is known as a shape preserving piecewise cubic interpolation because it has less oscillations than other interpolating methods (e.g., cubic spline) and better approximated the route by not introducing any curves in the path.
3. Calculate the first and second derivatives of the interpolated functions and then

calculate the instantaneous radius of curvature ( $R_c$ ) at each sampled coordinate using Eq. 4-1.

$$R_c = \frac{[(x')^2 + (y')^2]^{\frac{3}{2}}}{|x'y'' - x''y'|} \quad \text{Eq. 4-1}$$

4. Apply an upper threshold  $\rho_u$  and lower threshold  $\rho_l$ , ( $\rho_l < R_c[n] < \rho_u$ ) to the radius of curvature to obtain a list of candidate turning points.
5. See Chapter 4.2.3.1 step #4.
6. See Chapter 4.2.3.1 step #5.

#### 4.2.4 Performance Evaluation

The performance was evaluated using the six trips. The testing and training scheme followed a holdout cross validation (CV), where the first 5 smaller trips were chosen as the training cases to optimize the parameters and the 6<sup>th</sup> trip was chosen as the testing case to obtain the test scores to generalize the performances. For training, a parameter sweep was used to vary the thresholds and determine the parameters that achieved the best performances as follows:

1. Algorithm #1: Perform a parameter sweep from 0° to 90° in steps of 1° for the threshold angle  $\theta$ .
2. Algorithm #2: Perform a parameter sweep from 0 to 90° in steps of 1° for the angle  $\varphi$  and 1 to 30 m in steps of 1 m for the tolerance  $\tau$ . Disregard 0 m to avoid overlapping results with algorithm #1.
3. Algorithm #3: perform two iterations of the parameter sweep by first determining the upper threshold  $\rho_u$  from 0 to 50 m in steps of 1 m to remove gradual bends in the GPS trace and then repeat the parameter sweep for  $\rho_l$  from 0 m to  $\rho_u$  in steps of 1 m.

The experimental results of each test case were compared to the labeled trips and the results were summarized by the true positives (TP) (i.e., the number of correctly identified turns), false positives (FP) (i.e., the number of identified turns that were incorrect), false negatives (FN) (i.e., the number of missed turns), Precision (PC; Eq. 4-2), and Recall (RC; Eq. 4-3). The best parameters were chosen to maximize the weighted harmonic mean of the PC and RC, *F-score* (Eq. 4-4), where the  $\beta$  factor adjusts the weight between the FN and FP errors. Since it was desirable to have lower FP than FN and simultaneously a high TP,  $\beta$  was set to 0.5 to weigh the FN less to put a higher penalty on FP. It was more acceptable to have missed turns rather than include false turns because the missed turns would be represented by the large number of identified turns, as long as the FN were much lower than the TP.

$$PC = \frac{TP}{TP + FP} \quad \text{Eq. 4-2}$$

$$RC = \frac{TP}{TP + FN} \quad \text{Eq. 4-3}$$

$$Fscore = (1 + \beta^2) \left( \frac{PC \cdot RC}{(\beta^2 \cdot PC) + RC} \right) \quad \text{Eq. 4-4}$$

## 4.3 Results

### 4.3.1 Algorithm #1: Threshold Heading

The results from training on the 5 trips, using algorithm #1, were summarized in Figure 4-3.

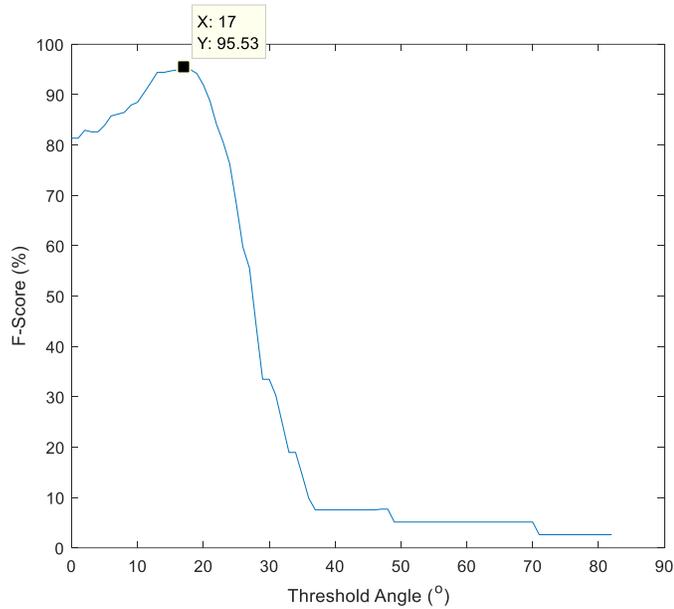


Figure 4-3. The performance of algorithm #1 on the 5 training trips showing the F-score for a parameter sweep of the threshold angle  $\theta$  between  $0^\circ$  to  $90^\circ$  in steps of  $1^\circ$ . The threshold angle  $\theta$  that maximized the F-score was  $17^\circ$ , as shown by the square marker.

From Figure 4-3, the largest F-score was 95.5% which corresponded to the  $\theta = 17^\circ$  with RC = 92.4%, PC = 96.3%, TP = 158, FN = 13, and FP = 6. The number of FN was 8.2% of all 268 turns. To verify the validity of the gold standard, RC was 100% when  $\theta = 0^\circ$ , which demonstrated that all the manually labeled turns were identifiable by the algorithm.

### 4.3.2 Algorithm #2: GPS Trace Reduction

The results from training on the 5 trips, using algorithm #2, were summarized in Figure 4-4.

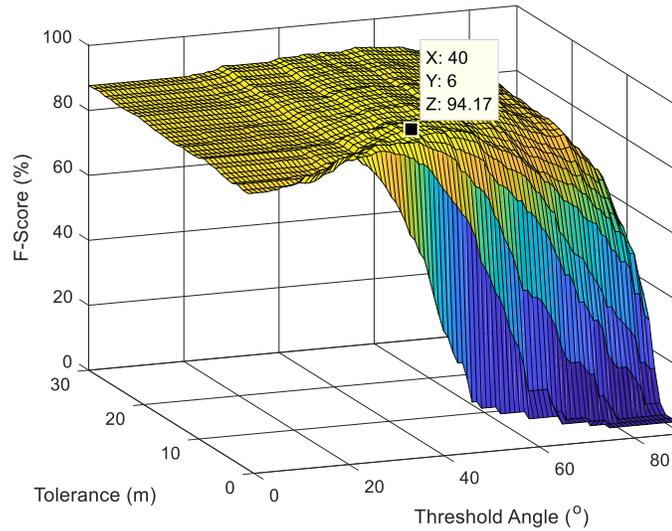


Figure 4-4. The performance of algorithm #2 on the 5 training trips showing the F-score for a parameter sweep of  $\varphi$  between  $0^\circ$  to  $90^\circ$  and 1 m to 30 m. The angle  $\varphi$  and tolerance  $\tau$  that yielded the best F-score was  $40^\circ$  and 6 m respectively, as shown by square marker.

From Figure 4-4, the largest F-score was 94.2% which corresponded to the threshold angle of  $\varphi = 40^\circ$  and  $\tau = 6$  m with RC = 90.6%, Precision = 95.1%, TP = 155, FN = 16, and FP = 8, where FN was 5.2% of all 268 turns.

### 4.3.3 Algorithm #3: Radius of Curvature

The results from training on the 5 trips, using algorithm #3, were summarized in Figure 4-5.

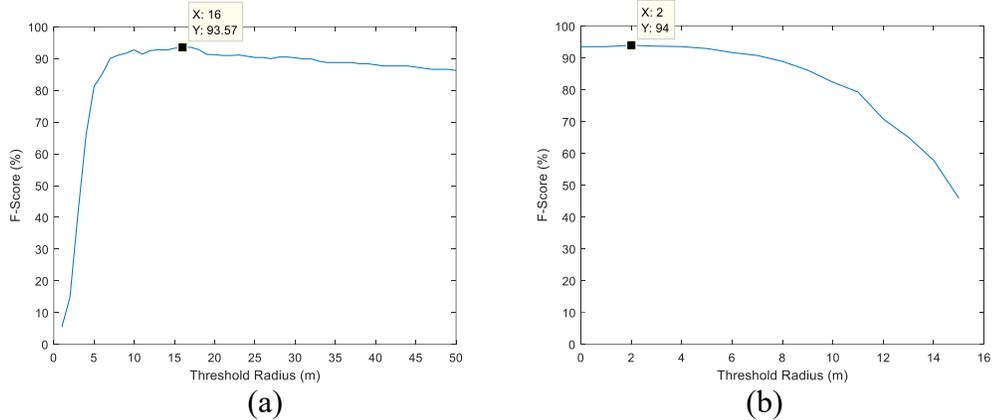


Figure 4-5. The performance of algorithm #3 on the 5 training trips showing the highest F-score as indicated by the square marker (a) for a parameter sweep of  $\rho_u$  between 0 to 50 m to determine the upper threshold, which was 16 m, and (b) the F-score for the lower threshold  $\rho_l$  was 2 m.

From Figure 4-5b, the largest F-score was 94.0% with Recall = 95.3%, Precision = 93.7%, TP = 163, FN = 8, and FP = 11, where FN was 3.0% of all 268 turns. The  $\rho_l = 2$  m and  $\rho_u = 16$  m thresholds were much smaller than the typical range of the actual geographical radius of turns (i.e., roughly 4 m to 70 m, please see Chapter 5). Since the pchip interpolation approximated the paths in between coordinates with nearly straight lines, the turns were essentially represented by many line segments and the instantaneous curvature at each coordinate  $(x[n], y[n])$  were much smaller than if a circle (i.e., curve) was fitted to a turn to calculate the approximate turn radius to describe the geographical size of the turn (Chapter 5).

#### 4.4 Discussion

The results from testing on trip #6 by using the optimized parameters from the training cases are summarized in Table 4-3.

Table 4-3. Comparison of the three turn-identification algorithms using the optimized parameters that yielded the largest  $F$ -score from training.

Algorithm	Parameters	TP	FN	FP	RC (%)	PC (%)	F-score (%)
1	17°	97	0	4	100	96.0	96.8
2	6 m, 40°	94	3	2	96.9	97.9	97.7
3	2 m, 16 m	97	0	9	100	91.5	93.1

As shown in Table 4-3, algorithm #2 had the largest F-score (97.7%) with the lowest FP, but highest FN. Algorithm #1 and #3 had the highest recall (100%), and algorithm #1 had the second highest F-score (96.8%) followed by algorithm #3 with the lowest F-score (93.1%). Since the penalty of FP was considered higher than FN, algorithm #2 was the highest ranked turn-identification algorithm among the three because it had the lowest FP, while still having a large TP (RC = 96.9%) to provide a large pool of cornering examples. The advantage of algorithm #2 was that the reduction of the GPS trace generally eliminated the small and gradual turns to reduce FP. Algorithm #1 and #3 tend to be more susceptible to the small and gradual curves. In general, all three algorithms had high PC (> 90%) and high RC (> 96%). When compared to the previous study by Wallace *et al.* [56], algorithm #2 had higher TP rate by about 23% (96% vs. 73%), lower FN rate, and similar FP rate from testing and training on about the same number of manually labeled turns (~200). Although algorithm #2 seemed to outperform the previous study, the difference in performance could also be contributed by the discrepancies in the testing methodology and the dissimilar datasets, which included different routes with different representative traffic scenarios.

Since the absolute GPS position was not available in the Candrive dataset to label the turns accurately, a true turn-identification performance cannot be assessed with the Candrive dataset. Instead, the three proposed turn-identification algorithms were evaluated for a single subject with a dataset created from the Candrive sensor, but the limitation on generalizability was from testing on only one participant, which can be improved in future studies by having more participants involved. To assess the generalizability on the Candrive dataset, algorithm #2 was tested on the Candrive dataset (as shown in Figure 5-1 in Chapter 5.2.1), and several randomly chosen trips across different drivers were manually inspected for the locations of the turning maneuvers, of which most of the turns (i.e., major corners and bends) were correctly identified and only a small proportion appeared to be misclassified. As long as the TP was much greater than the misclassification errors, the pool of examples of the turning maneuvers would be dominated by the true turns.

#### 4.4.1 Misclassification Errors

An example of a FP for algorithm #1, Figure 4-6, shows a sudden curve in the path that was caused by overtaking another vehicle. The  $\Delta h[n]$  within this curve was above the threshold and was categorized as a turn.



Figure 4-6. An example of FP in the GPS trace by algorithm #1 as indicated by the red marker on a map display with the trace of the route.

An example of a FP in Figure 4-7 for algorithm #2 shows that the reduced GPS trace altered the gentle curve into a prominent corner, which was recognized as a turn. An example of a FN in Figure 4-8 for algorithm #2 shows that the reduced GPS trace altered the corner, as shown by the single green marker and red star, such that the two turns in close vicinity was recognized as one turn.

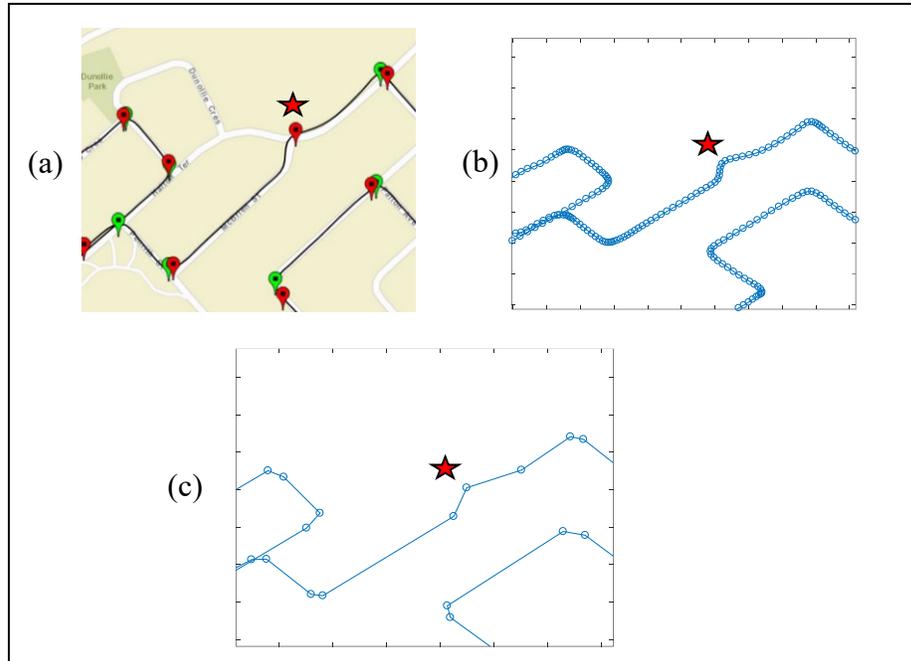


Figure 4-7. An example of a FP for algorithm #2 which was caused by a short curve in the road (red star) that was emphasized when reducing the GPS trace. (a) The map display shows the turns from the gold standard (green marker) and the identified turns by the algorithm (red marker). (b) The coordinates of the path. (c) The reduced GPS data of the same route.

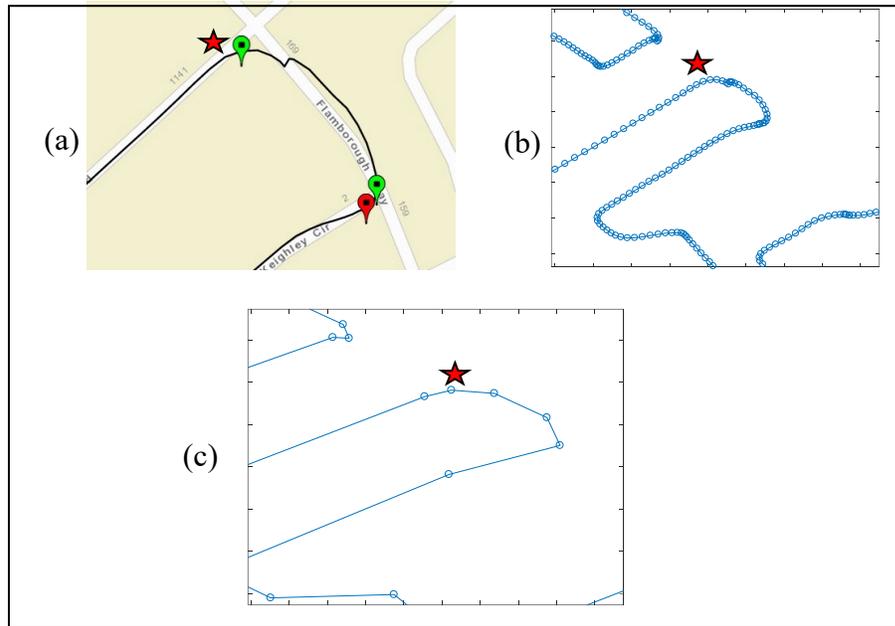


Figure 4-8. An example of a FN for algorithm #2 (red star). The turn (red star) was more gradual than the other turn (red marker), and the reduction of the GPS trace did not emphasize the corner. Since the two turns were in close vicinity, the turns were merged into one larger and sharper turn. (a) The map display shows the turns from the gold standard (green marker) and the identified turns by the algorithm (red marker). (b) The coordinates of the path. (c) The reduced GPS data of the same route.

An example of the FP for algorithm #3 in Figure 4-9 shows a short curve in the road where the  $R_c[n]$  within this curve was above the threshold and was recognized as a turn. This misclassification error was similar to the FP for algorithm #1 and #2.

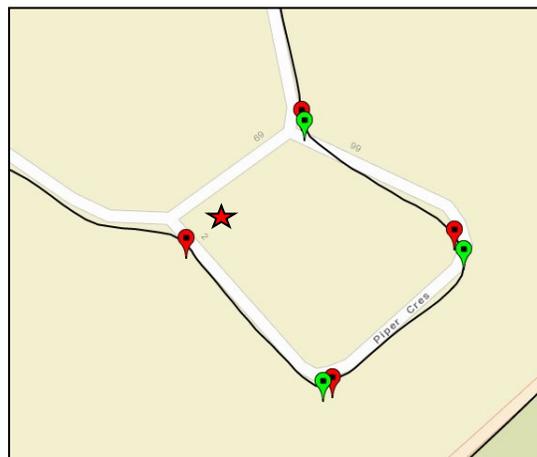


Figure 4-9. An example of a FP for algorithm #3 (red star), where a gentle curve was mistaken as a turn.

## 4.5 Conclusion

To identify turns in the GPS data, three algorithms were proposed and evaluated using six trips of independently collected data from the Candrive data. In general, all three algorithms had high PC ( $> 90\%$ ) and high RC ( $> 96\%$ ). Algorithm #2 had the best performance of the three because it had the highest F-score (97.7%) and the highest precision (97.9%). Algorithm #1 and #3 were more sensitive to small and gradual curves in the GPS trace, and algorithm #2 used the reduction of the GPS trace to remove many of these small or gradual curves before searching for turns. Algorithm #2 was implemented for the analysis of the turning maneuvers in the Candrive data to efficiently identify the examples of cornering with high accuracy for a large number of trips.

## 5 Chapter: Similarities and Differences in Maneuvering Behaviour

### 5.1 Introduction

Driving safety requires a proficiency in maneuvering the vehicle while obeying the traffic laws. Many basic driving maneuvers, such as slowing down the vehicle before a stop sign or negotiating a turn around a corner in the road, are encountered frequently in everyday driving and these maneuvers are influenced by individual preferences as well as the driving conditions. For example, one driver may prefer to brake abruptly before a stop sign for a thrilling ride, while another driver may prefer gradually braking for a gentler ride. These individual preferences could change or remain the same overtime depending on many factors including physical and cognitive health. Since older drivers' health can influence their driving behaviour (e.g., dementia) [34], changes in the drivers' physical or cognitive health may be reflected in how they operate their vehicle (e.g., maneuvering). As older drivers continue to age and experience deterioration in health, a change in their maneuvering behaviour could provide real evidence of at-risk behaviour and a decline in the ability to drive safety.

The purpose of this chapter was to explore the general similarities and differences in the maneuvering behaviour among the 12 stable-health older drivers and propose a method to measure the consistency of their maneuvering behaviour during their baseline year. Some similarities and difference among stable-health drivers were shown from their deceleration and acceleration events including different mean accelerations, and their acceleration and deceleration behaviours were modeled by a two-phase relationship by Wallace *et al.* [9]. By measuring the between-driver and within-driver variations, patterns that exist over the span of the baseline year can inform future studies on the typical

behaviour of each individual driver as well as the entire group. The consistency in their maneuvering behaviour can be described by the amount of variations between months and the yearlong variations during the entire year.

Four driving maneuvers were chosen for the analysis: acceleration from a stop, deceleration coming to a stop, right turn, and left turn. Since the GPS data were anonymized and the exact locations of specific traffic scenarios were unknown (e.g., stopping at a stop sign), these four driving maneuvers were chosen to represent different common traffic scenarios that occur abundantly in the data regardless of the differences in routes between drivers. The similarities and differences between these driving maneuvers can provide insight into their driving behaviour and provide a better understanding of how their maneuvering changes over time. Since these participants had a stable health condition during the baseline year, their maneuvering characteristics during the year provide examples of the natural variability in the driving behaviours of older drivers who do not have a decline in health. This chapter also examines the effects of environmental factors on their maneuvering behaviour: weather, season, traffic condition, and daylight. While there are many factors that affect driving behaviour, these four factors intuitively have large influence on driving behaviour, especially since the eastern province of Canada experiences four distinct seasons with a wide range of temperatures throughout the year. For example, a driver negotiating a turn in the summer on a sunny day would most likely turn at slower speeds on a snowy day in winter for the same turn to avoid the risk of MVC.

The first section defines the four driving maneuvers, their categories of events, and the driving conditions. Next, a list of steps for the analysis was presented to identify

the similarities and differences between drivers and measure the consistency during the baseline year. Afterwards, the results from the analysis of each type of driving maneuver were presented. Finally, the chapter concludes with a description of the key findings and a discussion on the utility of the driving maneuvers for long-term monitoring of driving behaviour.

## **5.2 Methodology**

For the analysis of the driving maneuvers, only the highest quality of the trips were used as described in Chapter 3.4, and each occurrence of a driving maneuver was called an *event*.

### **5.2.1 Driving Maneuvers**

The driving maneuvers under investigation were the ones that occurred abundantly in the GPS data regardless of the actual location, destination, or route taken by the driver. The maneuvers that frequently occurred while driving were accelerating, decelerating, and turning (i.e., cornering). Figure 5-1 shows a trip taken by a driver, and the highlighted portions of the trip indicated were the acceleration and deceleration events occurred as defined by Wallace *et al.* [9], including the identified turns using the algorithm from Chapter 4.2.3.2.

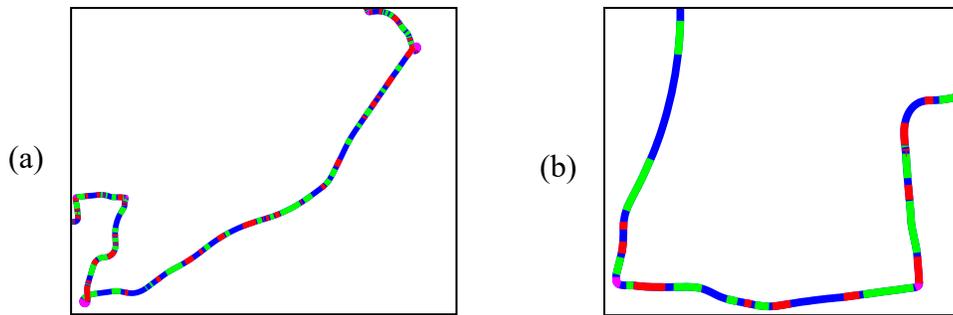


Figure 5-1. An example of a trip taken by an older driver showing the occurrence of common driving maneuvers during the trip, namely acceleration, deceleration, and turns. The acceleration and deceleration events were defined in [9] had changes in speed of  $\geq 4$  km/h. The vehicle was at steady speeds (blue); the vehicle was accelerating (green); the vehicle was decelerating (red); and the vehicle was turning (magenta). (a) It was evident that there were two expressway ramps (loops), which indicate that the long, straight path in the middle of the trip was the expressway and the other sections were lower speed roadways, which could be an example of the driver's trip between home and place of employment. (b) The magnified view of the lower speed roadways, which was characterized by more corners and bends than the expressway. The start and end destinations of this trip was excluded to preserve anonymity.

Since there can be many acceleration, deceleration, and turning events during a trip (Figure 5-1), a subset of these events were selected (Figure 5-2) to concentrate on specific traffic scenarios and lower the variability among events. Thus, the driving maneuvers were limited to specific subsets of events that involved a substantial change in speed or substantial change in heading to more accurately reflect the drivers' preference rather than responses to the surrounding traffic.

### 5.2.2 Definition of the Four Driving Maneuvers

The four maneuvers were chosen to be mutual exclusive to one another where each maneuver described a different portion of the drivers' trips. The four maneuvers are defined in Table 5-1.

The acceleration events (ACC) were limited to only straight paths, also known as *straightaways*, to focus on the driver's preferred changes in speed and less on steering; beginning from a stop (i.e.,  $< 10$  km/h to include rolling stops) to target the scenarios of

coming out of traffic stops (e.g., stop signs and traffic lights); continuously increasing in speed for at least 2 seconds to focus on the continuous act of pressing the gas pedal (i.e., accelerating the vehicle); and the final speed of the entire event must be at least 30 km/h to focus on the substantial increase in speed, which were more influenced by the driver's preferences rather than the minor speed adjustments in response to the surrounding traffic (from Chapter 2.9.1). An example of an acceleration event is shown in Figure 5-3a. Similarly, the deceleration events (DEC) were limited to straightaways; coming to a stop to target the scenarios of approaching traffic stops; continuously decreasing in speed for at least 2 seconds to focus on the continuous act of pressing the brake pedal; and the starting speed of the entire event must be at least 30 km/h to focus on the substantial decrease in speed.

The turns were divided into two types, right and left turns, to represent the different traffic scenarios that occurred between left and right turns and to incorporate the driver's steering behaviour while the vehicle was accelerating or decelerating. The right turn events (RT) were portions of the trip where the vehicle was travelling in a curved path; without stopping during the turn to avoid scenarios that were interrupted by oncoming traffic or pedestrians; continuously changing speeds consistently rightwards; duration must be at least 3 seconds to avoid short maneuvers such as changing lanes; and the vehicle must be travelling on a straightaway 5 seconds before and after the turn to filter out sporadic maneuvering. An example of a right turn is shown in Figure 5-3b. Similarly, the left turn events (LT) were portions of the trip where the vehicle was travelling in a curved path; without stops during the turn; continuously changing speeds



Table 5-1. Definitions of the four driving maneuvers

<b>Maneuver</b>	<b>Criteria</b>	<b>Definition</b>
Acceleration (ACC)	Duration was at least two seconds	$N \geq 2$ <i>for</i> $n = k, k + 1, \dots, k + N$
	Occurred on straightaways	$\Delta h[n] < 6^\circ$
	Starting at rest	$V_{GPS}[k] < 10 \text{ km/h}$
	Continuously increasing speeds	$a_t[n] > 0$
	Substantial change in speed	$V_{GPS}[k + N] \geq 30 \text{ km/h}$
Deceleration (DEC)	Duration was at least two seconds	$N \geq 2$
	Occurred on straightaways	$\Delta h[n] < 6^\circ$
	Ending at rest	$V_{GPS}[k + N] < 10 \text{ km/h}$
	Continuously decreasing speeds	$a_t[n] < 0$
	Substantial change in speed	$V_{GPS}[k] \geq 30 \text{ km/h}$
Right Turn (RT)	Satisfied Criteria of a Turn (Chapter 4)	Table 4-1
	Uninterrupted	$V_{GPS}[n] \neq 0 \text{ km/h}$
	Turn radius was between 4 and 70 m (see Chapter 5.2.3)	$4 \text{ m} < r_T < 70 \text{ m}$
	Travelling on straightaway 5 seconds before and after the turn.	$\Delta h[n] < 6^\circ$ AND $V_{GPS}[n] \neq 0 \text{ km/h}$ <i>for</i> $n = \{N, N + 1, \dots, N + 5\},$ $\{N, N - 1, \dots, N - 5\}$
Left Turn (LT)	Satisfied Criteria of a Turn (Chapter 4)	Table 4-1
	Uninterrupted	$V_{GPS}[n] \neq 0 \text{ km/h}$
	Turn radius was between 4 and 70 m (see Chapter 5.2.3)	$4 \text{ m} < r_T < 70 \text{ m}$
	Travelling on straightaway 5 seconds before and after the turn.	$\Delta h[n] < 6^\circ$ AND $V_{GPS}[n] \neq 0 \text{ km/h}$ <i>for</i> $n = \{N, N + 1, \dots, N + 5\},$ $\{N, N - 1, \dots, N - 5\}$

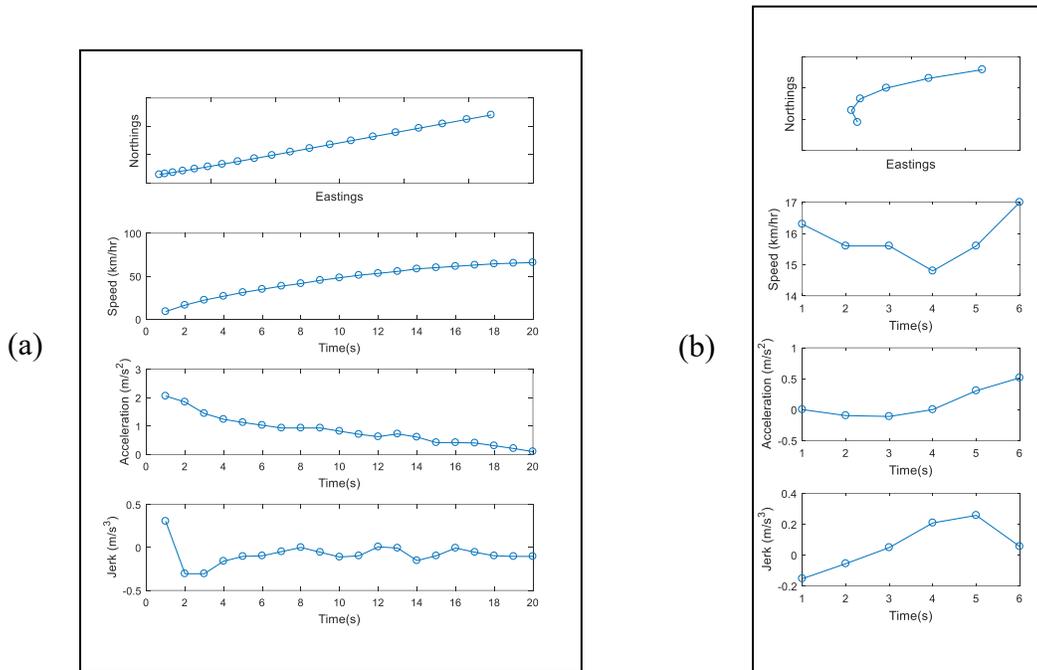


Figure 5-3. (a) An example of ACC event and (b) RT event showing the discrete-time signal of the speed, acceleration, and jerk values with linear interpolation for visualization.

### 5.2.3 Derived Variables

Each turn can be described in more detail by the radius and angle of the turn. The radius of the turn was the distance between the edge and center of a circle that approximately fitted the coordinates of the turn and described the size of the turn. Since a turn was an arc of coordinates, a unique circle approximately fitted the turn by using least-squares fitting of a circle to a series of coordinates around an arc [78]. Once each turn was approximated by a circle, the turn radius  $r_T$  can be found to calculate the instantaneous centripetal acceleration  $a_c[n]$  (i.e., lateral acceleration) using Eq. 5-1, which describes the aggressiveness of the cornering. Approaching a tighter turn at higher speeds would yield a larger centripetal acceleration, which indicated more aggressive cornering. The instantaneous net acceleration  $a_N[n]$  (Eq. 5-2) described the total magnitude of the acceleration during the turn.

$$a_c[n] = \frac{V_{GPS}[n]^2}{r_T[n]} \quad \text{Eq. 5-1}$$

$$a_N[n] = \sqrt{a_c[n]^2 + a_t[n]^2} \quad \text{Eq. 5-2}$$

The turn angle described the sharpness of the turn as shown in Figure 5-4. For example, a lower turn angle indicated a more gradual turn that involved less steering. The turn angle ( $\theta_T$ ) was the central angle created between the two radii that intersected at the center of the circle ( $h, z$ ) from the first (i.e., enter vector) and last coordinates (i.e., exit vector) of the turn to the center, using Eq. 5-3 for left turns and similarly for right turns but with the positive and negative conditions swapped.

$$\theta_T = \begin{cases} \cos^{-1} \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|}, & \text{sign } (\vec{A} \times \vec{B})_z = \text{positive} \\ 360^\circ - \cos^{-1} \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|}, & \text{sign } (\vec{A} \times \vec{B})_z = \text{negative} \end{cases} \quad \text{Eq. 5-3}$$

Where,

$$\vec{A} = (x[k] - h, y[k] - z)$$

$$\vec{B} = (x[k + N] - h, y[k + N] - z)$$

Figure 5-4 shows an example of a u-turn, which represented the sharpest possible turn occurring in the GPS data, an example of a gradual turn on the road, and an example of an on-ramp for the expressway.

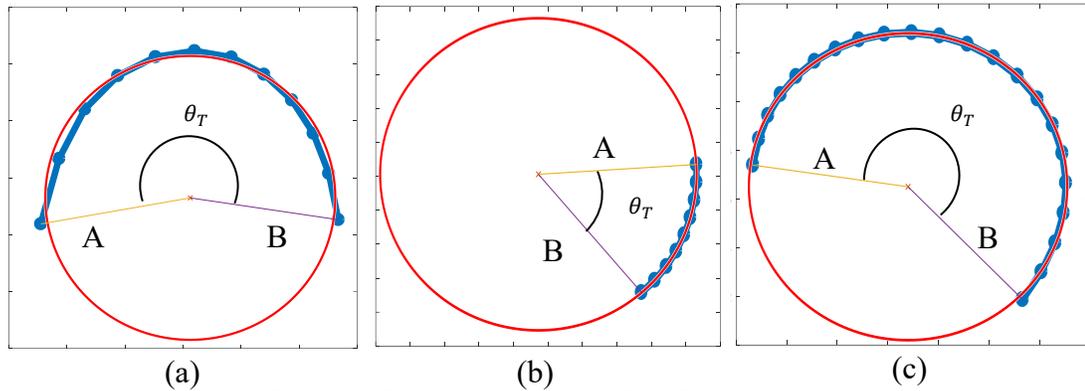


Figure 5-4. Example of three different turns identified from the GPS data. The path of the vehicle was shown by the blue GPS trace and the red circle approximately fitted the arc of the turn. The yellow vector A and purple vector B extended from the center to the edge of the circle and represented the start and ends of the turn respectively. (a) An example of a sharp turn from performing a u-turn with  $r_T = 4$  m and  $\theta_T > 180^\circ$ . (b) A large, gradual turn with  $r_T = 92$  m and  $\theta_T = 53^\circ$ . (c) An example of an on-ramp from the expressway with  $r_T = 62$  m and  $\theta_T = 217^\circ$ .

From the examples of turns in Figure 5-4, the large and gradual turns, such as Figure 5-4b, were filtered out by restricting the size of turns to between 4 m and 70 m to focus on the turns that required more steering, as defined in Table 5-1. By using Google Maps<sup>4</sup> distance measuring tool, a typical expressway ramp was 65 m and a typical residential turn was 12 m, which showed that the restriction on the turn radius was reasonable.

#### 5.2.4 Categories of Driving Events

The driving maneuvers were separated into different categories to focus on similar traffic conditions at similar speeds. For example, stopping at a stop sign would be more common on lower speed roadways (i.e., 40 km/h to 60 km/h), and a stop sign creates a queue for drivers to brake and accelerate one vehicle at a time. The 70 to 80 km/h roadways represented arterial roads which tend to carry more traffic at higher speeds and have fewer stop signs; thus the driving maneuvers would be more likely influenced by surrounding traffic, such as coming out of a red light behind another

<sup>4</sup> For more information, please visit [www.google.com/maps/](http://www.google.com/maps/)

vehicle. The right and left turns were categorized by the speed limits and turn angles. The speed limits described the size of the turns (e.g., turns on expressways were most likely on and off ramps) and the turn angles determined the amount of steering involved during the turn. The 40 to 60 km/h roadways represented residential and city driving, 70 to 80 km/h represented arterial roads (i.e., main streets) with higher speed limits, and 100 km/h for the expressways. The turns were further categorized by the turn angle: acute angle ( $0^\circ < \theta_T \leq 90^\circ$ ) represent gradual turns (i.e., smaller change in direction); obtuse angle ( $90^\circ < \theta_T \leq 180^\circ$ ) represent sharp turns (i.e., larger change in direction); and reflex angle ( $180^\circ < \theta_T$ ) represent u-turns and loops.

Table 5-2. Summary of the categories of driving events

Driving Maneuvers	Criteria (Speeds/Angles)		Scenarios
Acceleration (ACC)	40 to 60 km/h		Residential and city driving
	70 to 80 km/h		Arterial Roads
	100 km/h		Expressway
Deceleration (DEC)	40 to 60 km/h		Residential and city driving
	70 to 80 km/h		Arterial Roads
	100 km/h		Expressway
Right Turns (RT)	40 to 60 km/h	Acute Angle	Gradual corners
		Obtuse Angle	Sharp corners
		Reflex Angle	U-Turns and loops
	70 to 80 km/h	Acute Angle	Gradual corners
		Obtuse Angle	Sharp corners
		Reflex Angle	U-Turns and loops
	100 km/h	Acute Angle	Gradual ramps
		Obtuse Angle	Sharp ramps
		Reflex Angle	Loop ramps
Left Turns (LT)	40 to 60 km/h	Acute Angle	Gradual corners
		Obtuse Angle	Sharp corners
		Reflex Angle	U-Turns and loops
	70 to 80 km/h	Acute Angle	Gradual corners
		Obtuse Angle	Sharp corners
		Reflex Angle	U-Turns and loops
	100 km/h	Acute Angle	Gradual ramps
		Obtuse Angle	Sharp ramps
		Reflex Angle	loop ramps

From Table 5-2, there were a total of 24 categories of driving events based on the 4 types of driving maneuvers. Although there were 24 categories of events, the events that occurred on high speed roadways were much fewer than the events that occurred in the lower speed roadways as shown in Figure 5-5.

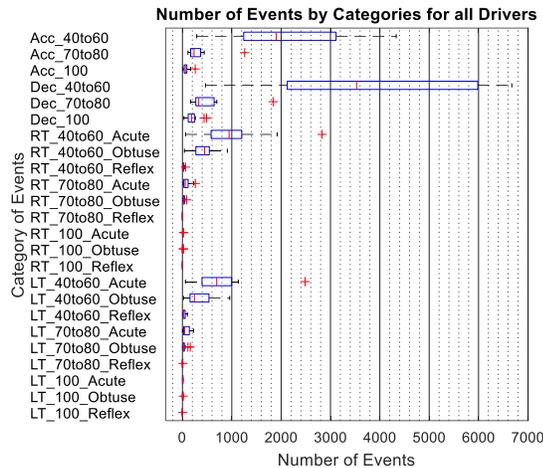


Figure 5-5. The distribution of the total number of each category of events for all drivers. The red lines represented the median value, the boxes marked the width of the IQR, and the + symbols represented the drivers who had a much larger number of events. The events that occurred on the lower speed roadways were more abundant than the other events.

Since the number of turns that occurred on the low-speed roadways was much greater than the other categories for all drivers and that the maneuvers were more likely influenced by traffic on higher speed roadways, the events were limited to 40 to 60 km/h speed limits to focus on this group of similar traffic scenarios. When compared to low speed roadways, there were far fewer cases of ACC and DEC events on the arterial roads and expressways since there were much fewer mandatory traffic stops, and older drivers may be deliberately avoiding the expressway. The turns with reflex angles (e.g. expressway ramps) occurred much less frequently than the obtuse and acute angle turns. By focusing on low-speed maneuvers, a more homogenous pool of examples was created to examine individual preferences during common and similar traffic scenarios.

### 5.2.5 Definition of the Driving Conditions

Each of the four environmental factors was divided into two polar opposite conditions to represent favourable and adverse driving conditions. The first factor is weather in which the conditions were clement weather and inclement weather. The second factor was seasons in which the conditions were summer (i.e., the warmest months of the year) and winter (i.e., the coldest months of the year). The third factor was the amount of daylight, daytime and nighttime, which described the level of visibility. The fourth factor was traffic conditions, which was described by off-peak and peak hours (i.e., rushhour).

Table 5-3. The criteria of the favourable and adverse driving conditions for each environmental factor

Factor	Favourable Condition	Adverse Condition
Weather	<u>Clement:</u> (Clear OR Mainly Clear OR Mostly Clear) AND (Temperature > 10°C)	<u>Inclement:</u> (Daily Total Precipitation > 10 mm) OR (Snow on the ground > 5 cm ) OR (Visibility < 10 km) OR (Temperature < -10°C)
Season	<u>Summer:</u> June, July, August	<u>Winter:</u> December, January, February
Daylight	<u>Daytime:</u> Between 30 min after sunrise and 30 min before sunset	<u>Nighttime:</u> Between 30 min after sunset and 30 min before sunrise
Traffic	<u>Off-peak times:</u> 9 a.m. to 4 p.m. EST 6 p.m. to 7 a.m. EST	<u>Peak times:</u> 7 a.m. to 9 a.m. EST 4 p.m. to 6 p.m. EST

### 5.2.6 Analysis of the Events

Each event can be characterized by its acceleration to focus on how the driver operated the vehicle and less on the chosen route. The acceleration described how the driver controlled the speed of the vehicle regardless of the distance travelled. The mean tangential acceleration ( $\bar{a}_t$ ) summarized the acceleration during each ACC and DEC

event and the mean net acceleration ( $\bar{a}_N$ ) for each acute and obtuse RT and LT event. The analysis steps are listed below:

- 1) Identify the differences between the drivers by their events during the baseline year including the number of significantly different pairs from the 66 possible pairs of drivers using ANOVA with Tukey's HSD ( $\alpha = .05$ ).
- 2) Represent the drivers' typical monthly behaviour by the median  $\bar{a}_t$  and  $\bar{a}_N$  of the events across each month, and then calculate the yearlong trend by using a linear regression model ( $Y = \beta_0 + \beta_1x + \varepsilon$ ) with iteratively reweighted least squares to account for outliers (i.e., unusual month).
- 3) Measure the consistency among the drivers during the year by calculating the intraclass correlation coefficient (ICC) (Eq. 5-4) using the mean squared error between-group (MSB) and mean squared error within-group (MSW) [79]. An ICC closer to 1 indicated high degree of consistency while an ICC closer to 0 indicated low degree of consistency between drivers.

$$ICC = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} = \frac{MSB - MSW}{MSB + (k - 1)MSW} \quad \text{Eq. 5-4}$$

- 4) Plot the absolute value of the yearlong trend  $|\beta_1|$  vs. the IQR of the residuals from step 2 to visualize the consistency during the year per driver.
- 5) Since a characteristic of the drivers was their annual mileage, examine the relationship between the averages of the median  $\bar{a}_t$  and  $\bar{a}_N$  (i.e., monthly behaviour) across the entire year and the annual mileage of the drivers using correlation coefficient (R).
- 6) Examine the differences between the accelerations during favourable and adverse driving conditions by comparing the mean  $\bar{a}_t$  and  $\bar{a}_N$  of the events

including a test for the hypothesis that drivers have lower mean accelerations in adverse conditions (and in favourable conditions for deceleration events) using a one-tailed, two sample t-test ( $\alpha = .05$ ) with the following hypotheses (opposite for deceleration events):

$$H_0: \mu_{favourable} = \mu_{adverse}$$

$$H_A: \mu_{favourable} > \mu_{adverse}$$

### 5.3 Results and Discussion

#### 5.3.1 Acceleration and Deceleration Events

The distribution of  $\bar{a}_t$  for ACC and DEC events for each driver during their baseline year is shown in Figure 5-6 and the corresponding table of values in Table 5-4 and Table 5-5. There were 54 of 66 (82%) and 61 of 66 (92%) significantly different pairs of drivers for ACC and DEC events respectively. Drivers 1, 2, and 3 had similar mean acceleration for ACC, but different mean deceleration for DEC. Driver 4 was among the highest annual mileage drivers, but had the fewest maneuvering events because this driver had many trips with unavailable speed limit data as indicated by  $\bar{V} = 0$  km/h. Driver 4 had the highest mean acceleration for ACC, while Drivers 8 had the highest mean deceleration for DEC during the year, which may indicate that these were aggressive or high confidence drivers. Driver 10 had the lowest mean acceleration and deceleration, which may indicate that this driver was a cautious or low confidence driver.

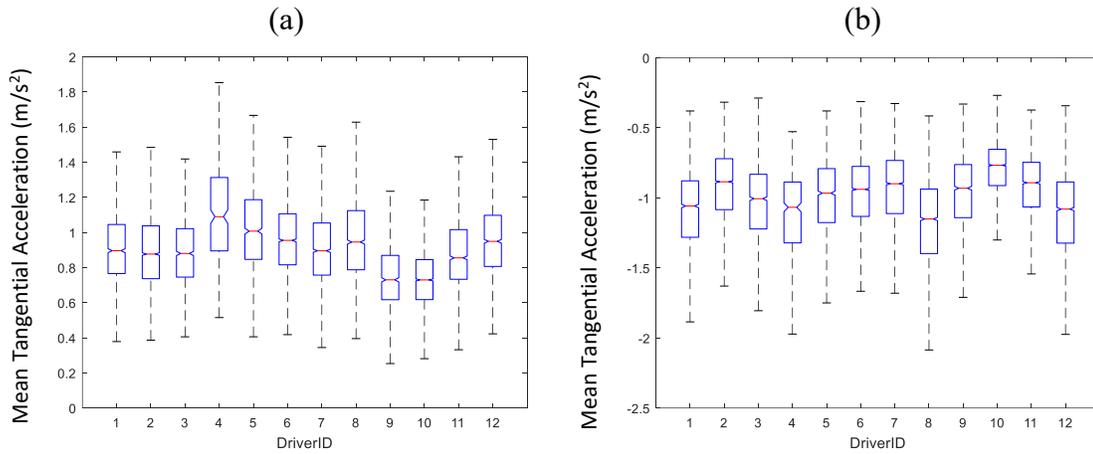


Figure 5-6. Comparison of the distributions of mean tangential acceleration ( $\bar{a}_t$ ) for ACC events (a) and DEC events (b) per driver across the entire baseline year. The red lines represented the median value and the boxes marked the width of the IQR. The outliers were removed by setting the lower limit to  $Q_1 - 1.5(IQR)$  and upper limit to  $Q_3 + 1.5(IQR)$ . By removing the outliers, the distribution was approximated by a normal distribution  $N(\mu, \sigma^2)$ .

Table 5-4. Summary of the mean  $\bar{a}_t$  for ACC events per driver over the entire baseline year including the group averages.

DriverID	N	$\mu$ (m/s <sup>2</sup> )	95% CI		$\sigma$ (m/s <sup>2</sup> )	IQR (m/s <sup>2</sup> )
1	1567	0.902	0.892	0.912	0.199	0.279
2	4279	0.890	0.884	0.896	0.211	0.301
3	1286	0.887	0.877	0.898	0.197	0.276
4	288	1.114	1.082	1.146	0.279	0.418
5	1026	1.011	0.997	1.025	0.233	0.340
6	3335	0.965	0.958	0.972	0.205	0.291
7	2802	0.903	0.896	0.911	0.210	0.297
8	2107	0.959	0.949	0.969	0.236	0.337
9	1164	0.741	0.731	0.751	0.180	0.252
10	3791	0.730	0.725	0.736	0.168	0.228
11	1834	0.872	0.863	0.882	0.205	0.283
12	1915	0.957	0.948	0.967	0.206	0.291
Group	25394	0.888	0.885	0.891	0.222	0.300

Table 5-5. Summary of the mean  $\bar{a}_t$  for DEC events per driver over the entire baseline year including the group averages.

DriverID	N	$\mu$ (m/s <sup>2</sup> )	95% CI		$\sigma$ (m/s <sup>2</sup> )	IQR (m/s <sup>2</sup> )
1	2746	-1.082	-1.092	-1.071	0.289	0.403
2	6323	-0.905	-0.911	-0.899	0.247	0.364
3	2110	-1.025	-1.037	-1.013	0.273	0.390
4	454	-1.106	-1.133	-1.078	0.297	0.434
5	1569	-0.986	-0.999	-0.973	0.267	0.385
6	6121	-0.952	-0.959	-0.946	0.251	0.357
7	5430	-0.922	-0.929	-0.915	0.262	0.379
8	3866	-1.167	-1.177	-1.157	0.321	0.460
9	2041	-0.955	-0.967	-0.944	0.265	0.379
10	6423	-0.779	-0.784	-0.775	0.181	0.259
11	3292	-0.907	-0.914	-0.899	0.220	0.319
12	3610	-1.108	-1.118	-1.098	0.299	0.437
Group:	43985	-0.959	-0.962	-0.957	0.281	0.380

Figure 5-7 shows the typical monthly acceleration behaviour from the ACC and DEC events per driver during their respective baseline year. The ICC of 0.752 and 0.846 for ACC and DEC events respectively indicated that there was a high degree of consistency among the drivers showing that individual preferences were distinct and consistent throughout the year. It was evident that a few drivers had missing data for some months, such as Driver 4 during July 2010, because they did not drive during those months most likely from taking vacations.

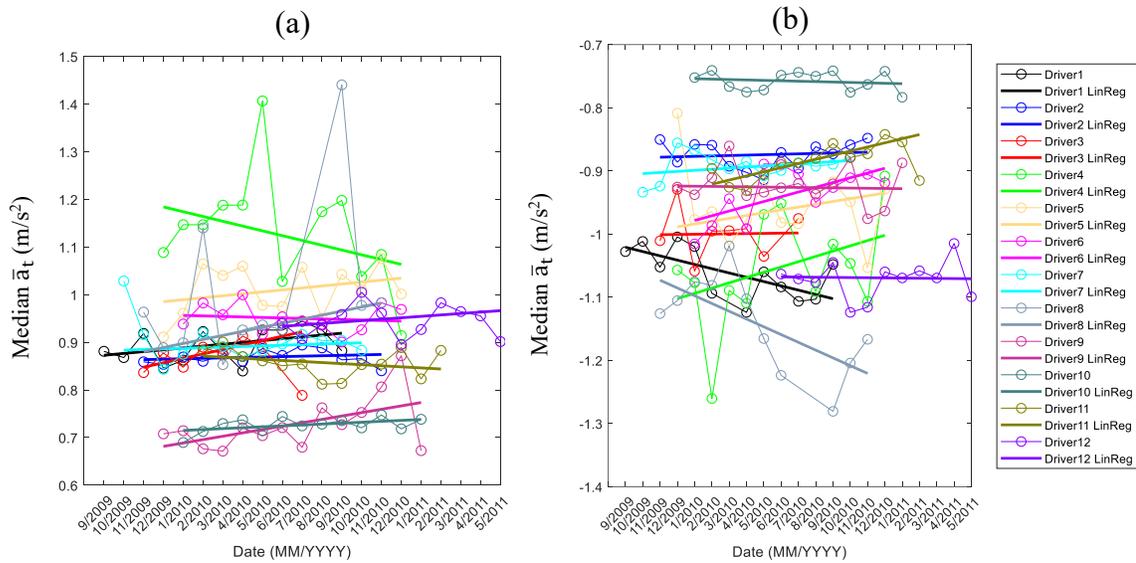


Figure 5-7. Comparison of the typical monthly behaviours from ACC (a) and DEC (b) events for each driver during their baseline year including the yearlong linear trend. Each month was represented by the median of the mean tangential accelerations ( $\bar{a}_t$ ) across all the events during each month and the linear interpolation imputed the missing in-between months.

From Figure 5-7, Driver 4 and Driver 8 had the largest variations and the large peaks, specifically during May and September 2010 respectively, corresponded to months with relatively low ( $<10$ ) number of events, which was attributed to driving less during those months. Drivers with higher average acceleration per month had larger variations during the year ( $R = 0.897$ ) as shown in Figure 5-8a. Figure 5-8b provided a method to measure and visualize the consistency in the maneuvering behaviour per driver by plotting the variations between months and the yearlong trend. Driver 1, 2, 6, 7, 10, 11, and 12 had relatively steady and similar behaviour during the year having low variations between around  $0.02$  to  $0.06$   $m/s^2$  between months and  $0$  to  $0.004$   $m/s^2$  for the yearlong trend  $|\beta_1|$ , whereas Driver 3, 4, 8, and 9 had greater variations during the year. Driver 5 had an overall steady behaviour during the year, but also distinct behaviours during different months as shown by the high variability between months and low change during the year.

Similar to ACC events, Driver 4 and Driver 8 had the largest variations in which the large peaks corresponded to months with relatively low ( $< 10$ ) number of events. Drivers with higher average deceleration per month had larger variations during the year ( $R = -0.944$ ) in Figure 5-9a. Driver 2, 3, 7, 9, 10, 12 formed a distinct cluster that had relatively steady behaviour during the year with low variations having between 0.02 to 0.04  $m/s^2$  between months and 0 to 0.002  $m/s^2$  for the yearlong trend  $|\beta_1|$  in Figure 5-9b, whereas Driver 1, 4, 5, 6, 8, and 11 had greater variations during the year.

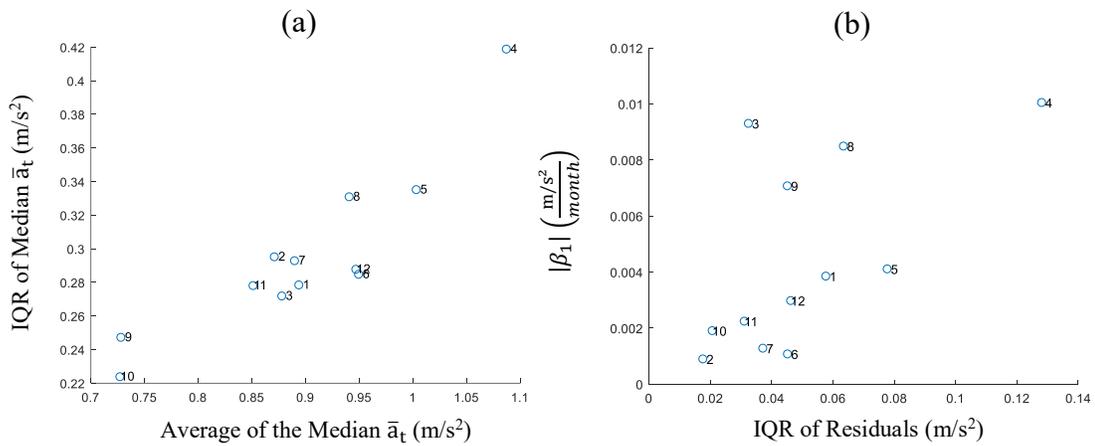


Figure 5-8. (a) The relationship between the IQR of the median  $\bar{a}_t$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_t$  across the entire year,  $R = 0.897$  ( $p < .001$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_t$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_t$ ,  $R = 0.603$  ( $p = .038$ ), for ACC events.

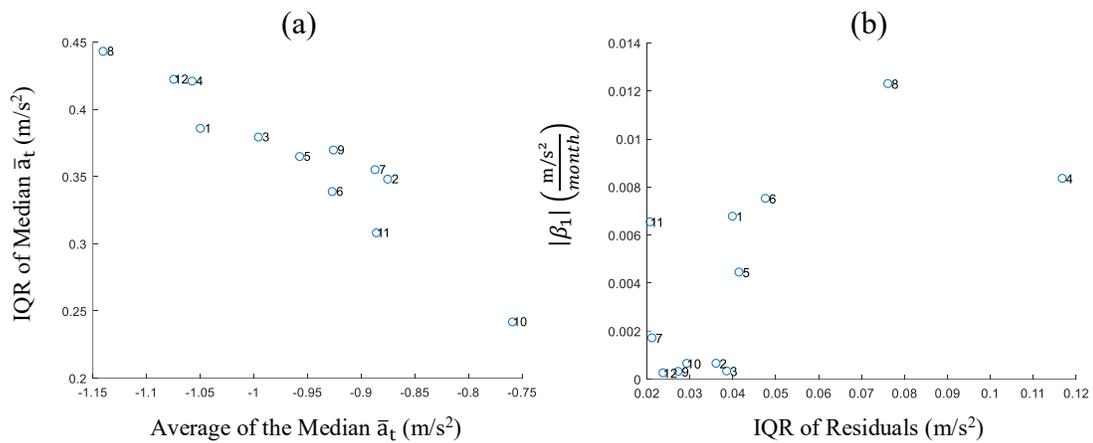


Figure 5-9. (a) The relationship between the IQR of the median  $\bar{a}_t$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_t$  across the entire year,  $R = -0.944$  ( $p < .001$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_t$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_t$ ,  $R = 0.654$  ( $p = .021$ ), for DEC events.

The annual mileage of the drivers did not have a strong relationship with the acceleration behaviour as shown in Figure 5-10. Most of the drivers had similar behaviours as seen by the distinct cluster except driver 9 and 10 who had relatively low average acceleration and Driver 4 who had much higher average acceleration and greater variations. The annual mileage of the drivers did not have a clear relationship with the deceleration behaviour as shown in Figure 5-11, where most of the drivers had similar behaviours as seen by the distinct cluster except Driver 10 who had relatively low average deceleration and Driver 4 had much higher variation.

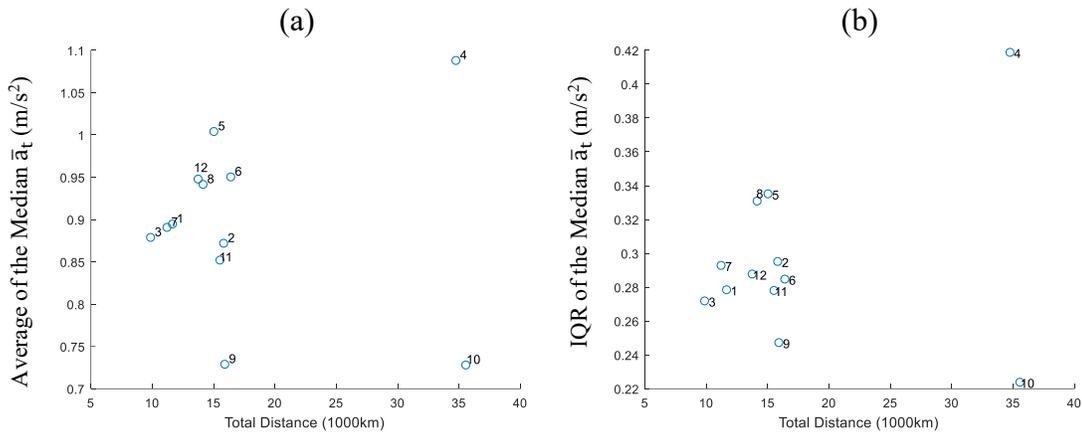


Figure 5-10. (a) The relationship between the average of the median  $\bar{a}_t$  (i.e., typical monthly accelerations) across the entire year and the annual mileage,  $R = 0.017$  ( $p = .956$ ), and (b) the relationship between the IQR of the median  $\bar{a}_t$  across the entire year and the annual mileage,  $R = 0.230$  ( $p = .472$ ), for ACC events.

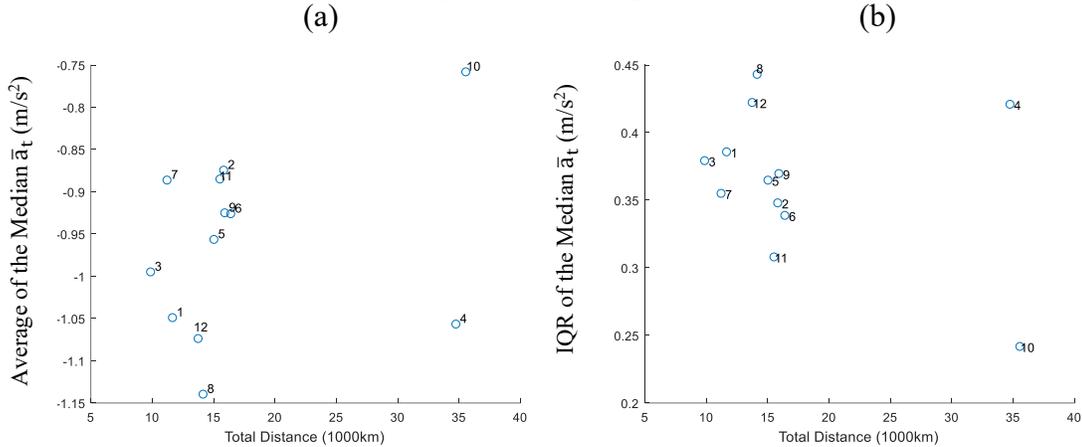


Figure 5-11. (a) The relationship between the average of the median  $\bar{a}_t$  (i.e., typical monthly accelerations) across the entire year and the annual mileage,  $R = 0.267$  ( $p = .402$ ), and (b) the relationship between the IQR of the median  $\bar{a}_t$  across the entire year and the annual mileage,  $R = -0.341$  ( $p = .278$ ), for DEC events.

The results from testing the mean  $\bar{a}_t$  for ACC and DEC events in different driving conditions for weather, season, traffic conditions, and daylight is shown in Appendix B.1 and Appendix B.2 respectively and summarized in Table 5-6. Although many of the hypothesis tests did not have enough evidence to reject the null hypothesis, there was evidence from the difference in mean accelerations that most drivers exhibited lower accelerations during adverse conditions. Daylight had the greatest impact on the

acceleration behaviour followed by the season based on the number of drivers who exhibited lower mean accelerations.

Table 5-6. Summary of the comparisons between the mean  $\bar{a}_t$  for ACC and DEC during favourable and adverse conditions per driver. From the group difference between the means ( $\Delta = \mu_{\text{adverse}} - \mu_{\text{favourable}}$ ), a negative value for ACC events indicated that the mean acceleration was lower during adverse condition, whereas a positive value for DEC events indicated a lower deceleration during adverse conditions.

	Weather		Seasons		Traffic		Daylight	
	ACC	DEC	ACC	DEC	ACC	DEC	ACC	DEC
Number of drivers showing statistical significance	1/12	2/12	5/12	1/12	2/12	0/12	5/12	0/12
Number of drivers with lower mean $ \bar{a}_t $ in adverse conditions	9/12	7/12	9/12	8/12	7/12	8/12	9/12	12/12
Average group difference of mean $\bar{a}_t$ (m/s <sup>2</sup> )	-0.015	0.014	-0.015	0.012	-0.004	0.018	-0.021	0.076

### 5.3.2 Right Turns Events

The distribution of  $\bar{a}_N$  from acute and obtuse RT events for each driver is shown in Figure 5-12 and the corresponding table of values in Table 5-7 and Table 5-8. There were 56/66 (85%) and 51/66 (77%) significantly different pairs of drivers for acute and obtuse RT events respectively. The mean  $\bar{a}_N$  was generally higher for obtuse turns (i.e., sharper turns) than the acute turns (i.e., gradual turns) because a sharper turn had larger centripetal acceleration component  $a_c[n]$  while the speeds between the acute and obtuse turns were generally similar.

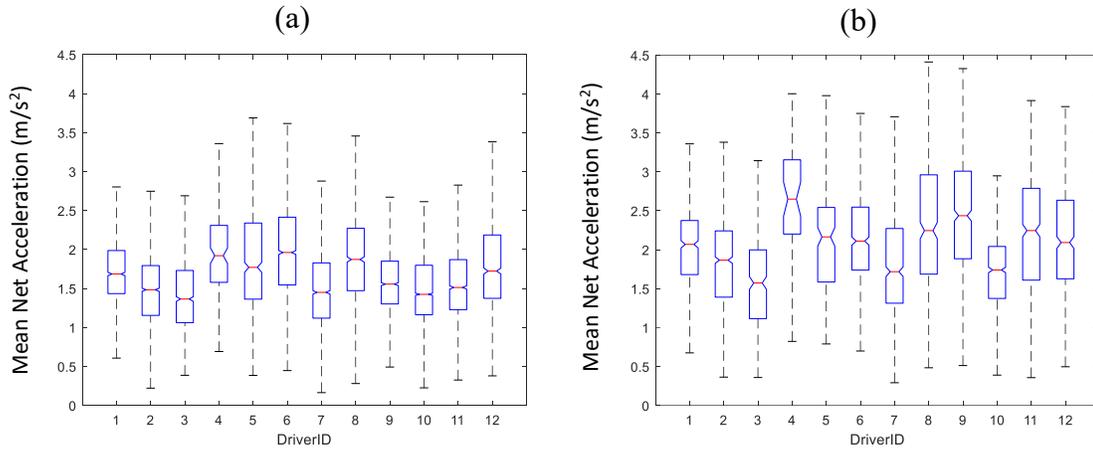


Figure 5-12. Comparison of the distributions of mean net acceleration ( $\bar{a}_N$ ) for acute RT events (a) and obtuse RT events (b) per driver for the entire baseline year. The red lines represented the median value and the boxes marked the width of the IQR. The outliers were removed by setting the lower limit to  $Q_1 - 1.5(IQR)$  and upper limit to  $Q_3 + 1.5(IQR)$ . By removing the outliers, the distribution was approximated by a normal distribution  $N(\mu, \sigma^2)$ .

Table 5-7. Summary of the mean  $\bar{a}_N$  for acute RT events per driver over the entire baseline year including the group averages.

DriverID	N	$\mu$ (m/s <sup>2</sup> )	95% CI		$\sigma$ (m/s <sup>2</sup> )	IQR (m/s <sup>2</sup> )
1	1863	1.708	1.690	1.727	0.416	0.553
2	2820	1.471	1.453	1.489	0.484	0.639
3	1198	1.409	1.383	1.436	0.467	0.669
4	126	1.932	1.832	2.031	0.564	0.730
5	583	1.863	1.807	1.919	0.690	0.977
6	3021	1.992	1.971	2.012	0.579	0.869
7	2540	1.474	1.453	1.494	0.520	0.709
8	1390	1.881	1.852	1.911	0.561	0.800
9	2122	1.562	1.544	1.579	0.414	0.549
10	5221	1.466	1.454	1.477	0.415	0.635
11	2091	1.534	1.512	1.555	0.495	0.642
12	1077	1.796	1.759	1.833	0.623	0.813
Group	24052	1.614	1.608	1.621	0.532	0.696

Table 5-8. Summary of the mean  $\bar{a}_N$  for obtuse RT events per driver over the entire baseline year including the group averages.

DriverID	N	$\mu$ (m/s <sup>2</sup> )	95% CI		$\sigma$ (m/s <sup>2</sup> )	IQR (m/s <sup>2</sup> )
1	559	2.017	1.973	2.061	0.527	0.695
2	1064	1.811	1.776	1.847	0.595	0.852
3	272	1.577	1.508	1.646	0.579	0.883
4	47	2.581	2.350	2.811	0.784	0.954
5	158	2.158	2.052	2.265	0.679	0.958
6	590	2.161	2.112	2.209	0.601	0.809
7	595	1.806	1.750	1.862	0.696	0.960
8	362	2.306	2.221	2.391	0.826	1.275
9	638	2.427	2.372	2.483	0.710	1.125
10	930	1.681	1.651	1.710	0.462	0.668
11	543	2.169	2.103	2.236	0.785	1.179
12	428	2.113	2.049	2.177	0.674	1.009
Group	6186	1.992	1.975	2.009	0.688	0.908

Figure 5-13 shows the typical monthly acceleration behaviour from the acute and obtuse RT events per driver during their respective baseline year. The ICC of 0.612 and 0.458 for acute and obtuse RT respectively indicated that there was moderate degree of consistency during the year, but less than ACC and DEC events.

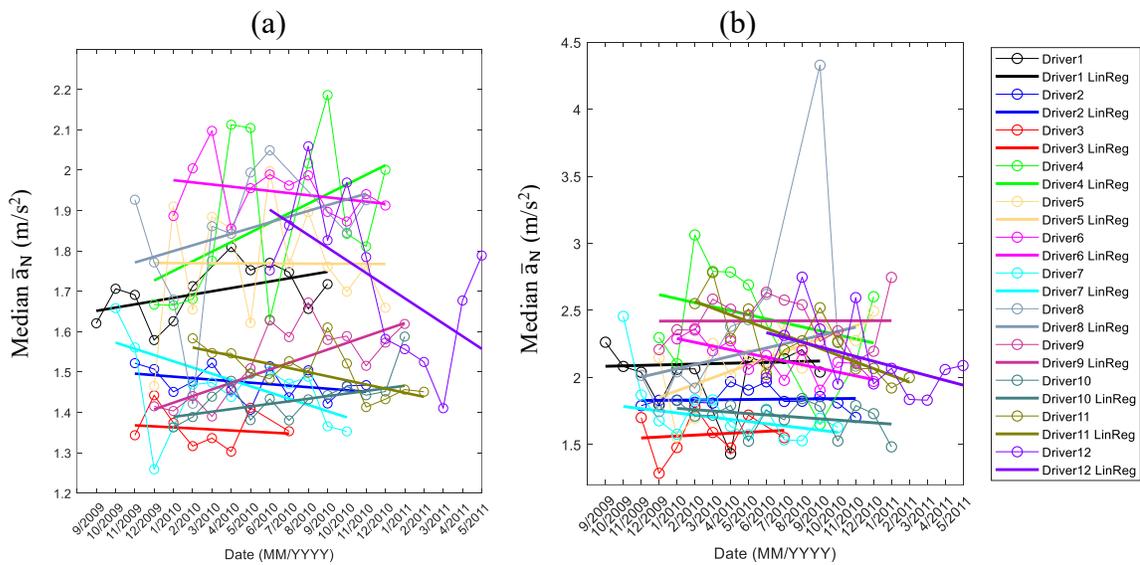


Figure 5-13. Comparison of the typical monthly behaviours from acute RT (a) and obtuse RT (b) events for each driver during their baseline year including the yearlong linear trend. Each month was represented by the median of the mean net accelerations ( $\bar{a}_N$ ) across all the events during each month and the linear interpolation imputed the missing in-between months.

From Figure 5-13, Driver 4 and Driver 8 had the largest variations for acute and obtuse RT respectively. There was a weaker positive relationship between average acceleration

per month and variations during the year for acute RT ( $R = 0.574$ ) in Figure 5-14a and obtuse RT ( $R = 0.491$ ) in Figure 5-15a when compared to ACC and DEC. In addition to the cluster of drivers, Driver 12 had uniquely large variations during the year in Figure 5-14b, and Driver 5 had a large variation between months indicating that there were distinct behaviours for different months. Driver 1, 2, 3, 6, 7, 9, and 10 formed a cluster for acute and obtuse RT in Figure 5-14b and Figure 5-15b as a group with relatively low variations during the year.

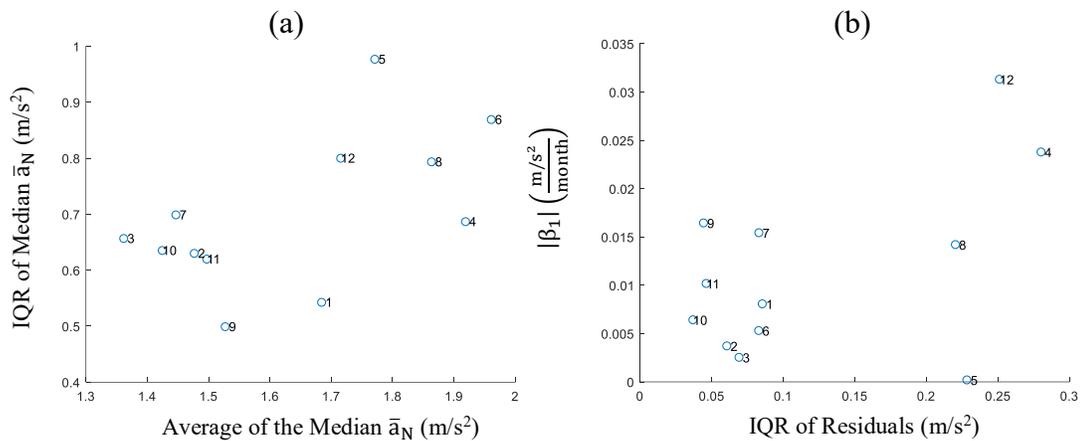


Figure 5-14. (a) The relationship between the IQR of the median  $\bar{a}_N$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_N$  across the entire year,  $R = 0.574$  ( $p = .051$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_N$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_N$ ,  $R = 0.524$  ( $p = .080$ ), for acute RT events.

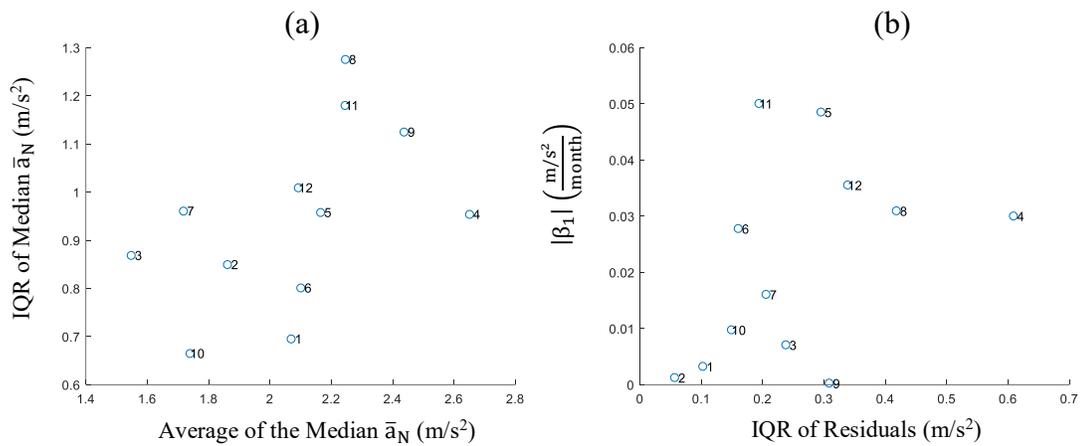


Figure 5-15. (a) The relationship between the IQR of the median  $\bar{a}_N$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_N$  across the entire year,  $R = 0.491$  ( $p = .105$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_N$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_N$ ,  $R = 0.400$  ( $p = .198$ ), for obtuse RT events.

Similar to ACC and DEC events, there was no significant relationship between annual mileage and acute and obtuse RT events as shown by  $R = 0.155$  ( $p = .630$ ) and  $R = 0.296$  ( $p = .351$ ) respectively for relationship between the average median  $\bar{a}_N$  vs. the annual mileage, and  $R = -0.086$  ( $p = .790$ ) and  $R = -0.275$  ( $p = .387$ ) respectively for the relationship between the IQR of the median  $\bar{a}_N$  vs. the annual mileage.

The results from testing the mean  $\bar{a}_N$  for acute and obtuse RT events in different driving conditions for weather, season, traffic, and daylight are shown in Appendix B.3 and Appendix B.4 and summarized in Table 5-9. Of the four factors, season and daylight had the most influence on the drivers, which was similar to ACC and DEC events.

Table 5-9. Summary of the comparisons between the mean  $\bar{a}_N$  for acute and obtuse RT during favourable and adverse conditions per driver. From the group difference between the means ( $\Delta = \mu_{\text{adverse}} - \mu_{\text{favourable}}$ ), a negative value indicated that the mean acceleration was lower during adverse condition.

	Weather		Seasons		Traffic		Daylight	
	Acute	Obtuse	Acute	Obtuse	Acute	Obtuse	Acute	Obtuse
Number of drivers showing statistical significance	3/12	3/12	4/12	4/12	1/12	0/12	6/12	4/12
Number of drivers with lower mean $\bar{a}_t$ in adverse conditions	7/12	10/12	7/12	9/12	6/12	5/12	8/12	11/12
Average group difference of mean $\bar{a}_t$ (m/s <sup>2</sup> )	-0.064	-0.125	-0.072	-0.123	0.036	0.058	-0.073	-0.107

### 5.3.3 Left Turns Events

The distribution of  $\bar{a}_N$  from acute and obtuse LT events for each driver is shown in Figure 5-16 and the corresponding table of values in Table 5-10 and Table 5-11. There were 53/66 (80%) and 47/66 (71%) significantly different pairs of drivers for acute and obtuse LT events respectively, which was less than the RT events. Similar to RT events, the mean  $\bar{a}_N$  was generally higher for obtuse turns (i.e., sharper turns) than the acute turns (i.e., gradual turns).

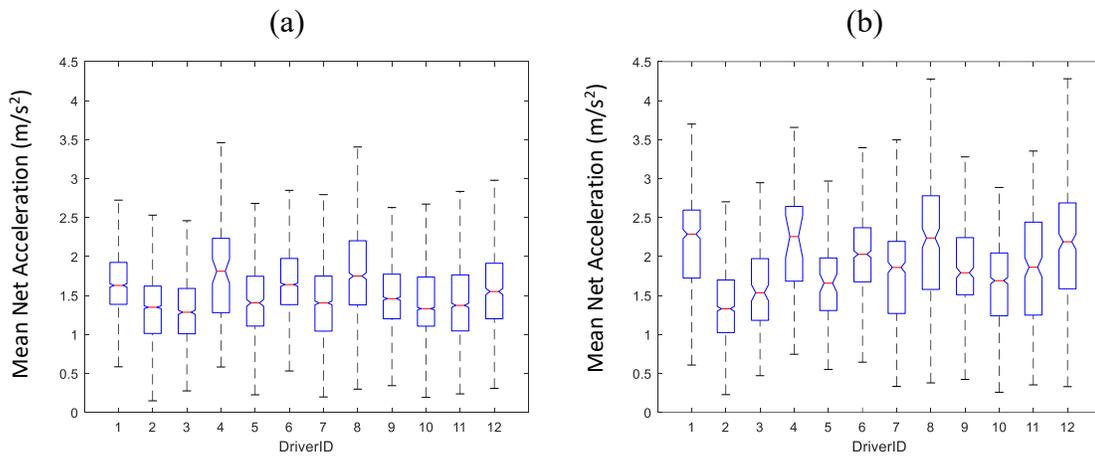


Figure 5-16. Comparison of the distributions of mean net acceleration ( $\bar{a}_N$ ) for acute LT events (a) and obtuse LT events (b) per driver for the entire baseline year. The red lines represented the median value and the boxes marked the width of the IQR. The outliers were removed by setting the lower limit to  $Q_1 - 1.5(IQR)$  and upper limit to  $Q_3 + 1.5(IQR)$ . By removing the outliers, the distribution was approximated by a normal distribution  $N(\mu, \sigma^2)$ .

Table 5-10. Summary of the mean  $\bar{a}_N$  for acute LT events per driver over the entire baseline year including the group averages.

DriverID	N	$\mu$ (m/s <sup>2</sup> )	95% CI		$\sigma$ (m/s <sup>2</sup> )	IQR (m/s <sup>2</sup> )
1	1459	1.654	1.633	1.675	0.407	0.538
2	2683	1.308	1.291	1.325	0.449	0.608
3	1025	1.310	1.284	1.336	0.430	0.581
4	93	1.814	1.679	1.948	0.653	0.953
5	500	1.435	1.395	1.474	0.450	0.639
6	1940	1.681	1.661	1.701	0.456	0.592
7	2141	1.400	1.379	1.422	0.510	0.707
8	997	1.787	1.748	1.825	0.618	0.823
9	1782	1.461	1.441	1.480	0.420	0.572
10	4225	1.402	1.389	1.415	0.441	0.630
11	1696	1.420	1.396	1.445	0.518	0.719
12	757	1.570	1.531	1.609	0.542	0.712
Group	19298	1.467	1.460	1.474	0.491	0.643

Table 5-11. Summary of the mean  $\bar{a}_N$  for obtuse LT events per driver over the entire baseline year including the group average.

DriverID	N	$\mu$ (m/s <sup>2</sup> )	95% CI		$\sigma$ (m/s <sup>2</sup> )	IQR (m/s <sup>2</sup> )
1	547	2.177	2.126	2.228	0.607	0.873
2	1232	1.357	1.329	1.386	0.508	0.675
3	141	1.590	1.498	1.682	0.553	0.791
4	31	2.169	1.939	2.399	0.627	0.959
5	80	1.673	1.557	1.789	0.521	0.674
6	616	2.022	1.981	2.063	0.519	0.694
7	480	1.763	1.702	1.825	0.685	0.925
8	236	2.193	2.093	2.293	0.779	1.202
9	304	1.861	1.795	1.927	0.584	0.732
10	999	1.627	1.595	1.658	0.509	0.806
11	227	1.851	1.754	1.948	0.739	1.192
12	293	2.178	2.087	2.270	0.796	1.101
Group	5186	1.764	1.746	1.782	0.661	0.942

Figure 5-17 displays the typical monthly acceleration behaviour from the LT events per driver during their respective baseline year. The ICC of 0.500 and 0.657 for acute and obtuse LT events indicated that there was moderate consistency among the drivers showing that individual preferences were mostly steady throughout the year, which was similar to RT events but less than ACC and DEC events.

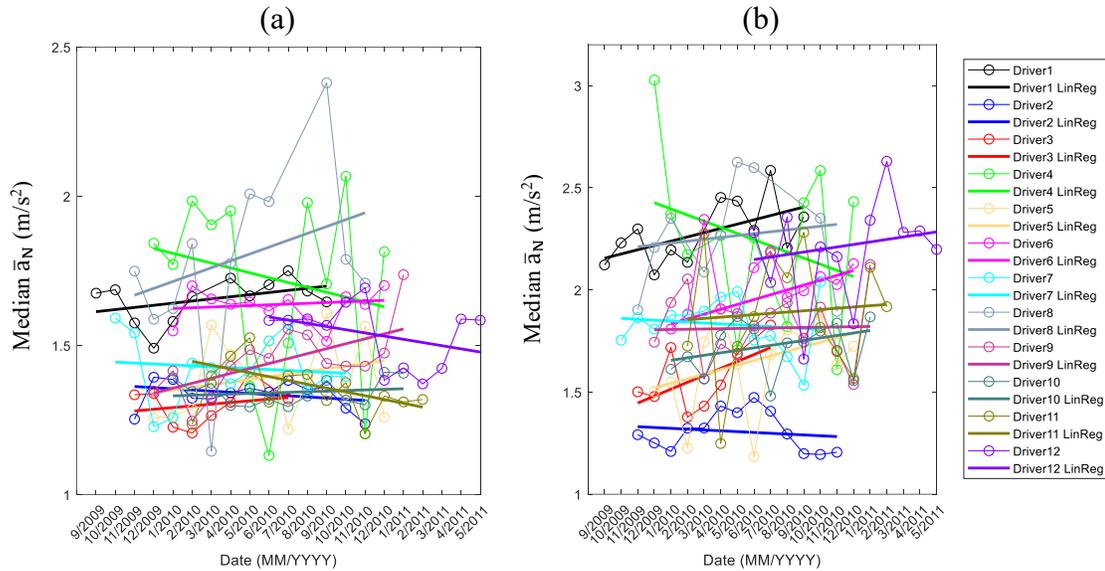


Figure 5-17. Comparison of the typical monthly behaviours from acute LT (a) and obtuse LT (b) events for each driver during their baseline year including the yearlong linear trend. Each month was represented by the median of the mean net accelerations ( $\bar{a}_N$ ) across all the events during each month and the linear interpolation imputed the missing in-between months.

From Figure 5-17, Driver 4 and Driver 8 had the largest variations for acute and obtuse LT respectively, which was similar to RT events. There was not a strong relationship between the average acceleration per month and variations during the year for acute LT ( $R = 0.561$ ) in Figure 5-18a or obtuse LT ( $R = 0.415$ ) in Figure 5-19a. Driver 4 and 8 had uniquely large variations during the year in Figure 5-18b for acute LT. Only Driver 4 had unusually large variation between months in Figure 5-19b for obtuse LT indicating that there may not be not a clear pattern for Driver 4 during obtuse LT events. Drivers who had much larger variations than other drivers may indicate low consistency and absence of patterns in their maneuvers.

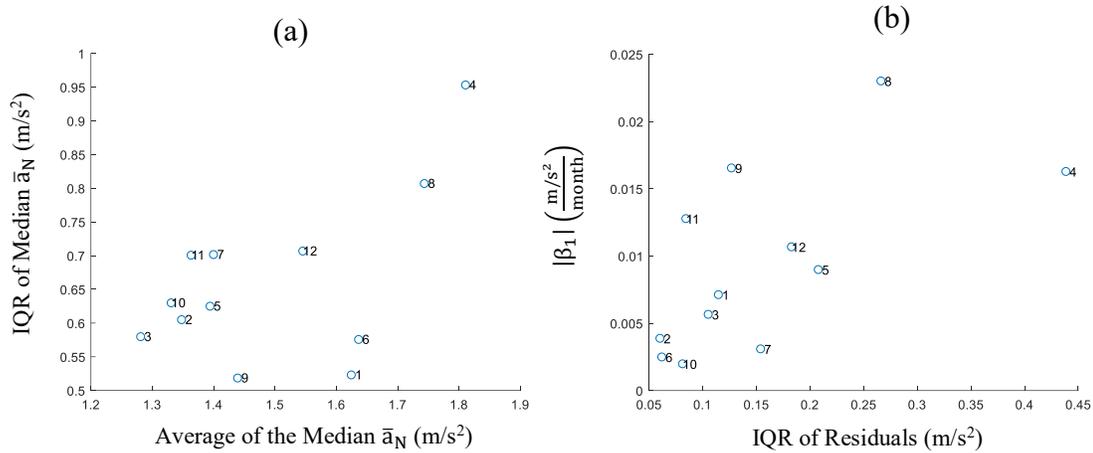


Figure 5-18. (a) The relationship between the IQR of the median  $\bar{a}_N$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_N$  across the entire year,  $R = 0.561$  ( $p = .058$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_N$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_N$ ,  $R = 0.634$  ( $p = .027$ ), for acute LT events.

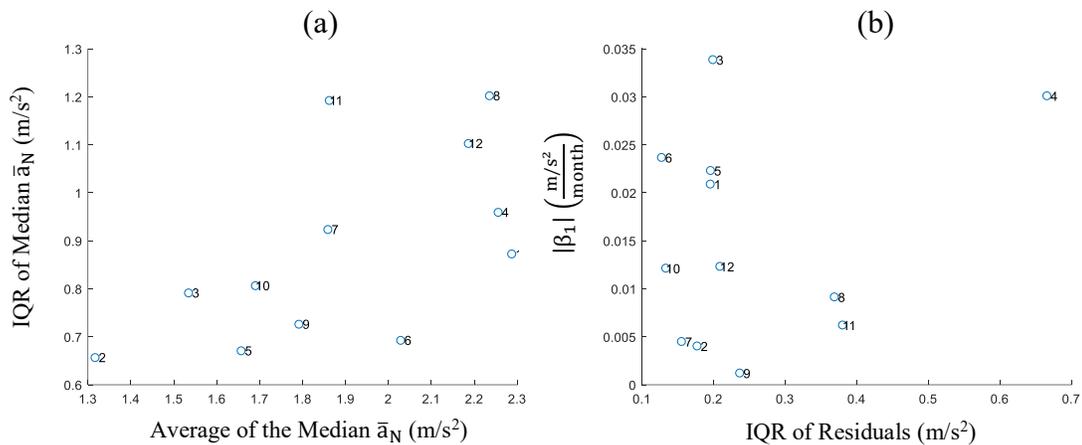


Figure 5-19. (a) The relationship between the IQR of the median  $\bar{a}_N$  (i.e., typical monthly accelerations) and the average of the median  $\bar{a}_N$  across the entire year,  $R = 0.415$  ( $p = .180$ ), and (b) the relationship between the magnitude of the yearlong change in the median  $\bar{a}_N$  per month and the IQR of the residuals from the linear regression on the median  $\bar{a}_N$ ,  $R = 0.226$  ( $p = .481$ ), for obtuse LT events.

Similar to RT events, there was no significant relationship between annual mileage and acute and obtuse LT behaviour as shown by  $R = 0.222$  ( $p = .487$ ) and  $R = 0.090$  ( $p = .783$ ) respectively for relationship between the average median  $\bar{a}_N$  vs. the annual mileage, and  $R = 0.464$  ( $p = .129$ ) and  $R = -0.034$  ( $p = .906$ ) respectively for the relationship between the IQR of the median  $\bar{a}_N$  vs. the annual mileage.

The results from testing the mean  $\bar{a}_N$  for acute and obtuse LT events in different driving conditions is shown in Appendix B.5 and Appendix B.6 and summarized in Table 5-12.

Table 5-12. Summary of the comparisons between the mean  $\bar{a}_N$  for acute and obtuse LT during favourable and adverse conditions per driver. From the group difference between the means ( $\Delta = \mu_{\text{adverse}} - \mu_{\text{favourable}}$ ), a negative value indicated that the mean acceleration was lower during adverse condition.

	Weather		Seasons		Traffic		Daylight	
	Acute	Obtuse	Acute	Obtuse	Acute	Obtuse	Acute	Obtuse
Number of drivers showing statistical significance	3/12	1/12	5/12	1/12	3/12	1/12	4/12	3/12
Number of drivers with lower mean $\bar{a}_t$ in adverse conditions	8/12	5/12	7/12	8/12	8/12	3/12	11/12	9/12
Average group difference of mean $\bar{a}_t$ (m/s <sup>2</sup> )	-0.014	0.028	-0.048	-0.053	-0.068	0.003	-0.101	-0.082

#### 5.4 Summary

This chapter examined the similarities and differences in driving behaviour among the stable health older drivers by using 4 types of driving maneuvers, which were designed to describe how they operated their vehicle and also capture their individual preferences. A relatively high acceleration during the year may indicate aggressive and high confidence behaviour and a relatively low acceleration may indicate difficulty in maneuvering the vehicle because the driver was more cautious and had lower confidence. To measure the consistency of the maneuvering behaviour during the year, a plot of the yearlong linear trend  $|\beta_1|$  vs. the IQR of the residuals of the typical monthly acceleration was proposed to visualize the behaviour of each individual driver as well as the clusters of drivers with similar behaviours. High variations between months may indicate difficulty in controlling the vehicle as the seasons and conditions change, and large yearly change may indicate general decline in maneuvering safely. The maneuvers that had

lower ICC values, such as RT and LT events, may indicate that drivers had less strong preferences towards those maneuvers and that the drivers tend to have similar turning behaviours.

For long-term monitoring, large deviations from the measured values of the previous years may be an indicator of a decline in driving ability and a decline in health. For all maneuvers, more than 70% of the pairs were significantly different and the ICC values showed moderate to high level of consistency among the drivers, which provided evidence that these maneuvers can reveal individual preferences, which had a stable pattern throughout the year. It may be worthwhile to monitor those drivers with high consistency and low variations during the year by their acceleration parameters,  $\bar{a}_t$  and  $\bar{a}_N$ , to observe changes in their normal, typical behaviour. The deceleration maneuver had the most distinct differences between drivers, and drivers with higher average acceleration during the year tend to have larger variations during the year for ACC and similarly with DEC events. On average, most drivers exhibited lower acceleration (deceleration for braking) in adverse driving conditions based on their mean accelerations. Daylight had the greatest impact on their maneuvering behaviour possibly because of a combination of reduced visibility and that more pedestrians, parked vehicles, and more traffic occurred at nighttime since the sun sets during rush hour in the winter.

By analyzing the characteristics of their maneuvers, these 12 older drivers provided insight into the typical behaviours of stable-health older drivers. The results from this analysis can be used to characterize their driving behaviour in their baseline year, which can be useful for the long-term monitoring, beyond the baseline year, to observe the effects of aging and detect at-risk behaviour.

## 6 Chapter: Driver Identification Using Vehicle Maneuvers

### 6.1 Introduction

Driving is a common task involving learned skill and individual preferences that can differ between drivers. Patterns can exist in the driving maneuvers that can reveal individual preferences between drivers, which can be beneficial for differentiating drivers of shared vehicles or detecting changes in driving behaviour. By using pattern recognition techniques, complex patterns can be discovered in the driving maneuvers.

The characteristics from motor vehicle operations were found to differentiate older drivers from the Candrive study [65], and the differences in acceleration and deceleration events between the older drivers were shown in [9][69]. By using a two-class classification approach, members within the pairs of older driver were distinguished from one another by using several supervised learning algorithms with linear and nonlinear decision boundaries [55]. The features that were included in the two-class classification were trip-level attributes, including the speed limit of the routes, and event-level attributes, including the two-phase deceleration characteristics of the drivers [9]. For a total of 162 features, 91 unique pairs from a pool of 14 drivers were compared and the best accuracy was 79.3% for the least distinguishable pair and 98.9% for the most distinguishable pair using a 75%/25% holdout and 10-fold CV testing scheme. Since some features had low discriminative ability, different subsets of the features were used for each pair because the most discriminative features depended on the characteristics of the included drivers.

This chapter extends this previous research by investigating the use of the drivers' vehicle maneuvers to identifying the driver from a group of drivers; the maneuvers were

acceleration from stopped, deceleration coming to a stop, right turns and left turns that occurred on 40 km/h to 60 km/h speed limit roadways as described in Chapter 5.2.2. A multiclass classification approach was used to train supervised learning algorithms on events from their trips and to test driver identification. The objectives of this chapter are as follows:

- 1) Propose and evaluate a method to identify drivers from their individual driving events. Namely, identify the driver by their acceleration events (ACC), deceleration events (DEC), right turns (RT), and left turns (LT) separately.
- 2) Propose and evaluate a method to identify the driver from their trips based solely on their maneuvering events by using all four types of maneuver.

This work was supported by the research presented in “Driver Identification Using Vehicle Acceleration and Deceleration from Naturalistic Driving of Older Adults” by Fung *et al.* [80] where it was shown that the acceleration and deceleration events collected by Wallace *et al.* [9] could be used to uniquely identify the driver from a group of 2 to 14 drivers using two expert classifiers (i.e., one classifier recognized drivers by their acceleration events whereas the other by deceleration events).

After the introduction, the methodology is presented describing the feature engineering and the testing scheme for the multiclass classification. Next, the results are presented from the classification of individual drivers using independent expert classifiers, and then combining the expert classifiers to form a multi-expert classifier. A discussion is presented to compare the individual expert classifiers’ performances and the performance of the multi-expert classifier. Lastly, the main results are summarized.

## 6.2 Methodology

The dataset contained all the extracted driving events for each of the 12 stable-health older drivers  $C_l$  ( $l = \{1, 2, \dots, L\}$ ) for  $L = 12$  drivers as described in Chapter 5.2.2, including acceleration events (ACC), deceleration events (DEC), right turn events (RT), and left turn events (LT). Each observation  $m_i$  ( $i = \{1, 2, \dots, I\}$ ) in the dataset was an individual driving event belonging to a particular driver, and each observation was assumed to be independent samples of driving events and had identical probability distributions. Features  $x_j$  ( $j = \{1, 2, \dots, J\}$ ) from the events were extracted to summarize each event, which belong to a particular trip  $T_n$  ( $n = \{1, 2, \dots, N\}$ ).

### 6.2.1 Training and Testing Scheme

Two supervised learning algorithms were used to investigate the multiclass classification performance: multiclass logistic regression (LR) and random forest (RF). A one vs. all scheme was used for LR with regularization to improve generalization and a learning rate of 0.01. The purpose of the LR was to obtain the benchmark performance and explore the performance of linear decision boundaries. The RF with 5 decision trees was used for the final proposed classifier because of its greater flexibility from capturing nonlinear patterns in data. The training and testing scheme was implemented in MATLAB 2017b, using the Statistics and Machine Learning Toolbox, and was separated into 3 parts:

1. Event Classifiers: the supervised learning algorithms were trained on the individual driving events to create an event classifier for each type of driving event: ACC, DEC, RT, and LT. The classifiers identified the driver by classifying the individual events using a 5-fold CV with 10 iterations (i.e., 10x5 CV).

2. Expert Trip Classifiers: a majority-vote (MV) using Borda count method [81] was used to vote for the predicted driver among the events belonging to the same trip. These independent expert classifiers  $e_k$  identified the driver by classifying the trips  $T_n$ , which contained the driving events.
3. Multi-Expert Classifier: MV was used to predict the driver by voting among all the expert classifiers  $e_k$ . The multi-expert (ME) classifier identified the driver by classifying the trips  $T_n$ , which contain events from any of the 4 types of maneuvers.

The testing and training scheme followed a 10x5 CV as shown in Figure 6-1. The list of steps is described below:

- 1) The in-vehicle sensor data was separated by their driver ID and segmented into individual trips as described in Chapter 3.5.
- 2) For each driver, stratified random sampling was used to split 80% of the trips for the training set and 20% of the trips for the testing set. This was repeated 5 times to create 5 folds for each of the 10 iterations.
- 3) The driving events were extracted from the training and testing trips to create the training and testing set of events. The smaller classes were oversampled with replacement to the size of the largest class to balance the training set and avoid bias towards a particular driver. In the testing set, the larger classes were undersampled to balance the frequency of all trips from each driver because the drivers were treated equally with equal prior distribution.
- 4) Supervised learning algorithms were trained on the events from the training set to create classifiers that recognized the driver by each type of maneuver, and the

testing set of events were used for performance evaluation. The 5-fold CV was iterated 10 times to account for the variability from the random splitting and sampling.

- 5) Steps 2 to 4 were repeated for one sequence of subsets of 2 to 12 drivers to test varying group sizes. Ideally, all combinations of drivers could be evaluated to assess the classification performance against increasing number of drivers, but this comes at a prohibitive cost of computing time for marginal insight. Given the fixed set of features across a pool of 12 stable-health drivers, it was assumed the subset sequence would not significantly impact the performance trend of the same classifier to handle additional drivers. Step 5 is the procedure for part 1.
- 6) Steps 2 to 5 were repeated using a MV (Eq. 6-2) among the driving events belonging to the same trip  $T_n$ . Step 6 is the procedure for part 2.
- 7) Steps 2 to 5 were repeated for a ME classifier using MV (Eq. 6-2) as the ensemble method. Step 7 is the procedure for part 3.

Borda count is an ensemble method of combining classifiers that is a type of majority-vote, where the opinions of the individual expert classifiers  $e_k$  are consolidated. A rank vector  $V_k = [v_k(1), v_k(2), \dots, v_k(L)]$  is a permutation of  $C_l$  where the first element  $v_k(1)$  represents the class with the most confidence (i.e., highest rank) and the subsequent elements were sorted by decreasing confidence. The rank was based on the posterior probabilities  $P(C_l|x_1, x_2, \dots, x_j)$  of each class, where  $V_k$  was created from classifying each event  $m_i$  by the  $k^{th}$  expert; the largest posterior probability was assigned rank 1 ( $v_k(1)$ ) and the lowest posterior probability assigned rank  $L$  ( $v_k(L)$ ). The Borda count  $r_k(l)$  is a score computed based on the rank vector  $V_k$  by assigning a larger score

for higher ranked classes, and the Borda counts were consolidated by summing the ranks  $R(l)$  across the individual expert classifiers.

$$r_k(l) = L - \arg[v_k(\lambda) = l] \quad \text{Eq. 6-1}$$

$$R(l) = \sum_{k=1}^K r_k(l) \quad \text{Eq. 6-2}$$

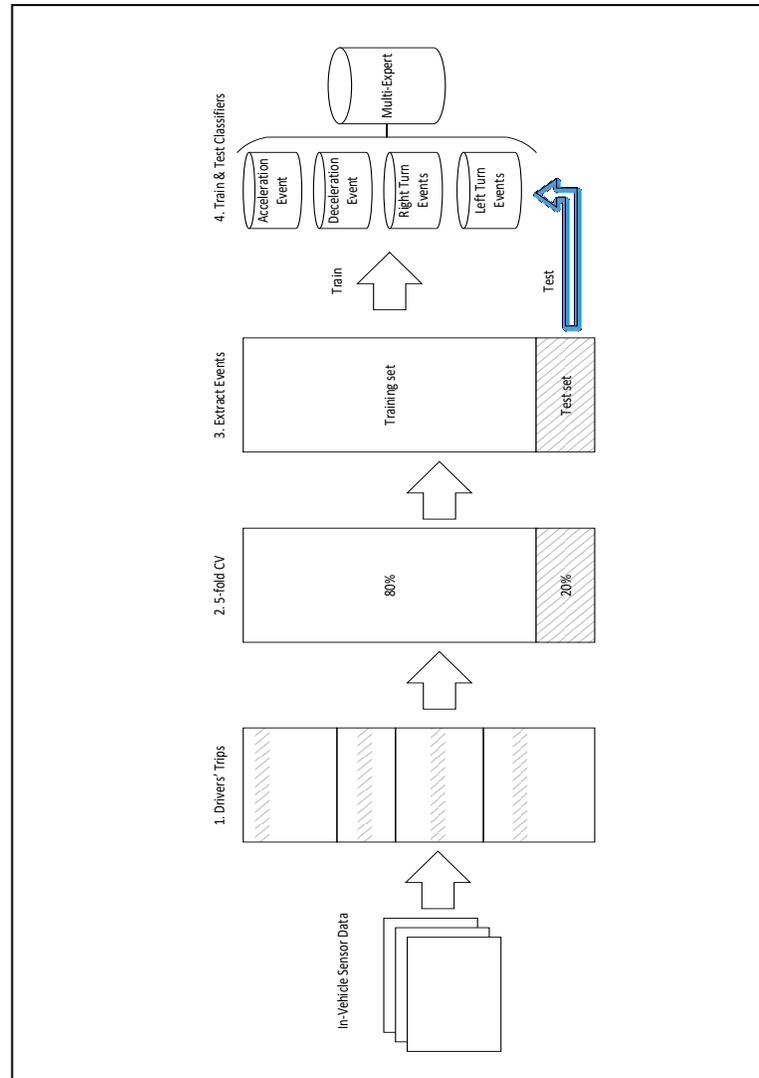


Figure 6-1. The training and testing scheme for the multiclass classification of drivers using maneuvering events. (1) The in-vehicle sensor data was prepared by organizing the dataset by the driver ID and segmented into individual trips. (2) The dataset of trips were split using 5-fold CV for the separate testing and training sets during the evaluation. (3) After extracting all maneuvering events from the trips, (4) the four independent expert classifiers were trained on each type of event, and then the expert classifiers were combined to create a multi-expert classifier to identify the drivers based on their maneuvering events.

## 6.2.2 Feature Selection

To summarize each event  $m_i$ , features were extracted to describe the characteristics of the event including descriptive statistics and morphological attributes of the discrete-time signals from the events as shown in Table 6-1. The discrete-time signals

$s[n]$  referred to the speed, tangential acceleration, centripetal acceleration, net acceleration, and jerk time-series values from the events. The features describing the environmental conditions (e.g., Hourly Temperature) were included only for testing the ME classifier for performance improvement.

Table 6-1. List of the features extracted from the driving events and the discrete-time signals  $s[n]$ .

Features	Description
dur	The duration (s) of the event.
dist	The distance (m) the vehicle travelled during the event.
avg	The average speed of the entire event (km/h).
max	The maximum value of the signal.
min	The minimum value of the signal.
mean	The mean value of the signal.
median	The median value of the signal.
sd	The standard deviation value of the signal.
kurt	The kurtosis value of the signal.
skew	The skew value of the signal.
rms	The root mean square value of the signal.
range	The difference between the max and min values of the signal.
s2e	The difference between the first and last values of the signal.
posAmp	The difference between the max and mean value of the signal.
negAmp	The difference between the mean and min value of the signal.
energy	The energy of the signal. $energy = \sum s[n]^2$
ent	The power spectral entropy of the signal [82] calculated by: 1. Calculate $S(\omega_i) = FFT(s[n])$ 2. Calculate $\hat{P}(\omega_i) = \frac{1}{N}  S(\omega_i) ^2$

	3. Normalize $p_i = \frac{\hat{P}(\omega_i)}{\sum_i \hat{P}(\omega_i)}$ 4. Calculate $H = -\sum_{i=1}^n p_i \ln(p_i)$
turnAngle	The angle of the enter vector and exit vector (degrees) of a turn. (Chapter 5.2.3)
turnRadius	The size of a turn by its radius (m) of the fitted circle. (Chapter 5.2.3)
locMinSpeed	Whether the minimum speed $V_{GPS}[n]$ occurred {before, after, during} the apex of the turn
<b>Environmental Features</b>	
Hourly Temperature	The temperature during the hour of the event.
Weather Outlook	The weather condition during the hour of the event.
Total Precipitation	The daily total precipitation of the event.
isDaytime	A status flag {0,1} whether the event occurred in daytime or nighttime.
isRushHour	A status flag {0,1} whether the event occurred during the high traffic density hours or not.

There were a total of 44 features for acceleration and deceleration events, and 74 features for right and left turn events. Some of these features were expected to have less discriminative ability than others, and patterns may exist in different subsets of these features depending on the included drivers in the group. For LR, the most important features were selected to remove the redundant features and reduce the dimensionality of the dataset before testing. Prior to LR, RF was used to perform feature selection to obtain the top 10 features for the classification of the group of 12 drivers. In MATLAB 2017b, the “PermutedVarDeltaError” metric was used for determining the importance of each feature. For each feature in each individual decision tree of RF, the observations were randomly reordered by permutation while the other predictors were left in their normal order, and then the model was retrained to determine the difference in accuracy between the new model and previous model. If permuting the predictor variable did not have an

impact on the accuracy, then the predictor was not as important as the others. Each predictor was scored from averaging the independent measures from each decision tree and across all trees [83]. Feature selection with RF was performed using 10-fold CV.

### **6.2.3 Performance Evaluation**

One sequence of subsets of drivers from sizes 2 to 12 drivers were tested to evaluate the performance trends across a pool of 12 drivers, and the same sequence was used for all classifiers. For each iteration of CV, the TP, FP, and FN were accumulated across the 5 folds to represent one instance of training and testing over the entire dataset of driving events. Since the CV splitting and over/undersampling were performed at random, the multiple iterations covered the variations between CV instances. Performance metrics were calculated from the aggregate confusion matrix.

The classification performance was evaluated by the average classification accuracy (score) and 2 times the standard deviation ( $2\sigma$ ) to describe the variability from the random sampling during CV and compared to the null model, which represented the classification accuracy of random selection (i.e., absence of patterns). The average accuracy from 5x10 CV were compared between classifiers using two sample t-test ( $\alpha = .05$ ). To examine the average classification performance per class, the means of the macro-averaging of the PC (Eq. 6-3), RC (Eq. 6-4), and Fscore (Eq. 6-5) [84] were shown and the confusion matrices for all 12 drivers were shown.

$$PC_M = \frac{\sum_{i=1}^L \frac{TP_i}{TP_i + FP_i}}{L} \quad \text{Eq. 6-3}$$

$$RC_M = \frac{\sum_{i=1}^L \frac{TP_i}{TP_i + FN_i}}{L} \quad \text{Eq. 6-4}$$

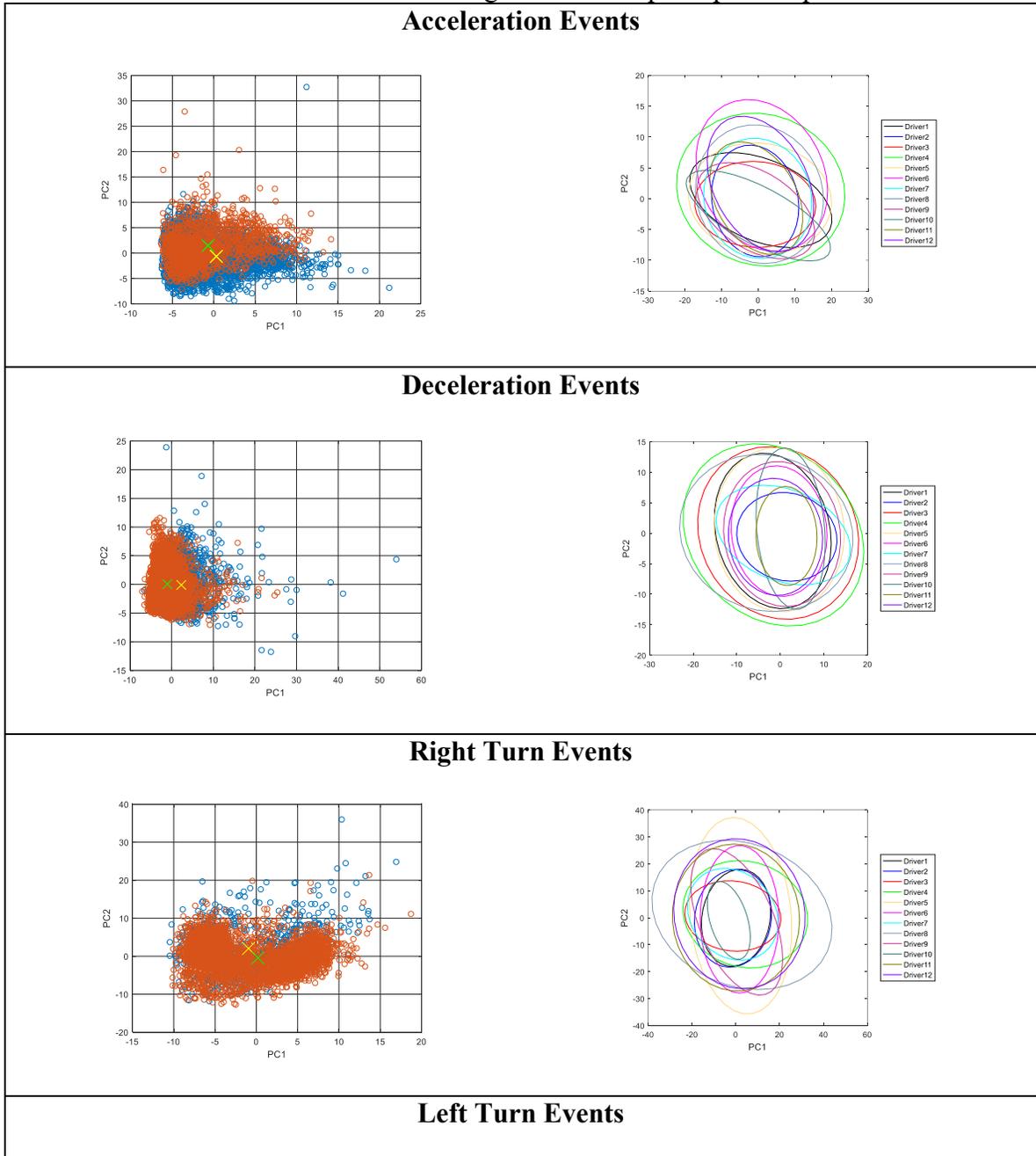
$$F_M = (2) \left( \frac{PC_M \cdot RC_M}{PC_M + RC_M} \right) \quad \text{Eq. 6-5}$$

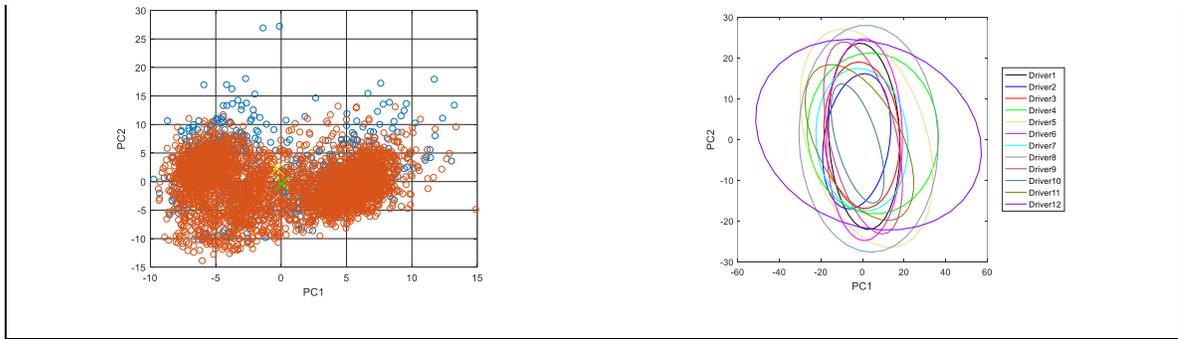
## 6.3 Results

### 6.3.1 Data Visualization

To visualize the class separation among the 12 drivers for each type of maneuver, principal component analysis (PCA) was used to reduce the dimension of the feature set to 2D for visualization of two example drivers. In addition, all 12 drivers were compared visually by assuming a Gaussian distribution of the features for each class using PCA and showing the 95% confidence ellipse. From Table 6-2, there exist some class separations and differences in mean values for each driver. The overlaps were expected since the features (e.g., mean acceleration) had similar ranges between drivers.

Table 6-2. (Left) An example of the class separation of two high mileage drivers, #2 (red) and #10 (blue), using the first two principal components to visualize the events. The green x marker represents the mean of Driver 2 and the yellow x marker represents the mean of Driver 10. (Right) The group of 12 drivers showing 95% confidence ellipse from a Gaussian distribution of the events using the first two principal components from PCA.





### 6.3.2 Event Classifiers

Table 6-3 lists the top 10 features from feature selection using RF for testing the LR. The top 10 features for each type of maneuver were ranked in order of predictor importance. For LT and RT, the turn angle and turn radius were important features as expected. For all types of maneuvers, different combinations of features from the speed, acceleration, and jerk values had the highest discriminative ability.

Table 6-3. Top 10 features from feature selection prior to training the LR event classifier

Rank	ACC	DEC	RT	LT
1	speed_ent	speed_final	turnRadius	jerk_s2e
2	jerk_skew	speed_ent	turnAngle	turnAngle
3	acc_s2e	acc_s2e	jerk_s2e	turnRadius
4	speed_median	jerk_s2e	jerk_median	jerk_max
5	jerk_mean	jerk_mean	acc_kurt	acc_mean
6	jerk_ent	speed_posAmp	speed_avg	jerk_skew
7	jerk_kurt	acc_skew	acc_median	dist
8	acc_sd	jerk_median	netAcc_s2e	acc_kurt
9	acc_ent	acc_kurt	centAcc_sd	netAcc_median
10	acc_skew	acc_posAmp	dist	netAcc_kurt

With the features listed in Table 6-3, event classifiers were trained using LR on the driving events and the results were summarized in Appendix C.1 for ACC, DEC, RT, and LT. RF was used to improve the event classifiers by training on the entire feature set and the results were summarized in Appendix C.1 for ACC, DEC, RT, and LT events. The performance of LR for all maneuvers and subsets were generally slightly above the null score with the highest score of 69.7% for DEC event classifier with a subset of 2

drivers. In contrast, the RF event classifiers outperformed the LR consistently across all maneuvers and all subsets with as much as 23.8% increase for LT events with 2 drivers. For RF, the best score was from the DEC event classifier for 2 drivers (76.5%) and ACC event classifier for 12 drivers (29.8%).

### **6.3.3 Expert Trip Classifiers**

Each type of event was grouped by their respective trips and a MV among the same type of events within the trips was used to distinguish the drivers. From the event classification using RF, a MV was implemented in each event classifier to create the expert classifiers that classified trips instead of individual events to identify the drivers by their trips. The results from classification of trips for subsets of 2 to 12 drivers using the expert classifiers were shown in Appendix C.2. The performances of the expert classifiers were similar to the event classifiers, which demonstrated that the MV did not negatively impact the classification performance. The best score was the expert DEC trip classifier for 2 drivers at 80.7% and ACC Classifier for 12 drivers at 34.6% accuracies. For the group of 12 drivers, the confusion matrices in Appendix C.2 showed the per class performances for one iteration of CV for the ACC, DEC, RT, and LT expert classifiers at 34.8%, 32.4%, 34.3%, and 32.4% accuracies, respectively. The diagonal entries, shaded in grey, were the correct classifications (TP). Driver 4 was consistently the least distinguishable driver with the lowest recall at 5.1% on average, which may indicate an absence of patterns in the driver's maneuvers or individual preferences, which was also supported by the low consistency and large variations in mean accelerations during the year as shown in Chapter 5. The drivers with lower recall were less distinguishable because they may have had less distinct patterns in their events from having similar

behaviours to other drivers. Driver 10 was among the most distinguishable drivers with the recall of 62.6% and 43.2% precision on average. Drivers 2 and 6 were also among the most distinguishable drivers with high recall and may have had the most unique patterns among the 12 drivers.

#### **6.3.4 Multi-Expert Trip Classifier**

After creating the expert classifiers, they were combined to create the ME classifier to identify the driver by their trips using all 4 types of maneuvers. The performance from the classification of trips with subsets of 2 to 12 drivers are shown in Appendix C.3 and the confusion matrix for one instance of CV for 12 drivers is shown in Appendix C.3 with 50.2% average accuracy. The performance of the ME generally improved from combining the expert classifiers, but including the environmental features did not significantly increase the performance. Driver 2, 6, 9, and 10 were the most distinguishable with the highest recall (>60%) and Driver 4 was again the least distinguishable driver with the lowest recall (3.0%).

For the group of 12 drivers, the minimum number of 6 events improved the performance from 49.4% to 54.0% as shown in Figure 6-2 from varying the minimum number of events per trip, which was about the same as the minimum of 5 events as shown by Wallace [55] for classification using acceleration and deceleration events.

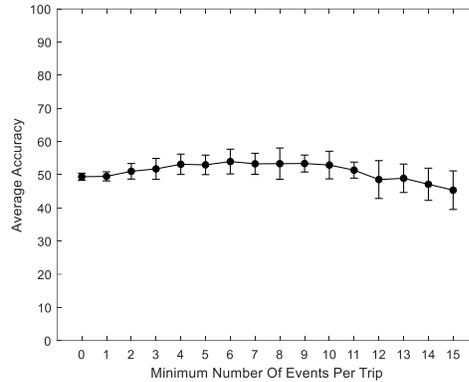


Figure 6-2. The average classification accuracies of the ME trip classifier for the group of 12 drivers and the improvement from removing trips with few events.

The results from rank classification for all 12 drivers are shown in Figure 6-3. The average rank accuracy is the proportion of test cases that correctly identified the driver as one of the top predicted drivers (e.g., Rank 5 is the top 5 drivers), and the accuracies increase monotonically with rank. For a rank of 7, the ME trip classifier performed at 89.4%, which was double the performance at rank 1 (44.7%) and more than half of the drivers could be distinguished with a high degree of confidence.

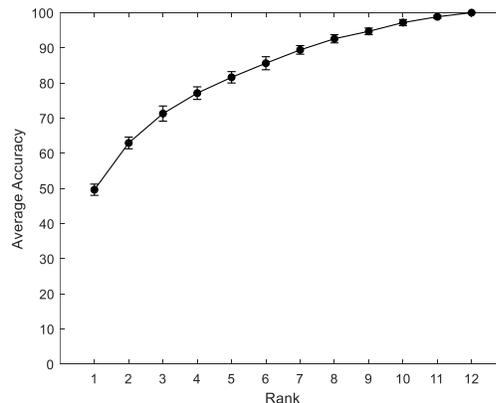


Figure 6-3. The average classification accuracies from rank classification of 12 drivers for the ME trip classifier. At rank 1, the average accuracies were the lowest having only one choice for the predicted driver, and at rank 12 the accuracies were 100% since all the drivers were predicted.

## 6.4 Discussion

The classification performance from LR provided the benchmark scores for each

subset of drivers from using a linear classifier. The results from LR showed that it was possible to distinguish drivers by their maneuvers, albeit only slightly better than the null model of random selection. By using RF, the classification performance was significantly higher for all subset and maneuvers (Table 6-4) and demonstrated that nonlinear patterns exist in the data.

Table 6-4. Summary of the p-values from significance testing between classification scores of LR and RF for each type of maneuver.

<b>Subsets</b>	<b>ACC</b>	<b>DEC</b>	<b>RT</b>	<b>LT</b>
2	< .001	< .001	< .001	< .001
3	< .01	< .001	< .001	< .001
4	.004	< .001	< .001	< .001
5	< .001	< .001	< .001	< .001
6	< .001	< .001	< .001	< .001
7	< .001	< .001	< .001	< .001
8	< .001	< .001	< .001	< .001
9	< .001	< .001	< .001	< .001
10	< .001	< .001	< .001	< .001
11	< .001	< .001	< .001	< .001
12	< .001	< .001	< .001	< .001

Similarly, the performance of the expert classifiers showed that implementing MV to classify trips instead of events did not decrease classification performance and demonstrated that it was possible to distinguish drivers by their trips using only their vehicle maneuvers. When comparing among individual experts, the DEC trip classifier had the best performance for the smaller subsets, while all four experts had similar performance for larger subsets. The deceleration maneuver had the most class separation for smaller groups of drivers, as demonstrated by the high degree of consistency among drivers (Figure 5-7), because the drivers may have had stronger individual preferences for braking than accelerating or turning. By combining the individual experts to form the ME trip classifier, the classification performance was significantly higher for most of the subsets and maneuvers (Table 6-5), and demonstrated that an ensemble approach to

distinguishing drivers can improve performance as shown in Figure 6-4 to Figure 6-7 for each type of maneuver.

Table 6-5. Summary of the p-values from significance testing between the classification scores of the expert trip classifiers and ME trip classifier for each type of maneuver.

Subsets	ACC	DEC	RT	LT
2	< .001	0.916	< .001	< .001
3	< .001	0.366	< .001	< .001
4	< .001	0.977	< .001	< .001
5	< .001	0.053	< .001	0.037
6	< .001	< .001	< .001	< .001
7	< .001	< .001	< .001	< .001
8	< .001	< .001	< .001	< .001
9	< .001	< .001	< .001	< .001
10	< .001	< .001	< .001	< .001
11	< .001	< .001	< .001	< .001
12	< .001	< .001	< .001	< .001

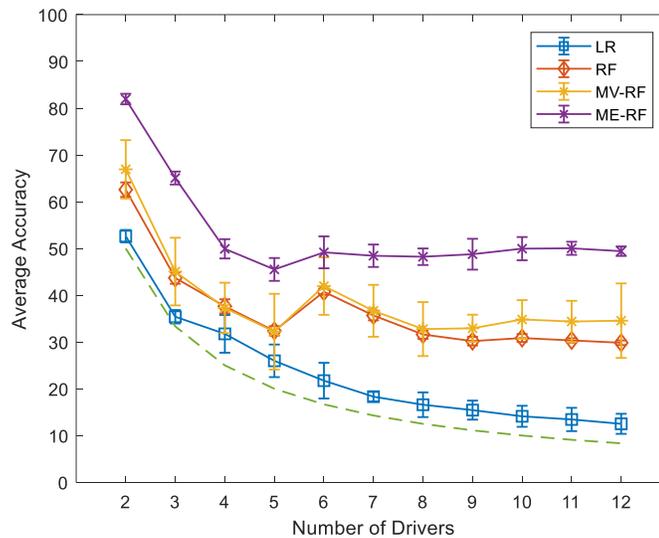


Figure 6-4. The average classification accuracies for size 2 to 12 subsets of drivers showing that majority vote (MV) applied to the ACC classifiers consistently improved performance across all subsets of drivers, and the ME classifier consistently had higher performance. The LR and RF identified drivers based on individual driving events, whereas the MV-RF and ME-RF based on trips.

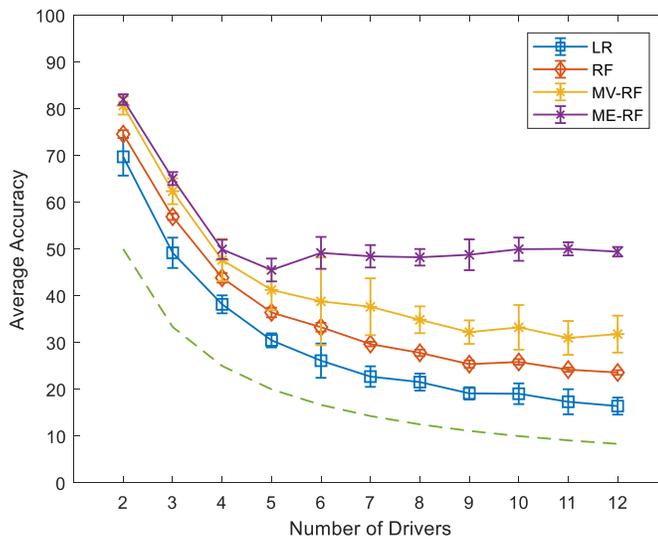


Figure 6-5. The average classification accuracies for size 2 to 12 subsets of drivers showing that majority vote (MV) applied to the DEC classifiers consistently improved performance across all subsets of drivers, and the ME classifier had higher performance for larger subsets of 6 or more drivers. The LR and RF identified drivers based on individual driving events, whereas the MV-RF and ME-RF based on trips.

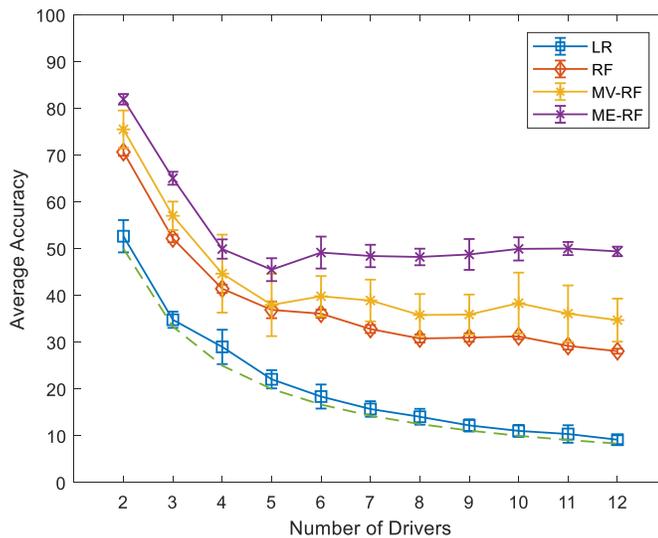


Figure 6-6. The average classification accuracies for size 2 to 12 subsets of drivers showing that majority vote (MV) applied to the RT classifiers consistently improved performance across all subsets of drivers, and the ME classifier consistently had higher performance. The LR and RF identified drivers based on individual driving events, whereas the MV-RF and ME-RF based on trips.

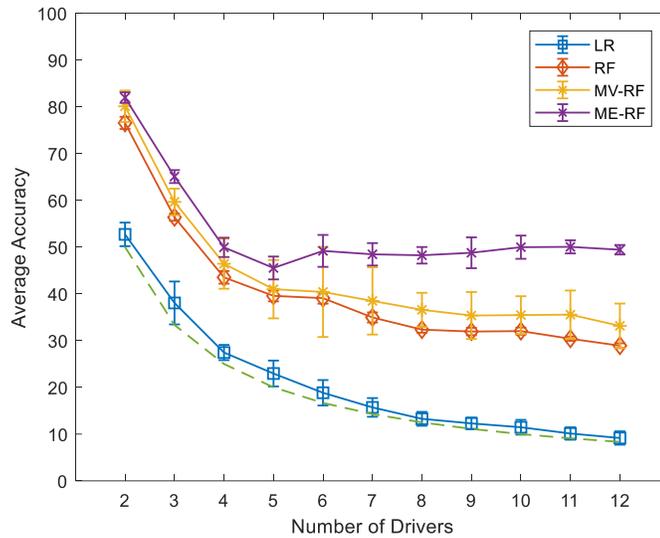


Figure 6-7. The average classification accuracies for size 2 to 12 subsets of drivers showing that majority vote (MV) applied to the LT classifiers consistently improved performance across all subsets of drivers, and the ME classifier consistently had higher performance. The LR and RF identified drivers based on individual driving events, whereas the MV-RF and ME-RF based on trips.

When compared to previous studies by Wallace [55], who used the Candrive data for two-class classification, the ME accuracy of 82.0% for a group of 2 drivers is within the range of accuracies between 79.3% and 98.9%. For future studies, computing the ME classifier accuracies for all combinations of pairs of drivers would provide the highest and lowest classification performances for groups of 2 drivers, which would be more comparable for the given range. It was expected that the highest accuracy would not be as high as 98.9% since the dataset of maneuvers did not include route attributes, which can greatly improve the identification of a driver. Hallac *et al.* [41] achieved an average of 76.9% and 50.1% for groups of 2 and 5 drivers respectively from driver identification using only turns. In comparison, the expert LT classifier achieved 80.1% and 41.0% for a group of 2 and 5 drivers respectively, which was similar in performance. The differences were attributed to the differences in testing methodology where Hallac *et al.* [41] compared drivers by the same geographical turns using the CAN-bus sensors, whereas

this chapter compared drivers by a large pool of turn examples from different geographical locations. Enev *et al.* [40] claimed 100% accuracy in differentiating 15 drivers, which was the highest accuracy of all the previous studies in driver identification. Their testing methodology included using CAN-bus sensors and extracting features from overlapping sliding windows on the time-series sensor data for each driver, and all drivers drove on a designated 50 mile course through the city which reduced the variability in the data and helped in differentiation. The high accuracy may also be attributed to data leakage from information from the training set mixing with the testing set, and the use of ensemble classifiers may be producing results overfitted on their dataset and not generalizable to naturalistic driving data.

## **6.5 Conclusion**

Driver identification using vehicle maneuvers was investigated by using a multiclass classification approach for subset sizes of 2 to 12 drivers from a pool of 12 stable-health older drivers. Specifically, LR and RF supervised learning algorithms were trained on the features extracted from the driving events to identify the drivers by their trips. A method was proposed and evaluated to identify drivers from their individual driving events by creating four independent event classifiers that were trained to recognize the driver by their acceleration, deceleration, right turn, and left turn events that occurred on low speed roadways. By using a majority-vote, the expert trip classifiers were combined to form a multi-expert trip classifier that generally improved performance over the individual experts. The ME average classification accuracy was 82.0% for a group of 2 drivers and 49.4% for the group of 12 drivers, which was about 6 times larger than the null model of random selection (8.3%).

Distinguishing between drivers demonstrated that there were distinct patterns and discernible characteristics among the drivers in how they maneuver their vehicle regardless of their chosen routes. Furthermore, these patterns may be consistent over time and that changes in their driving patterns may indicate a change in driving behaviour and health, which can be beneficial for assessing medical fitness to drive. The methods presented for driver identification could be adapted as methods to differentiate between the same driver during different time periods to detect changes in driving patterns.

## **7 Chapter: Summary**

### **7.1 Overview**

By analyzing the four types of the low-speed driving maneuvers, a portion of the older drivers' complex driving behaviour was studied to gain further insight into their individual preferences and typical behaviour during their respective baseline year. Identifying individual preferences based on common, reoccurring driving scenarios was important because these characteristics form a strong pattern in the driving data that remain consistent over time and thus considered a better representation of their driving behaviour than patterns that were more likely influenced by external factors such as traffic. The analysis of driving maneuvers is situated in the lowest level of organization hierarchy (Figure 1-1) from which there are many directions to take for future studies to bring Candrive closer to their goal of identifying predictors of MVC and the effects of aging for developing a screening tool.

There are mainly two branches of research topics that are extensions to the characterization of the baseline behaviour. Firstly, continuing to describe and measure the individual preferences among different older drivers using different statistical and pattern recognitions methods will reveal more about their baseline driving behaviour to help identify which specific parameters in the in-vehicle sensor data are relevant to finding the predictors of MVC. Secondly, subsequent years can be analyzed to observe the changes in behaviours over long-term, beyond the baseline year, and the effects of aging especially when MVC occurs.

### **7.2 Summary of Analysis**

This is the first study on analyzing the low-speed driving maneuvers using the

Candrive data from stable-health older drivers and using the patterns in the maneuvers to differentiate between drivers. This work is an extension of Wallace [55] by using the processed in-vehicle sensor data and the improved turn-identification algorithm to analyze a subset of driving events to concentrate on a pool of events that represent similar traffic scenarios to more accurately describe their individual preferences, rather than behaviour that was more likely a response to external factors such as traffic. From the previous work on two-driver classification by Wallace [55], this study proposed a novel approach to multiclass classification by using a multi-expert classifier to identify drivers by their trips based only on the patterns in their maneuvers, as opposed to including trip-level attributes. The results showed evidence that these driving maneuvers revealed individual driving preferences and may be important for monitoring changes in driving behaviour.

### **7.3 Challenges and Limitations**

One of the main challenges is conducting analysis using anonymized GPS data. On one hand, protecting the participants by preserving their anonymity for confidentiality reasons is of utmost importance for research studies. On the other, the data was a rich and detailed historical record of their everyday driving, but anonymization of the GPS data reduced the capabilities of analyzing specific traffic scenarios. Hence, creative methods must be invented to infer and deduce likely behaviours exhibited from the anonymized GPS data to estimate and generalize their driving behaviour.

Another challenge is working with unbalanced data, where some drivers had much lower number of trips than others and some drivers chose not to drive for some months. Also, missing data such as a posted limit of 0 km/h (i.e., indicated unavailable

speed limit data) reduced the number of examples of low-speed driving maneuvers. Ideally, the pool of examples that represented their driving behaviour were as large as possible to average out the effects of the unobserved factors, and thus the criteria for the types of driving maneuvers had to be reasonable to generate many events while being narrow enough to reduce the variability.

Several assumptions about the data including approximating the weather by using the weather data from the city's main airport and assuming all drivers remained within the same UTM zone (i.e., city limits) contributed to errors in the analysis.

## **7.4 Future Work**

### **7.4.1 Characterization of Baseline Year**

On the same topic of motor vehicle operation, there are other specific driving events that can be explored such as preferred cruising speeds on the higher speed roadways, characterizing expressway driving by looking at frequency of lane changing during certain hours of the day, or analyzing the cornering on expressway on and off ramps. By increasing granularity, the timeline for analysis could be expanded to identify weekly or daily trends to account for more specific behaviours such as weekend driving vs. weekday driving behaviours, but a challenge is that older drivers may not be driving everyday. Since many drivers follow a routine schedule, their maneuvering behaviours during their frequent routes could be analyzed to compare the same routes and control for the trip distance and traffic conditions. With a larger sample of stable-health older drivers, generalization of driving behaviour can be improved by obtaining more examples of their typical behaviours. Different groups of older drivers with similar traits could be compared such as analyzing the maneuvering behaviour of 70 and 80 year-old drivers to

identify differences between age groups and common trends among older groups to identify the effects of aging.

#### **7.4.2 Analysis of Subsequent Years**

Monitoring the maneuvering behaviour of the stable-health older drivers for many years could lead to observing changes in behaviour and these changes could be cross-referenced with the driver's health report and lifestyle choices to understand the causes of the changes in driving behaviour. For the low-speed driving maneuvers, the plot of the yearlong trend vs. the variations between months could be used to summarize the consistency in behaviour during each year, and the changes in consistency between years can be measured by using a distance metric (e.g., Euclidean) to calculate the differences over the years.

#### **7.5 Conclusion**

This thesis research proposed the use of low-speed driving maneuvers to contribute to the characterization of the baseline driving behaviours of stable-health older drivers using the in-vehicle sensor data with anonymized GPS data. The characteristics of their driving maneuvers can provide more information on their typical driving behaviour, and provide parameters of interest that can be monitored over long-term to detect changes in driving behaviour. A change in their maneuvering behaviour may indicate a decline in cognitive and physical health leading to the discovery of predictors of MVC, and thereby improving the screening process of older drivers by providing evidence from their actual driving behaviour.

## Appendices

### Appendix A : Data Validation

#### A.1 Criteria of Triple-Stable Older Drivers

List of the criteria for triple-stable [54]:

- “Cognitive stability was defined as a Montreal Cognitive Assessment (MOCA) score greater than 24 at baseline and no more than a 2 point decline at year 1 assessment.”
- “Cognitive decline was defined as a decrease in MOCA score of 3 or more points between baseline and year 1 assessment.”
- “Physical stability was defined as a Timed Up & Go (TUG) score at year 1 that was less than or equal to the TUG score at baseline assessment, where TUG scores were classified as “1” for times less than 10 s, “2” for 10 to 19 s, and “3” for greater than or equal to 20 s.”
- “Physical decline was defined as a TUG score at year 1 that was greater than the TUG score at baseline assessment.”
- “General health stability criteria was defined as a Cumulative Illness Rating Scale (CIRS) total score that did not increase by more than 2 points between baseline and year 1 assessment.”
- “General health decline was defined as a CIRS total score that increased by 3 or more points between baseline and year 1 assessment.”

#### A.2 In-Vehicle Sensor Data

Table A-1. List of the sensor signals from the in-vehicle sensor data recorded by the Candrive data logging device [64].

#	Source	Parameter	Description	Units/Format
1	Device	Sample Number	The sample number, increments for every new sample @ 1Hz	Integer
2	Device	Serial Number	Serial number of the central unit	Integer
3	Device	Trip Number	Increments for every trip	Integer
4	Device	Key Tag	RFID tag of the current driver	Integer
5	GPS	Date And Time	Date and time stamps in UTC	M/DD/YYYY HH:MM
6	GPS	Latitude	GPS coordinates (Accuracy $\pm 10$ m)	Decimal Degrees
7	GPS	Longitude	GPS coordinates (Accuracy $\pm 10$ m)	Degrees Degrees
8	GPS	Fix Status	The number of satellites connected	{0,3,4}
9	GPS	DOP	Dilution of Precision is measurement of the error for GPS coordinates	Floating point
10	GPS	GPS Speed	Speed of the Vehicle (Accuracy 1.0 km/h)	km/h
11	GIS	Speed Limit	Posted Speed Limit of the roadway	km/h
12	GIS	Alerts	Road Signs/Hazards in proximity	Nominal
13	OBDII	Engine coolant temperature	-	$^{\circ}$ C
14	OBDII	Fuel Trim	% changes to the fuel mixture	%
15	OBDII	Engine RPM	-	Integer
16	OBDII	Vehicle speed sensor	Speed reported by the engine computer	km/h

17	OBDII	Intake air temperature	-	°C
18	OBDII	Ambient Air Temperature	-	°C
19	OBDII	Absolute throttle position	-	Floating point
20	OBDII	Relative throttle position	-	%

Table A-2. Summary of the data verification and imputation methods in the processing pipeline for the raw CSV files from Candrive.

#	Variable	Valid Criteria	Invalid Values Occurrence Rate	Imputation Method/Action
1	Record Number	Monotonically Increasing	0 %	-
2	Date/Time	Unique and non-empty	< 1 %	Discarded
	Year	$2009 \leq \text{Year} \leq 2015$		
	Month	$1 \leq \text{Month} \leq 12$		
	Day	$1 \leq \text{Day} \leq 31$		
	Hour	$1 \leq \text{Hour} \leq 59$		
	Minute	$1 \leq \text{Minute} \leq 59$		
	Second	$1 \leq \text{Second} \leq 59$		
3	Trips	< 120 s OR < 100 m	< 1%	Discarded
4	GPS Longitude and Latitude	<ul style="list-style-type: none"> <li>Fix Status <math>\neq</math> "No Fix"</li> <li>Realistic changes in position (&lt; 150 km/h)</li> <li>DOP is &lt; 5</li> </ul>	~6%	Categorized the trips as {High, Medium, Low} quality
5	High Quality Trips	No invalid data	~98% of Trips	-
6	Medium Quality Trips	Some invalid data	~1% of Trips	<ul style="list-style-type: none"> <li>Nearest Neighbour for the beginning and end of trips</li> <li>Linear Interpolation During the trips</li> </ul>
7	Low Quality Trips	Too much invalid data	~1% of Trips	Discarded
8	Speed Difference	$ relDiff[n]  < 3\sigma_{relDiff[n]}$	~6%	$V_{GPS}[n] = V_{OBDII}[n]$
9	GPS Drift	$V_{OBDII}[n] = 0 \text{ km/hr}$ AND $V_{GPS}[n] = 0 \text{ km/hr}$	~2%	$V_{GPS}[n] = 0 \text{ km/hr}$ , Longitude and latitude replaced with previous valid coordinates

### A.3 Weather Data

Table A-3. Summary of the data verification and imputation methods for pre-processing of the hourly reported weather CSV files.

#	Variable	Data Verification	Invalid Values Occurrence Rate	Imputation/Action
1	Date/Time	Unique	0 %	-
	Year	$2009 \leq \text{Year} \leq 2015$		
	Month	$1 \leq \text{Month} \leq 12$		
	Day	$1 \leq \text{Day} \leq 31$		
	Hour	$1 \leq \text{Hour} \leq 23$		
2	Data Quality	-	-	Removed
3	Temp	$-40^{\circ}\text{c} \leq \text{Temp} \leq 40^{\circ}\text{c}$	< 1%	Linear Interpolation
4	Temp Flag	-	-	Removed
5	Dew Point Temp	$-40^{\circ}\text{c} \leq \text{DewPointTemp} \leq 40^{\circ}\text{c}$	< 1%	Linear Interpolation

6	Dew Point Temp Flag	-	-	Removed
7	Rel Hum	$0\% \leq \text{Rel Hum} \leq 100\%$	< 1%	Linear Interpolation
8	Rel Hum Flag	-	-	Removed
9	Win Dir	$0^\circ \leq \text{Win Dir} \leq 360^\circ$	~2 %	Nearest Neighbour
10	Win Dir Flag	-	-	Removed
11	Wind Spd	$0 \text{ km/h} \leq \text{Wind Spd} \leq 80 \text{ km/h}$	< 1%	Linear Interpolation
12	Wind Spd Flag	-	-	Removed
13	Visibility	$0 \text{ km} \leq \text{Visibility} \leq 50 \text{ km}$	< 1%	Linear Interpolation
14	Visibility Flag	-	-	Removed
15	Stn Press	$95 \text{ kPa} \leq \text{Stn Press} \leq 108 \text{ kPa}$	< 1%	Linear Interpolation
16	Stn Press Flag	-	-	Removed
17	Hmdx	$0^\circ\text{c} \leq \text{Hmdx} \leq 50^\circ\text{c}$	~87%	Nearest Neighbour (Summer Months Only)
18	Hmdx Flag	-	-	Removed
19	Wind Chill	$-50^\circ\text{c} \leq \text{Wind Chill} \leq 40^\circ\text{c}$	~71 %	Nearest Neighbour (Winter Months Only)
20	Wind Chill Flag	-	-	Removed
21	Weather	String Data Type	~28%	Nearest Neighbour

Table A-4. Summary of the data verification and imputation methods for pre-processing the daily reported weather CSV files.

#	Variable	Data Verification	Invalid Values Occurrence Rate	Imputation Method/Action
1	Date/Time	Unique	< 1 %	Discarded
2	Year	$2009 \leq \text{Year} \leq 2015$		
3	Month	$1 \leq \text{Month} \leq 12$		
4	Day	$1 \leq \text{Day} \leq 31$		
5	Hour	$1 \leq \text{Hour} \leq 59$		
6	Minute	$1 \leq \text{Minute} \leq 59$		
7	Second	$1 \leq \text{Second} \leq 59$		
8	Max Temp	$-40^\circ\text{c} \leq \text{MaxTemp} \leq 40^\circ\text{c}$	0 %	-
9	Max Temp Flag	-	-	Removed
10	Min Temp	$-40^\circ\text{c} \leq \text{MinTemp} \leq 40^\circ\text{c}$	0 %	-
11	Min Temp Flag	-	-	Removed
12	Mean Temp	$-40^\circ\text{c} \leq \text{Mean Temp} \leq 40^\circ\text{c}$	0 %	-
13	Mean Temp Flag	-	-	Removed
14	Heat Deg Day	$-40^\circ\text{c} \leq \text{Heat Deg Day} \leq 40^\circ\text{c}$	0 %	-
15	Heat Deg Days Flag	-	-	Removed
16	Cool Deg Day	$-40^\circ\text{c} \leq \text{Cool Deg Day} \leq 40^\circ\text{c}$	0 %	-
17	Cool Degree Days Flag	-	-	Removed
18	Total Rain	$0 \text{ mm} \leq \text{Total Rain} \leq 100 \text{ mm}$	0 %	-
19	Total Rain Flag	-	-	Removed
20	Total Snow	$0 \text{ cm} \leq \text{Total Snow} \leq 40 \text{ cm}$	0 %	-

21	Total Snow Flag	-	-	Removed
22	Total Precipitation	$0 \text{ mm} \leq \text{Total Precipitation} \leq 100 \text{ mm}$	0 %	
23	Total Precipitation Flag	-	-	Removed
24	Snow on Grnd	$0 \text{ cm} \leq \text{Snow on Grnd} \leq 100 \text{ cm}$ Summer months have 0 cm snow	~38 %	Nearest Neighbour
25	Snow on Grnd Flag	-	-	Removed
26	Dir of Max Gust	$-50 \leq \text{Dir of Max Gust} \leq 40$	~41 %	Nearest Neighbour
27	Dir of Max Gust Flag	-	-	Removed
28	Spd of Max Gust	$0 \text{ km/h} \leq \text{Spd of Max Gust} \leq 80 \text{ km/h}$	~39 %	Linear Interpolation
29	Spd of Max Gust Flag	-	-	Removed

## Appendix B : Comparison of Favourable and Adverse Driving Conditions

### B.1 Acceleration Events

Table B-1. Summary of the comparisons between the mean  $\bar{a}_t$  for ACC events during clement and inclement weather per driver including the groups averages. From the difference between the means ( $\Delta = \mu_{\text{inclement}} - \mu_{\text{clement}}$ ), a negative value indicated that the mean acceleration was lower during inclement weather. The relationship between  $\Delta$  and the annual mileage was  $R = -0.259$  ( $p = .416$ ).

DriverID	Clement			Inclement			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	208	0.963	0.207	664	0.880	0.210	<.001*	-0.083
2	725	0.902	0.233	1772	0.897	0.232	.314	-0.005
3	69	0.855	0.192	535	0.888	0.220	.902	0.033
4	36	1.177	0.221	120	1.110	0.301	.074	-0.067
5	229	1.010	0.234	323	1.002	0.240	.355	-0.008
6	660	0.980	0.237	1084	0.982	0.221	.583	0.002
7	494	0.922	0.238	1268	0.911	0.228	.178	-0.012
8	204	0.971	0.250	877	0.944	0.252	.087	-0.026
9	161	0.756	0.190	513	0.740	0.200	.181	-0.016
10	491	0.727	0.173	1394	0.723	0.179	.331	-0.004
11	353	0.862	0.223	626	0.882	0.214	.920	0.021
12	188	0.955	0.201	891	0.943	0.216	.236	-0.012
Group	318	0.923	0.217	839	0.909	0.226		-0.015

Symbol \* indicated statistical significance

Table B-2. Summary of the comparisons between the mean  $\bar{a}_t$  for ACC events during summer and winter per driver including the groups averages. From the difference between the means ( $\Delta = \mu_{\text{winter}} - \mu_{\text{summer}}$ ), a negative value indicated that the mean acceleration was lower during winter. The relationship between  $\Delta$  and the annual mileage was  $R = -0.110$  ( $p = .735$ ).

DriverID	Summer			Winter			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	314	0.947	0.203	475	0.865	0.211	<.001*	-0.082
2	1198	0.916	0.236	928	0.896	0.236	.025*	-0.020
3	72	0.805	0.196	424	0.889	0.211	.999	0.084
4	67	1.107	0.236	108	1.090	0.287	.337	-0.017
5	251	1.008	0.236	186	0.978	0.240	.100	-0.030
6	978	0.976	0.237	413	0.996	0.232	.931	0.020
7	791	0.934	0.224	777	0.895	0.224	<.001*	-0.039
8	232	0.982	0.240	517	0.931	0.248	.004*	-0.051
9	260	0.762	0.202	363	0.741	0.195	.093	-0.021
10	1008	0.738	0.173	806	0.709	0.181	<.001*	-0.030
11	516	0.874	0.263	403	0.894	0.221	.893	0.020
12	417	0.959	0.198	540	0.941	0.218	.085	-0.019
Group	509	0.917	0.220	495	0.902	0.225		-0.015

Symbol \* indicated statistical significance

Table B-3. Summary of the comparisons between the mean  $\bar{a}_t$  for ACC events during off-peak and peak traffic per driver including the groups averages. From the difference between the means ( $\Delta = \mu_{\text{peak}} - \mu_{\text{offpeak}}$ ), a negative value indicated that the mean acceleration was lower during peak traffic. The relationship between  $\Delta$  and the annual mileage was  $R = 0.313$  ( $p = .322$ ).

DriverID	Off-Peak			Peak			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	1203	0.906	0.205	371	0.898	0.205	.240	-0.009
2	2871	0.896	0.229	1461	0.907	0.240	.931	0.011
3	987	0.890	0.212	317	0.913	0.219	.949	0.023
4	250	1.121	0.285	40	1.115	0.302	.450	-0.006
5	813	1.024	0.254	228	1.018	0.245	.370	-0.006
6	2746	0.973	0.222	638	0.985	0.230	.878	0.012
7	2072	0.914	0.229	770	0.914	0.224	.512	0.000
8	1721	0.976	0.255	413	0.939	0.242	.003*	-0.037
9	1098	0.748	0.192	80	0.731	0.191	.218	-0.017
10	3289	0.730	0.175	548	0.761	0.185	.999	0.031
11	1459	0.890	0.224	408	0.871	0.273	.093	-0.020
12	1686	0.969	0.218	250	0.935	0.211	.009*	-0.034
Group	1683	0.920	0.225	460	0.915	0.231		-0.004

Symbol \* indicated statistical significance

Table B-4. Summary of the comparisons between the mean  $\bar{a}_t$  for ACC events during daytime and nighttime per driver including the groups averages. From the difference between the means ( $\Delta = \mu_{\text{nighttime}} - \mu_{\text{daytime}}$ ), a negative value indicated that the mean acceleration was lower during nighttime. The relationship between  $\Delta$  and the annual mileage was  $R = 0.273$  ( $p = .392$ ).

DriverID	Daytime			Nighttime			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	1300	0.903	0.205	185	0.920	0.198	.858	0.017
2	3492	0.903	0.229	434	0.887	0.213	.078	-0.016
3	939	0.894	0.213	210	0.872	0.201	.083	-0.022
4	259	1.117	0.286	29	1.176	0.286	.848	0.058
5	901	1.025	0.254	97	0.985	0.230	.056	-0.040
6	2378	0.981	0.222	745	0.960	0.225	.016*	-0.020
7	2130	0.922	0.226	475	0.877	0.228	<.001*	-0.046
8	1700	0.984	0.254	297	0.909	0.224	<.001*	-0.075
9	988	0.744	0.195	150	0.766	0.164	.935	0.022
10	3441	0.738	0.177	286	0.684	0.166	<.001*	-0.054
11	1637	0.886	0.238	158	0.885	0.224	.484	-0.001
12	1837	0.966	0.217	62	0.883	0.176	<.001*	-0.083
Group	1750	0.922	0.226	261	0.900	0.211		-0.021

Symbol \* indicated statistical significance

## B.2 Deceleration Events

Table B-5. Summary of the comparisons between the mean  $\bar{a}_t$  for DEC events during clement and inclement weather per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{inclement}} - \mu_{\text{clement}}$ ), a positive value indicated that the mean deceleration was lower during inclement weather. The relationship between  $\Delta$  and the annual mileage was  $R = -0.438$  ( $p = .155$ ).

DriverID	Clement			Inclement			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	397	-1.163	0.343	1161	-1.090	0.339	.999	0.073
2	1115	-0.953	0.315	2629	-0.925	0.299	.994	0.028
3	118	-1.067	0.326	927	-1.055	0.346	.639	0.011
4	52	-1.120	0.347	222	-1.149	0.370	.300	-0.029
5	340	-0.961	0.296	461	-1.005	0.317	.020*	-0.045
6	1144	-0.974	0.300	2027	-1.005	0.345	.004*	-0.031
7	1008	-0.978	0.332	2439	-0.943	0.334	.998	0.035
8	357	-1.260	0.382	1605	-1.168	0.418	.999	0.092
9	323	-0.972	0.297	895	-0.951	0.285	.864	0.021
10	887	-0.795	0.223	2466	-0.802	0.244	.192	-0.008
11	628	-0.908	0.245	1155	-0.919	0.253	.173	-0.012
12	394	-1.139	0.338	1670	-1.108	0.326	.951	0.031
Group	564	-1.024	0.312	1471	-1.010	0.323		0.014

Symbol \* indicated statistical significance

Table B-6. Summary of the comparisons between the mean  $\bar{a}_t$  for DEC events during summer and winter per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{winter}} - \mu_{\text{summer}}$ ), a positive value indicated that the mean deceleration was lower during winter. The relationship between  $\Delta$  and the annual mileage was  $R = -0.413$  ( $p = .182$ ).

DriverID	Summer			Winter			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	597	-1.157	0.345	771	-1.089	0.346	.999	0.069
2	1749	-0.931	0.298	1371	-0.923	0.299	.762	0.008
3	123	-1.017	0.310	725	-1.051	0.349	.134	-0.034
4	100	-1.083	0.341	186	-1.165	0.372	.031	-0.082
5	407	-1.008	0.320	252	-0.992	0.330	.737	0.017
6	1767	-0.959	0.298	802	-1.033	0.343	<.001*	-0.073
7	1571	-0.963	0.319	1501	-0.928	0.329	.998	0.034
8	422	-1.294	0.511	984	-1.152	0.396	<.001*	0.142
9	503	-0.962	0.291	571	-0.972	0.289	.285	-0.010
10	1769	-0.805	0.252	1397	-0.802	0.243	.649	0.003
11	904	-0.932	0.272	765	-0.915	0.262	.891	0.016
12	863	-1.153	0.357	993	-1.104	0.317	.999	0.050
Group	898	-1.022	0.326	860	-1.010	0.323		0.012

Symbol \* indicated statistical significance

Table B-7. Summary of the comparisons between the mean  $\bar{a}_t$  for DEC events during off-peak and peak traffic per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{peak}} - \mu_{\text{offpeak}}$ ), a positive value indicated that the mean deceleration was lower during peak traffic. The relationship between  $\Delta$  and the annual mileage was  $R = -0.188$  ( $p = .559$ ).

DriverID	Off-Peak			Peak			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	2128	-1.118	0.340	686	-1.086	0.360	.979	0.032
2	4358	-0.929	0.299	2126	-0.936	0.300	.211	-0.006
3	1637	-1.061	0.331	540	-1.049	0.336	.771	0.012
4	404	-1.144	0.342	63	-1.106	0.453	.736	0.038
5	1315	-1.026	0.326	303	-0.989	0.326	.960	0.036
6	5141	-0.992	0.327	1195	-0.968	0.312	.992	0.024
7	4133	-0.970	0.339	1501	-0.937	0.322	.999	0.033
8	3195	-1.210	0.401	769	-1.152	0.347	.999	0.059
9	1904	-0.966	0.286	165	-0.988	0.278	.163	-0.022
10	5698	-0.806	0.231	976	-0.820	0.264	.065	-0.014
11	2708	-0.927	0.255	667	-0.937	0.294	.202	-0.010
12	3217	-1.134	0.331	466	-1.099	0.363	.973	0.034
Group	2987	-1.024	0.317	788	-1.006	0.330		0.018

Symbol \* indicated statistical significance

Table B-8. Summary of the comparisons between the mean  $\bar{a}_t$  for DEC events during daytime and nighttime per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{nighttime}} - \mu_{\text{daytime}}$ ), a positive value indicated that the mean deceleration was lower during nighttime. The relationship between  $\Delta$  and the annual mileage was  $R = 0.159$  ( $p = .622$ ).

DriverID	Daytime			Nighttime			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	2352	-1.120	0.340	282	-1.066	0.336	.994	0.053
2	5208	-0.945	0.300	675	-0.862	0.269	.999	0.082
3	1516	-1.063	0.330	371	-1.031	0.339	.947	0.032
4	420	-1.153	0.366	42	-1.020	0.243	.999	0.133
5	1456	-1.028	0.323	115	-0.931	0.366	.997	0.098
6	4553	-1.004	0.312	1345	-0.946	0.371	.999	0.057
7	4185	-0.966	0.333	986	-0.945	0.335	.961	0.021
8	3100	-1.222	0.402	593	-1.100	0.332	.999	0.123
9	1742	-0.973	0.283	250	-0.939	0.299	.954	0.034
10	6030	-0.811	0.237	463	-0.778	0.231	.998	0.033
11	2952	-0.940	0.263	288	-0.820	0.235	.999	0.120
12	3512	-1.134	0.337	108	-1.012	0.259	.999	0.122
Group	3086	-1.030	0.319	460	-0.954	0.301		0.076

Symbol \* indicated statistical significance

### B.3 Acute Right Turn Events

Table B-9. Summary of the comparisons between the mean  $\bar{a}_N$  for acute RT events during clement and inclement weather per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{inclement}} - \mu_{\text{clement}}$ ), a negative value indicated that the mean acceleration was lower during inclement weather. The relationship between  $\Delta$  and the annual mileage was  $R = -0.323$  ( $p = .306$ ).

DriverID	Clement			Inclement			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	166	1.811	0.428	521	1.764	0.428	.111	-0.047
2	178	1.604	0.435	515	1.629	0.457	.739	0.025
3	30	1.369	0.427	223	1.506	0.441	.945	0.137
4	7	2.144	0.397	30	1.861	0.574	.073	-0.284
5	58	1.871	0.579	92	1.904	0.559	.632	0.032
6	386	2.180	0.534	602	2.149	0.530	.188	-0.031
7	197	1.595	0.467	542	1.565	0.485	.227	-0.029
8	90	2.117	0.690	331	1.924	0.593	.008*	-0.194
9	196	1.901	0.560	384	1.730	0.488	<.001*	-0.171
10	413	1.547	0.396	1088	1.552	0.388	.598	0.006
11	142	1.743	0.650	331	1.755	0.605	.574	0.012
12	77	2.038	0.728	286	1.820	0.581	.008*	-0.219
Group	162	1.827	0.524	412	1.763	0.511		-0.064

Symbol \* indicated statistical significance

Table B-10. Summary of the comparisons between the mean  $\bar{a}_N$  for acute RT events during summer and winter per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{winter}} - \mu_{\text{summer}}$ ), a negative value indicated that the mean acceleration was lower during winter. The relationship between  $\Delta$  and the annual mileage was  $R = 0.365$  ( $p = .243$ ).

DriverID	Summer			Winter			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	245	1.846	0.416	358	1.745	0.417	.002*	-0.101
2	275	1.668	0.432	273	1.636	0.463	.203	-0.032
3	31	1.395	0.450	174	1.584	0.460	.981	0.189
4	9	1.614	0.527	27	1.808	0.568	.819	0.195
5	80	2.042	0.620	47	1.890	0.533	.075	-0.151
6	605	2.147	0.526	231	2.191	0.533	.861	0.045
7	342	1.566	0.500	312	1.533	0.455	.184	-0.034
8	94	2.223	0.652	230	1.924	0.597	<.001*	-0.298
9	343	1.909	0.536	212	1.663	0.474	<.001*	-0.246
10	731	1.525	0.411	645	1.548	0.385	.861	0.023
11	232	1.755	0.588	223	1.761	0.549	.547	0.006
12	150	2.196	0.742	171	1.743	0.544	<.001*	-0.454
Group	261	1.824	0.533	242	1.752	0.498		-0.072

Symbol \* indicated statistical significance

Table B-11. Summary of the comparisons between the mean  $\bar{a}_N$  for acute RT events during off-peak and peak traffic per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{peak}} - \mu_{\text{offpeak}}$ ), a negative value indicated that the mean acceleration was lower during peak traffic. The relationship between  $\Delta$  and the annual mileage was  $R = -0.114$  ( $p = .725$ ).

DriverID	Off-Peak			Peak			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	951	1.803	0.429	241	1.829	0.421	.801	0.026
2	803	1.614	0.460	351	1.660	0.426	.947	0.045
3	411	1.499	0.467	123	1.551	0.452	.870	0.053
4	55	2.000	0.625	8	1.962	0.807	.450	-0.038
5	266	1.978	0.603	45	1.808	0.629	.048*	-0.170
6	1572	2.134	0.545	347	2.134	0.497	.489	-0.001
7	948	1.573	0.479	266	1.566	0.441	.410	-0.007
8	710	2.006	0.613	156	1.975	0.579	.271	-0.032
9	887	1.786	0.515	110	1.872	0.486	.958	0.086
10	2438	1.536	0.390	397	1.588	0.415	.989	0.052
11	707	1.786	0.577	190	1.768	0.642	.364	-0.018
12	545	1.871	0.621	94	2.307	0.751	.999	0.436
Group	858	1.799	0.527	194	1.835	0.545		0.036

Symbol \* indicated statistical significance

Table B-12. Summary of the comparisons between the mean  $\bar{a}_N$  for acute RT events during daytime and nighttime per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{nighttime}} - \mu_{\text{daytime}}$ ), a negative value indicated that the mean acceleration was lower during nighttime. The relationship between  $\Delta$  and the annual mileage was  $R = -0.276$  ( $p = .386$ ).

DriverID	Daytime			Nighttime			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	977	1.815	0.427	144	1.711	0.428	.003*	-0.104
2	840	1.640	0.454	147	1.549	0.451	.012*	-0.091
3	366	1.534	0.490	96	1.447	0.379	.030*	-0.088
4	55	2.036	0.656	8	1.714	0.486	.062	-0.322
5	274	1.947	0.605	22	1.953	0.684	.516	0.006
6	1363	2.134	0.548	436	2.145	0.514	.647	0.011
7	884	1.591	0.499	238	1.490	0.364	<.001*	-0.101
8	618	2.023	0.636	187	1.916	0.529	.011*	-0.108
9	873	1.807	0.512	98	1.704	0.493	.026*	-0.104
10	2541	1.539	0.401	213	1.568	0.317	.895	0.029
11	808	1.778	0.596	52	1.779	0.540	.509	0.002
12	621	1.936	0.666	12	1.924	0.463	.466	-0.012
Group	852	1.815	0.541	138	1.742	0.471		-0.073

Symbol \* indicated statistical significance

#### B.4 Obtuse Right Turn Events

Table B-13. Summary of the comparisons between the mean  $\bar{a}_N$  for obtuse RT events during clement and inclement weather per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{inclement}} - \mu_{\text{clement}}$ ), a negative value indicated that the mean acceleration was lower during inclement weather. The relationship between  $\Delta$  and the annual mileage was  $R = -0.077$  ( $p = .813$ ).

DriverID	Clement			Inclement			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	72	2.051	0.519	247	1.997	0.527	.223	-0.053
2	154	1.903	0.625	364	1.871	0.575	.293	-0.032
3	15	1.675	0.512	101	1.626	0.576	.368	-0.049
4	4	2.507	0.623	22	2.379	0.832	.367	-0.128
5	27	2.011	0.757	48	1.934	0.571	.324	-0.077
6	94	2.222	0.685	157	2.254	0.620	.642	0.032
7	62	1.816	0.643	193	1.819	0.707	.511	0.003
8	31	2.623	0.889	140	2.028	0.793	.001*	-0.596
9	110	2.621	0.657	206	2.382	0.781	.002*	-0.239
10	118	1.696	0.516	303	1.662	0.478	.270	-0.034
11	78	2.308	0.687	171	2.089	0.819	.015*	-0.220
12	44	2.157	0.629	206	2.053	0.672	.164	-0.104
Group	67	2.133	0.645	180	2.008	0.663		-0.125

Symbol \* indicated statistical significance

Table B-14. Summary of the comparisons between the mean  $\bar{a}_N$  for obtuse RT events during summer and winter per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{winter}} - \mu_{\text{summer}}$ ), a negative value indicated that the mean acceleration was lower during winter. The relationship between  $\Delta$  and the annual mileage was  $R = 0.393$  ( $p = .207$ ).

DriverID	Summer			Winter			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	125	2.091	0.531	160	1.969	0.528	.027*	-0.122
2	230	1.922	0.553	194	1.834	0.529	.049*	-0.087
3	12	1.658	0.599	73	1.546	0.601	.279	-0.112
4	3	2.349	0.180	20	2.598	0.803	.876	0.249
5	37	2.084	0.642	24	2.047	0.709	.420	-0.036
6	153	2.217	0.642	62	2.277	0.693	.721	0.060
7	102	1.778	0.684	120	1.813	0.654	.653	0.036
8	35	2.694	0.725	99	1.952	0.767	<.001*	-0.742
9	177	2.658	0.660	115	2.236	0.765	<.001*	-0.423
10	208	1.706	0.503	184	1.679	0.450	.288	-0.027
11	132	2.221	0.741	128	2.077	0.852	.074	-0.144
12	90	2.199	0.683	116	2.068	0.647	.081	-0.131
Group	109	2.131	0.595	108	2.008	0.666		-0.123

Symbol \* indicated statistical significance

Table B-15. Summary of the comparisons between the mean  $\bar{a}_N$  for obtuse RT events during off-peak and peak traffic per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{peak}} - \mu_{\text{offpeak}}$ ), a negative value indicated that the mean acceleration was lower during peak traffic. The relationship between  $\Delta$  and the annual mileage was  $R = 0.509$  ( $p = .091$ ).

DriverID	Off-Peak			Peak			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	423	2.020	0.528	104	2.006	0.559	.408	-0.014
2	626	1.859	0.602	285	1.962	0.560	.994	0.103
3	168	1.664	0.630	54	1.631	0.542	.355	-0.033
4	36	2.525	0.769	6	2.923	0.568	.915	0.398
5	118	2.085	0.629	23	2.427	0.779	.971	0.342
6	432	2.219	0.652	100	2.253	0.588	.694	0.034
7	320	1.843	0.714	101	1.974	0.740	.939	0.130
8	270	2.349	0.842	53	2.150	0.810	.054	-0.200
9	523	2.494	0.730	52	2.505	0.629	.545	0.011
10	685	1.667	0.497	101	1.765	0.496	.967	0.098
11	393	2.242	0.805	92	2.148	0.710	.135	-0.093
12	381	2.130	0.681	33	2.052	0.557	.225	-0.079
Group	365	2.091	0.673	84	2.150	0.628		0.058

Symbol \* indicated statistical significance

Table B-16. Summary of the comparisons between the mean  $\bar{a}_N$  for obtuse RT events during daytime and nighttime per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{nighttime}} - \mu_{\text{daytime}}$ ), a negative value indicated that the mean acceleration was lower during nighttime. The relationship between  $\Delta$  and the annual mileage was  $R = 0.696$  ( $p = .012$ ).

DriverID	Daytime			Nighttime			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	443	2.047	0.516	53	1.804	0.566	.002*	-0.243
2	695	1.913	0.599	121	1.842	0.559	.104	-0.070
3	146	1.664	0.629	42	1.615	0.586	.318	-0.049
4	40	2.543	0.738	2	3.364	0.791	.814	0.821
5	124	2.143	0.678	10	2.028	0.592	.284	-0.116
6	399	2.249	0.652	92	2.171	0.613	.139	-0.078
7	295	1.925	0.765	85	1.670	0.535	<.001*	-0.255
8	231	2.439	0.871	63	2.018	0.647	<.001*	-0.421
9	495	2.505	0.721	60	2.403	0.708	.149	-0.102
10	706	1.686	0.503	62	1.647	0.416	.243	-0.039
11	429	2.242	0.792	27	1.920	0.728	.017*	-0.322
12	399	2.132	0.673	7	1.717	0.580	.054	-0.415
Group	367	2.124	0.678	52	2.016	0.610		-0.107

Symbol \* indicated statistical significance

### B.5 Acute Left Turn Events

Table B-17. Summary of the comparisons between the mean  $\bar{a}_N$  for acute LT events during clement and inclement weather per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{inclement}} - \mu_{\text{clement}}$ ), a negative value indicated that the mean acceleration was lower during inclement weather. The relationship between  $\Delta$  and the annual mileage was  $R = 0.571$  ( $p = .052$ ).

DriverID	Clement			Inclement			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	148	1.797	0.441	426	1.693	0.438	.007*	-0.104
2	188	1.461	0.405	393	1.562	0.417	.997	0.101
3	29	1.523	0.407	187	1.512	0.432	.443	-0.012
4	13	1.812	0.464	29	2.031	0.537	.905	0.219
5	36	1.548	0.540	84	1.540	0.534	.470	-0.008
6	236	1.805	0.479	386	1.846	0.494	.852	0.042
7	107	1.638	0.535	350	1.633	0.544	.467	-0.005
8	51	2.121	0.695	213	1.919	0.683	.033*	-0.201
9	163	1.757	0.528	333	1.641	0.529	.011*	-0.116
10	364	1.491	0.455	978	1.508	0.433	.726	0.017
11	114	1.545	0.653	254	1.520	0.622	.361	-0.026
12	41	1.805	0.607	191	1.729	0.653	.238	-0.076
Group	124	1.692	0.517	319	1.678	0.526		-0.014

Symbol \* indicated statistical significance

Table B-18. Summary of the comparisons between the mean  $\bar{a}_N$  for acute LT events during summer and winter per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{winter}} - \mu_{\text{summer}}$ ), a negative value indicated that the mean acceleration was lower during winter. The relationship between  $\Delta$  and the annual mileage was  $R = 0.615$  ( $p = .033$ ).

DriverID	Summer			Winter			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	243	1.799	0.447	280	1.699	0.432	.005*	-0.100
2	237	1.525	0.409	208	1.560	0.415	.816	0.035
3	18	1.538	0.350	153	1.563	0.448	.609	0.025
4	12	1.801	0.571	24	2.187	0.530	.968	0.386
5	63	1.517	0.450	54	1.433	0.483	.166	-0.085
6	365	1.837	0.500	140	1.835	0.478	.482	-0.002
7	197	1.680	0.502	200	1.580	0.557	.032*	-0.099
8	67	2.250	0.618	130	1.866	0.699	<.001*	-0.384
9	274	1.737	0.552	186	1.567	0.465	<.001*	-0.170
10	630	1.491	0.434	597	1.520	0.431	.875	0.028
11	161	1.553	0.685	173	1.570	0.662	.590	0.017
12	85	1.889	0.504	118	1.658	0.686	.003*	-0.231
Group	196	1.718	0.502	189	1.670	0.524		-0.048

Symbol \* indicated statistical significance

Table B-19. Summary of the comparisons between the mean  $\bar{a}_N$  for acute LT events during off-peak and peak traffic per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{peak}} - \mu_{\text{offpeak}}$ ), a negative value indicated that the mean acceleration was lower during peak traffic. The relationship between  $\Delta$  and the annual mileage was  $R = -0.356$  ( $p = .256$ ).

DriverID	Off-Peak			Peak			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	834	1.747	0.441	219	1.727	0.458	.281	-0.020
2	661	1.525	0.429	286	1.553	0.421	.821	0.028
3	331	1.554	0.432	84	1.473	0.438	.067	-0.081
4	56	2.009	0.608	5	1.500	0.482	.038*	-0.509
5	207	1.609	0.531	39	1.479	0.395	.040*	-0.130
6	940	1.830	0.494	193	1.793	0.454	.155	-0.037
7	573	1.673	0.551	169	1.623	0.496	.128	-0.051
8	450	2.026	0.708	89	1.965	0.685	.224	-0.061
9	766	1.704	0.540	89	1.758	0.520	.820	0.054
10	2165	1.497	0.437	331	1.602	0.451	.999	0.105
11	541	1.643	0.665	125	1.482	0.618	.005*	-0.160
12	356	1.762	0.668	33	1.813	0.634	.670	0.051
Group	657	1.715	0.542	139	1.647	0.504		-0.068

Symbol \* indicated statistical significance

Table B-20. Summary of the comparisons between the mean  $\bar{a}_N$  for acute LT events during daytime and nighttime per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{nighttime}} - \mu_{\text{daytime}}$ ), a negative value indicated that the mean acceleration was lower during nighttime. The relationship between  $\Delta$  and the annual mileage was  $R = 0.354$  ( $p = .259$ ).

DriverID	Daytime			Nighttime			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	849	1.763	0.441	132	1.621	0.403	<.001*	-0.141
2	698	1.557	0.440	121	1.493	0.393	.054	-0.064
3	277	1.564	0.449	92	1.497	0.410	.092	-0.067
4	51	1.979	0.615	10	1.908	0.622	.373	-0.071
5	211	1.592	0.499	27	1.508	0.616	.251	-0.084
6	811	1.846	0.505	244	1.766	0.430	.008*	-0.080
7	513	1.669	0.542	168	1.608	0.535	.100	-0.061
8	373	2.016	0.743	115	1.978	0.601	.290	-0.038
9	736	1.720	0.539	90	1.596	0.496	.014*	-0.124
10	2151	1.507	0.445	280	1.529	0.399	.802	0.022
11	587	1.648	0.656	43	1.322	0.570	<.001*	-0.326
12	373	1.771	0.669	12	1.598	0.558	.158	-0.172
Group	636	1.719	0.545	111	1.619	0.503		-0.101

Symbol \* indicated statistical significance

### B.6 Obtuse Left Turn Events

Table B-21. Summary of the comparisons between the mean  $\bar{a}_N$  for obtuse LT events during clement and inclement weather per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{inclement}} - \mu_{\text{clement}}$ ), a negative value indicated that the mean acceleration was lower during inclement weather. The relationship between  $\Delta$  and the annual mileage was  $R = 0.485$  ( $p = .110$ ).

DriverID	Clement			Inclement			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	78	2.322	0.584	237	2.106	0.651	.003*	-0.215
2	170	1.416	0.498	377	1.429	0.504	.612	0.013
3	8	1.605	0.565	46	1.676	0.575	.624	0.071
4	4	1.882	0.479	8	2.311	0.466	.905	0.429
5	14	1.739	0.587	22	1.605	0.534	.248	-0.134
6	112	2.025	0.496	183	2.073	0.553	.783	0.049
7	63	1.916	0.549	169	1.823	0.636	.137	-0.093
8	18	2.430	0.799	84	2.312	0.732	.284	-0.118
9	38	1.964	0.631	83	2.058	0.634	.775	0.094
10	112	1.689	0.514	365	1.650	0.516	.244	-0.039
11	39	1.540	0.680	85	1.711	0.768	.892	0.171
12	37	2.004	0.801	149	2.116	0.834	.772	0.112
Group	58	1.878	0.599	151	1.906	0.617		0.028

Symbol \* indicated statistical significance

Table B-22. Summary of the comparisons between the mean  $\bar{a}_N$  for obtuse LT events during summer and winter per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{winter}} - \mu_{\text{summer}}$ ), a negative value indicated that the mean acceleration was lower during winter. The relationship between  $\Delta$  and the annual mileage was  $R = 0.460$  ( $p = .132$ ).

DriverID	Summer			Winter			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	110	2.306	0.568	150	2.096	0.652	.003*	-0.210
2	257	1.455	0.492	190	1.426	0.505	.274	-0.029
3	3	1.731	0.921	36	1.651	0.588	.448	-0.080
4	4	2.137	0.464	6	2.325	0.486	.720	0.187
5	17	1.667	0.636	17	1.614	0.486	.393	-0.053
6	181	2.083	0.552	61	1.990	0.485	.106	-0.093
7	101	1.759	0.637	113	1.838	0.607	.821	0.079
8	23	2.475	0.623	58	2.243	0.702	.077	-0.231
9	56	2.009	0.648	52	2.021	0.634	.538	0.012
10	209	1.674	0.499	204	1.633	0.504	.203	-0.041
11	54	1.881	0.746	55	1.700	0.745	.104	-0.181
12	62	2.168	0.749	90	2.175	1.143	.519	0.007
Group	90	1.945	0.628	86	1.893	0.628		-0.053

Symbol \* indicated statistical significance

Table B-23. Summary of the comparisons between the mean  $\bar{a}_N$  for obtuse LT events during off-peak and peak traffic per driver including group averages. From the difference between the means ( $\Delta = \mu_{\text{peak}} - \mu_{\text{offpeak}}$ ), a negative value indicated that the mean acceleration was lower during peak traffic. The relationship between  $\Delta$  and the annual mileage was  $R = -0.600$  ( $p = .039$ ).

DriverID	Off-Peak			Peak			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	402	2.185	0.610	113	2.248	0.614	.832	0.063
2	773	1.412	0.484	192	1.429	0.498	.661	0.017
3	82	1.664	0.545	25	1.680	0.519	.553	0.016
4	21	2.329	0.552	2	1.610	0.093	<.001*	-0.719
5	57	1.761	0.569	11	1.692	0.527	.351	-0.069
6	463	2.084	0.530	99	1.995	0.540	.070	-0.089
7	276	1.889	0.669	101	1.923	0.578	.687	0.034
8	173	2.191	0.820	42	2.517	0.625	.997	0.326
9	193	1.968	0.635	15	2.084	0.714	.725	0.116
10	729	1.648	0.528	135	1.683	0.530	.759	0.035
11	164	1.769	0.762	48	1.774	0.781	.514	0.005
12	259	2.133	0.830	36	2.439	1.413	.894	0.306
Group	299	1.920	0.628	68	1.923	0.619		0.003

Symbol \* indicated statistical significance

Table B-24. Summary of the comparisons between the mean  $\bar{a}_N$  for obtuse LT events during daytime and nighttime per driver including the group averages. From the difference between the means ( $\Delta = \mu_{\text{nighttime}} - \mu_{\text{daytime}}$ ), a negative value indicated that the mean acceleration was lower during nighttime. The relationship between  $\Delta$  and the annual mileage was  $R = -0.218$  ( $p = 0.497$ ).

DriverID	Daytime			Nighttime			P-value	$\Delta$ (m/s <sup>2</sup> )
	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )	N	$\mu$ (m/s <sup>2</sup> )	$\sigma$ (m/s <sup>2</sup> )		
1	396	2.213	0.625	85	2.055	0.535	.009*	-0.158
2	733	1.452	0.477	123	1.246	0.450	<.001*	-0.206
3	82	1.679	0.544	12	1.662	0.634	.467	-0.017
4	21	2.292	0.572	2	1.999	0.605	.308	-0.293
5	60	1.727	0.554	6	1.856	0.706	.659	0.128
6	410	2.089	0.550	116	2.022	0.479	.100	-0.067
7	293	1.898	0.633	51	1.792	0.685	.152	-0.106
8	139	2.106	0.752	49	2.509	0.832	.998	0.403
9	167	1.926	0.623	29	2.103	0.648	.910	0.177
10	780	1.662	0.528	63	1.539	0.534	.041*	-0.123
11	182	1.790	0.763	19	1.563	0.653	.085	-0.227
12	282	2.159	0.848	6	1.662	0.696	.071	-0.496
Group	295	1.916	0.622	47	1.834	0.621		-0.082

Symbol \* indicated statistical significance

## Appendix C : Classification Performance

### C.1 Event Classifiers

Table C-1. Classification performance of the ACC event classifier using LR and RF for subsets of 2 to 12 drivers from a pool of 12 drivers.

Subsets	LR (%)					RF (%)					I	Null (%)
	Score	2 $\sigma$	PC <sub>M</sub>	RC <sub>M</sub>	F <sub>M</sub>	Score	2 $\sigma$	PC <sub>M</sub>	RC <sub>M</sub>	F <sub>M</sub>		
2	52.6	1.3	53.2	52.6	52.9	62.6	1.5	65.0	62.6	63.8	5577	50.0
3	35.4	1.4	36.2	35.4	35.8	43.7	1.3	45.0	43.7	44.3	6804	33.3
4	31.7	4.0	31.4	31.7	31.5	37.6	1.5	48.2	37.6	42.2	7068	25.0
5	26.0	3.5	26.7	26.0	26.3	32.4	1.0	41.0	32.4	36.2	8063	20.0
6	21.7	3.8	22.8	21.7	22.2	40.7	0.7	46.2	40.7	43.3	11194	16.7
7	18.3	1.0	19.7	18.3	18.9	35.7	1.1	40.0	35.7	37.7	13881	14.3
8	16.6	2.6	16.9	16.6	16.7	31.6	0.9	34.6	31.6	33.0	15824	12.5
9	15.4	2.0	16.5	15.4	15.9	30.2	0.9	34.5	30.2	32.2	16938	11.1
10	14.1	2.2	13.3	14.1	13.7	30.9	0.9	33.6	30.9	32.2	20526	10.0
11	13.4	2.5	12.9	13.4	13.0	30.3	0.5	32.7	30.3	31.5	22295	9.1
12	12.5	2.1	12.8	12.5	12.6	29.8	0.5	32.5	29.8	31.1	24148	8.3

Table C-2. Classification performance of the DEC event classifier using LR and RF for subsets of 2 to 12 drivers from a pool of 12 drivers.

Subsets	LR (%)					RF (%)					I	Null (%)
	Score	2 $\sigma$	PC <sub>M</sub>	RC <sub>M</sub>	F <sub>M</sub>	Score	2 $\sigma$	PC <sub>M</sub>	RC <sub>M</sub>	F <sub>M</sub>		
2	69.7	4.0	70.2	69.7	69.9	74.6	0.8	75.7	74.6	75.1	8174	50.0
3	49.2	3.2	49.0	49.2	49.1	56.9	0.7	60.0	56.9	58.4	10057	33.3
4	38.2	1.9	38.9	38.2	38.5	43.8	1.0	50.7	43.8	47.0	10478	25.0
5	30.5	1.5	30.9	30.5	30.6	36.4	1.0	42.0	36.4	39.0	11893	20.0
6	26.1	3.7	27.5	26.1	26.8	33.3	0.9	37.0	33.3	35.0	17529	16.7
7	22.7	2.2	26.6	22.7	24.4	29.7	0.5	32.4	29.7	31.0	22601	14.3
8	21.5	1.8	21.6	21.5	21.5	27.8	0.7	30.1	27.8	28.9	26072	12.5
9	19.1	1.3	20.8	19.1	19.9	25.4	0.8	27.6	25.4	26.4	27887	11.1
10	19.1	2.2	18.7	19.1	18.9	25.9	0.6	27.7	25.9	26.8	33637	10.0
11	17.3	2.7	17.5	17.3	17.4	24.2	0.5	25.8	24.2	25.0	36536	9.1
12	16.4	1.8	16.9	16.4	16.6	23.6	0.5	25.1	23.6	24.3	39845	8.3

Table C-3. Classification performance of the RT event classifier using LR and RF for subsets of 2 to 12 drivers from a pool of 12 drivers.

Subsets	LR (%)					RF (%)					I	Null (%)
	Score	2 $\sigma$	PC <sub>M</sub>	RC <sub>M</sub>	F <sub>M</sub>	Score	2 $\sigma$	PC <sub>M</sub>	RC <sub>M</sub>	F <sub>M</sub>		
2	52.7	3.5	52.9	52.7	52.8	70.7	0.8	71.2	70.7	70.9	8174	50.0
3	34.8	1.7	36.1	34.8	35.4	52.2	0.8	53.9	52.2	53.0	10057	33.3
4	29.0	3.7	28.7	29.0	28.7	41.4	0.9	53.0	41.4	46.5	10478	25.0
5	22.1	1.9	22.6	22.1	22.3	36.9	1.8	46.7	36.9	41.2	11893	20.0
6	18.4	2.6	18.1	18.4	18.1	36.1	0.9	42.2	36.1	38.9	17529	16.7
7	15.7	1.7	16.0	15.7	15.7	32.8	0.9	38.1	32.8	35.2	22601	14.3
8	14.1	1.7	13.5	14.1	13.6	30.8	0.9	34.4	30.8	32.5	26072	12.5
9	12.2	1.3	11.7	12.2	11.9	31.0	0.9	34.4	31.0	32.6	27887	11.1
10	11.0	1.3	8.9	11.0	9.7	31.3	0.6	34.2	31.3	32.7	33637	10.0
11	10.4	1.9	10.8	10.4	10.5	29.2	0.8	31.9	29.2	30.5	36536	9.1
12	9.2	0.9	8.5	9.2	8.8	28.1	0.5	30.9	28.1	29.4	39845	8.3

Table C-4. Classification performance of the LT event classifier using LR and RF for subsets of 2 to 12 drivers from a pool of 12 drivers.

Subsets	LR (%)					RF (%)					I	Null (%)
	Score	2 $\sigma$	PC <sub>M</sub>	RC <sub>M</sub>	F <sub>M</sub>	Score	2 $\sigma$	PC <sub>M</sub>	RC <sub>M</sub>	F <sub>M</sub>		
2	52.7	2.5	53.3	52.7	53.0	76.5	1.3	77.3	76.5	76.9	5901	50.0
3	38.0	4.6	38.5	38.0	38.2	56.4	0.8	58.7	56.4	57.5	7047	33.3
4	27.4	1.6	28.5	27.4	27.8	43.5	1.3	54.8	43.5	48.5	7173	25.0
5	23.0	2.8	24.2	23.0	23.5	39.6	1.2	50.1	39.6	44.2	7764	20.0
6	18.8	2.7	19.3	18.8	18.9	39.1	1.2	47.2	39.1	42.8	10347	16.7
7	15.7	2.0	15.6	15.7	15.4	34.9	1.1	39.3	34.9	36.9	12900	14.3
8	13.3	1.5	11.0	13.3	11.9	32.3	0.7	35.7	32.3	33.9	14187	12.5
9	12.3	1.2	12.1	12.3	12.1	31.9	0.8	35.4	31.9	33.6	16352	11.1
10	11.5	1.6	11.1	11.5	11.2	32.0	0.8	34.7	32.0	33.3	21508	10.0
11	10.1	1.3	9.9	10.1	9.9	30.4	0.5	33.1	30.4	31.7	23342	9.1
12	9.2	1.4	8.3	9.2	8.6	28.9	0.5	31.2	28.9	30.0	24414	8.3

## C.2 Expert Trip Classifiers

Table C-5. Classification performance of the expert ACC trip classifier using MV among the ACC events from the same trip

Subsets	Score (%)	2 $\sigma$ (%)	PC <sub>M</sub> (%)	RC <sub>M</sub> (%)	F <sub>M</sub> (%)	N (%)	Null (%)
2	66.9	6.2	67.8	66.0	66.9	2183	50.0
3	45.1	7.2	47.2	45.7	46.5	2611	33.3
4	37.3	5.4	49.3	37.4	42.3	2696	25.0
5	32.2	8.1	45.5	33.9	38.8	3093	20.0
6	42.0	6.2	49.8	42.3	45.7	4641	16.7
7	36.7	5.6	42.8	37.2	39.8	5645	14.3
8	32.8	5.8	38.6	34.0	36.1	6331	12.5
9	32.9	2.9	39.7	33.4	36.3	6895	11.1
10	34.8	4.1	37.8	34.3	36.0	8428	10.0
11	34.4	4.4	38.0	34.9	36.4	9193	9.1
12	34.6	8.0	39.8	35.0	37.2	9922	8.3

Table C-6. Classification performance of the expert DEC trip classifier using MV among the DEC events from the same trip

Subsets	Score (%)	2 $\sigma$ (%)	PC <sub>M</sub> (%)	RC <sub>M</sub> (%)	F <sub>M</sub> (%)	N (%)	Null (%)
2	80.7	1.9	81.0	80.5	80.8	2538	50.0
3	62.3	2.8	67.5	61.8	64.5	3052	33.3
4	47.7	4.5	52.5	47.1	49.2	3165	25.0
5	41.3	4.4	45.2	41.1	42.8	3604	20.0
6	38.8	9.4	42.8	38.1	40.2	5489	16.7
7	37.7	6.1	42.3	37.3	39.6	6702	14.3
8	34.9	2.9	37.2	34.9	36.0	7521	12.5
9	32.2	2.5	35.9	32.7	34.2	8260	11.1
10	33.2	4.8	36.4	33.0	34.6	10228	10.0
11	31.0	3.6	34.5	31.8	33.1	11167	9.1
12	31.8	4.0	33.7	31.2	32.4	12070	8.3

Table C-7. Classification performance of the expert RT trip classifier using MV among the RT events from the same trip

Subsets	Score (%)	2 $\sigma$ (%)	PC <sub>M</sub> (%)	RC <sub>M</sub> (%)	F <sub>M</sub> (%)	N (%)	Null (%)
2	75.5	4.0	75.1	75.1	75.1	2562	50.0
3	57.0	3.0	61.2	57.4	59.2	3212	33.3
4	44.7	8.3	45.9	44.8	44.7	3302	25.0
5	38.0	6.7	50.1	39.3	43.9	3699	20.0
6	39.8	4.3	45.2	40.4	42.5	5485	16.7
7	38.9	4.5	44.1	37.7	40.6	6730	14.3
8	35.8	4.5	41.4	35.7	38.3	7434	12.5
9	35.9	4.3	41.8	37.8	39.7	8393	11.1
10	38.4	6.5	41.7	37.9	39.7	10500	10.0
11	36.1	6.0	40.7	36.2	38.3	11524	9.1
12	34.7	4.6	39.6	34.7	37.0	12261	8.3

Table C-8. Classification performance of the expert LT trip classifier using MV among the LT events from the same trip

Subsets	Score (%)	2 $\sigma$ (%)	PC <sub>M</sub> (%)	RC <sub>M</sub> (%)	F <sub>M</sub> (%)	N (%)	Null (%)
2	80.1	3.4	80.0	79.9	79.9	2429	50.0
3	59.6	2.8	63.5	59.5	61.4	3006	33.3
4	46.4	5.3	41.2	46.3	43.2	3079	25.0
5	41.0	6.2	47.2	41.0	43.7	3438	20.0
6	40.4	9.6	43.2	40.5	41.6	4902	16.7
7	38.5	7.2	40.4	38.7	39.5	6050	14.3
8	36.6	3.6	37.5	36.0	36.7	6668	12.5
9	35.4	5.0	36.4	35.7	36.0	7613	11.1
10	35.4	4.0	39.0	36.4	37.6	9649	10.0
11	35.5	5.2	37.0	35.7	36.3	10516	9.1
12	33.1	4.8	36.6	33.4	34.9	11142	8.3

Table C-9. Confusion matrix of the classification performance of the expert ACC trip classifier

Driver	Predicted												Recall (%)	
	1	2	3	4	5	6	7	8	9	10	11	12		
Actual	1	20	27	6	1	7	1	6	8	3	6	0	0	23.5
	2	7	54	2	0	4	0	6	4	3	3	1	1	63.5
	3	11	27	7	0	2	2	11	7	3	12	1	2	8.2
	4	4	23	3	8	6	3	14	13	0	0	7	4	9.4
	5	14	27	4	3	15	1	3	13	0	3	0	2	17.6
	6	0	2	0	0	2	72	1	0	0	0	2	6	84.7
	7	14	27	6	1	4	0	14	4	5	7	0	3	16.5
	8	9	28	1	3	5	1	10	18	1	4	4	1	21.2
	9	7	18	3	0	1	1	6	3	26	18	1	1	30.6
	10	11	5	8	1	1	0	2	1	8	48	0	0	56.5
	11	3	12	1	1	0	14	4	0	1	1	41	7	48.2
	12	4	11	1	0	3	17	0	5	1	0	11	32	37.6
Precision (%)	19.2	20.7	16.7	44.4	30.0	64.3	18.2	23.7	51.0	47.1	60.3	54.2		

Table C-10. Confusion matrix for the classification performance of the expert DEC trip classifier

Driver		Predicted												Recall (%)
		1	2	3	4	5	6	7	8	9	10	11	12	
Actual	1	52	14	2	2	1	16	4	13	2	2	3	2	46.0
	2	12	59	11	0	2	10	8	2	0	8	1	0	52.2
	3	7	16	13	0	8	15	24	6	3	9	9	3	11.5
	4	35	23	4	3	4	16	6	12	2	4	1	3	2.7
	5	12	23	9	0	8	13	15	9	4	6	7	7	7.1
	6	8	15	9	0	1	63	0	5	1	6	4	1	55.8
	7	1	29	4	0	0	15	43	7	1	10	1	2	38.1
	8	23	10	3	0	3	11	5	41	7	4	0	6	36.3
	9	5	15	9	0	7	13	11	7	21	16	5	4	18.6
	10	3	19	3	1	3	7	5	0	5	65	1	1	57.5
	11	2	29	5	0	3	11	13	4	2	15	24	5	21.2
	12	5	11	9	0	2	14	6	10	0	3	6	47	41.6
Precision (%):		31.5	22.4	16.0	50.0	19.0	30.9	30.7	35.3	43.8	43.9	38.7	58.0	

Table C-11. Confusion matrix for the classification performance of the expert RT trip classifier

Driver		Predicted												Recall (%)
		1	2	3	4	5	6	7	8	9	10	11	12	
Actual	1	41	5	3	0	2	6	5	3	8	14	2	1	45.6
	2	10	38	2	0	0	9	5	5	2	10	6	3	42.2
	3	12	16	13	0	2	5	10	3	4	18	5	2	14.4
	4	9	10	3	6	2	17	11	18	1	3	6	4	6.7
	5	11	12	3	1	15	14	6	4	5	11	7	1	16.7
	6	10	17	2	0	2	42	4	3	4	2	1	3	46.7
	7	13	23	5	1	1	4	28	4	1	10	0	0	31.1
	8	11	22	1	2	1	16	6	21	1	3	3	3	23.3
	9	4	7	6	0	0	4	3	1	58	5	2	0	64.4
	10	6	10	5	0	1	2	2	1	1	61	1	0	67.8
	11	10	18	3	0	1	7	9	2	3	8	26	3	28.9
	12	16	8	3	0	2	17	3	8	2	2	8	21	23.3
Precision (%):		26.8	20.4	26.5	60.0	51.7	29.4	30.4	28.8	64.4	41.5	38.8	51.2	

Table C-12. Confusion matrix for the classification performance of the expert LT trip classifier

Driver		Predicted												Recall (%)
		1	2	3	4	5	6	7	8	9	10	11	12	
Actual	1	35	11	3	0	1	2	4	1	1	9	3	3	47.9
	2	5	37	3	0	0	1	4	1	2	13	6	1	50.7
	3	11	18	6	0	2	2	8	0	5	17	4	0	8.2
	4	12	12	1	1	1	11	5	12	4	4	7	3	1.4
	5	19	11	5	0	17	4	6	3	1	6	1	0	23.3
	6	11	8	1	1	4	37	4	3	0	4	0	0	50.7
	7	2	25	3	0	2	4	15	2	9	4	4	3	20.5
	8	9	10	2	0	3	7	8	23	4	3	3	1	31.5

	<b>9</b>	5	8	6	0	0	4	5	2	33	5	4	1	45.2
	<b>10</b>	3	13	1	0	0	0	2	0	3	50	0	1	68.5
	<b>11</b>	6	17	4	0	1	6	8	3	4	4	19	1	26.0
	<b>12</b>	10	15	3	1	1	7	8	6	2	5	4	11	15.1
<b>Precision (%)</b> :		27.3	20.0	15.8	33.3	53.1	43.5	19.5	41.1	48.5	40.3	34.5	44.0	

### C.3 Multi-Expert Trip Classifier

Table C-13. Classification performance of the ME trip classifier showing the results from including the environmental features

Subsets	Without environmental features (%)					With environmental features (%)					N	Null (%)
	Score	2SD	PC <sub>M</sub>	RC <sub>M</sub>	FM	Score	2SD	PC <sub>M</sub>	RC <sub>M</sub>	FM		
2	82.0	1.1	84.1	82.0	83.1	81.7	1.2	83.2	81.3	82.2	3085	50.0
3	65.1	1.4	71.9	64.1	67.8	65.0	1.3	73.3	65.4	69.1	3905	33.3
4	49.9	2.1	66.8	51.9	58.4	50.8	2.8	43.7	52.1	47.5	4038	25.0
5	45.5	2.4	62.3	44.5	51.9	45.9	4.4	59.5	42.3	49.4	4595	20.0
6	49.2	3.4	46.5	47.2	46.9	50.2	2.5	47.3	48.6	48.0	6663	16.7
7	48.4	2.4	59.6	49.0	53.8	48.4	2.7	49.8	50.5	50.1	8014	14.3
8	48.2	1.8	57.7	47.3	52.0	47.6	3.6	55.9	45.9	50.4	8877	12.5
9	48.8	3.3	49.4	50.5	49.9	48.9	2.1	54.1	48.0	50.8	9907	11.1
10	50.0	2.5	52.6	51.0	51.8	48.9	2.0	53.3	49.3	51.2	12210	10.0
11	50.0	1.4	57.4	50.4	53.7	49.2	2.4	46.0	48.2	47.1	13324	9.1
12	49.4	1.0	52.5	50.2	51.3	49.0	1.8	56.3	49.6	52.7	14317	8.3

Table C-14. Confusion matrix for the classification performance of the ME trip classifier

Driver	Predicted												Recall (%)	
	1	2	3	4	5	6	7	8	9	10	11	12		
Actual	<b>1</b>	68	15	2	1	1	6	7	9	5	13	3	3	51.1
	<b>2</b>	5	86	0	0	3	6	10	3	6	12	1	1	64.7
	<b>3</b>	8	32	11	1	2	2	21	7	9	26	10	4	8.3
	<b>4</b>	14	25	4	4	2	23	16	32	2	4	4	3	3.0
	<b>5</b>	7	30	8	0	20	9	14	18	1	16	5	5	15.0
	<b>6</b>	2	5	0	0	1	112	4	1	3	2	0	3	84.2
	<b>7</b>	1	21	4	0	1	5	76	3	0	13	7	2	57.1
	<b>8</b>	10	12	0	1	3	9	12	73	2	5	4	2	54.9
	<b>9</b>	2	8	1	0	1	3	4	3	88	17	5	1	66.2
	<b>10</b>	1	3	2	0	0	1	3	1	3	118	0	1	88.7
	<b>11</b>	1	21	4	0	1	7	5	4	4	15	69	2	51.9
	<b>12</b>	4	14	1	0	0	19	8	5	2	1	3	76	57.1
<b>Precision (%)</b> :		55.3	31.6	29.7	57.1	57.1	55.4	42.2	45.9	70.4	48.8	62.2	73.8	

## References

- [1] Statistics Canada, “The Daily — Canada’s population estimates: Age and sex, July 1, 2015,” *The Daily*, 2015. [Online]. Available: <http://www.statcan.gc.ca/daily-quotidien/150929/dq150929b-eng.htm>.
- [2] S. C. Marshall, “The Role of Reduced Fitness to Drive Due to Medical Impairments in Explaining Crashes Involving Older Drivers,” *Traffic Injury Prevention*, vol. 9, no. 4, pp. 291–298, 2008.
- [3] Ontario Ministry of Transport, “Senior driver’s licence renewal program,” *Driver*, 2017. [Online]. Available: <http://www.mto.gov.on.ca/english/driver/senior-driver-licence-renewal-program.shtml>.
- [4] Government of British Columbia, “Senior Drivers,” *Driver Medical Fitness*, 2017. [Online]. Available: <http://www2.gov.bc.ca/gov/content/transportation/driving-and-cycling/driver-medical/driver-medical-fitness/senior-drivers>.
- [5] S. C. Marshall, *et al.*, “Protocol for Candrive II / Ozcandrive , a multicentre prospective older driver cohort study,” *Accident Anal. and Prevention*, vol. 61, pp. 245–252, 2013.
- [6] A. M. Hollis, H. Duncanson, L. R. Kapust, P. M. Xi, and M. G. O’Connor, “Validity of the Mini-Mental State Examination and the Montreal Cognitive Assessment in the Prediction of Driving Test Outcome,” *Journal of the American Geriatrics Society*, vol. 63, no. 5, pp. 988–992, 2015.
- [7] J. C. W. Kwok, I. Gélinas, D. Benoit, and G. Chilingaryan, “Predictive validity of the Montreal Cognitive Assessment (MoCA) as a screening tool for on-road driving performance,” *British Journal of Occupational Therapy*, vol. 78, no. 2, pp. 100–108, 2015.
- [8] W. Lindstrom-Forneri, H. A. Tuokko, D. Garrett, and F. Molnar, “Driving as an everyday competence: A model of driving competence and behavior,” *Clinical Gerontologist*, vol. 33, no. 4, pp. 283–297, 2010.
- [9] B. Wallace, *et al.*, “Measurement of Distinguishing Features of Stable Cognitive and Physical Health Older Drivers,” *IEEE Trans. Instrum. Meas.*, vol. 65, no. 9, pp. 1990–2001, 2016.
- [10] D. B. Hogan, “Which older patients are competent to drive? Approaches to office-based assessment,” *Canadian Family Physician*, vol. 51, no. March, pp. 362–368, 2005.
- [11] F. J. Molnar, A. M. Byszewski, S. C. Marshall, and M. Man-sun-hing, “In-office evaluation of medical fitness to drive: Practical approaches for assessing older people,” *Canadian Family Physician*, vol. 51, pp. 372–379, 2005.
- [12] Z. Vlahodimitrakou, *et al.*, “Development and evaluation of a Driving Observation Schedule (DOS) to study everyday driving performance of older drivers,” *Accident Anal. and Prevention*, vol. 61, pp. 253–260, 2013.
- [13] T. Imamura, H. Yamashita, R. bin Othman, Z. Zhang, and T. Miyake, “Driving behavior classification Driver sensing based on vehicle steering wheel operations,” in *Proc. SICE Annual Conference*, 2008, pp. 2714–2718.
- [14] H. C. Lee, A. H. Lee, D. Cameron, and C. Li-tsang, “Using a driving simulator to identify older drivers at inflated risk of motor vehicle crashes,” *J. of Safety Research*, vol. 34, pp. 453–459, 2003.
- [15] M. Bedard, M. Parkkari, B. Weaver, J. Riendeau, and M. Dahlquist, “Assessment

- of Driving Performance Using a Simulator Protocol: Validity and Reproducibility,” *The American Journal of Occupational Therapy*, vol. 64, no. 2, pp. 336–340, 2010.
- [16] S. Lee, J. A. Lee, and H. Choi, “Driving Trail Making Test part B: a variant of the TMT-B,” *Journal of Physical Therapy Science*, vol. 28, no. 1, pp. 148–153, 2016.
- [17] L. J. Molnar, *et al.*, “Driving avoidance by older adults: Is it always self-regulation?,” *Accident Analysis and Prevention*, vol. 61, pp. 272–280, 2013.
- [18] A. M. Crizzle, A. M. Myers, and Q. J. Almeida, “Parkinsonism and Related Disorders Self-regulatory practices of drivers with Parkinson’s disease: Accuracy of patient reports,” *Parkinsonism and Related Disorders*, vol. 19, no. 2, pp. 176–180, 2013.
- [19] M. M. Porter, *et al.*, “Older Driver Estimates of Driving Exposure Compared to In-Vehicle Data in the Candrive II Study,” *Traffic Injury Prevention*, vol. 16, no. 1, pp. 24–27, 2015.
- [20] Z. Nasreddine, *et al.*, “The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment,” *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.
- [21] A. M. Colman, “Mini-Mental State Examination,” in *A Dictionary of Psychology*, Oxford University Press, 2015.
- [22] C. R. Bowie and P. D. Harvey, “Administration and interpretation of the Trail Making Test,” *Nature Protocols*, vol. 1, no. 5, pp. 2277–2281, 2006.
- [23] B. M. Dobbs and D. Schopflocher, “The Introduction of a New Screening Tool for the Identification of Cognitively Impaired Medically At-Risk Drivers: The SIMARD A Modification of the DemTect.,” *Journal of primary care & community health*, vol. 1, no. 2, pp. 119–127, 2010.
- [24] K. K. Ball and C. Owsley, “The useful field of view test: A new technique for evaluating age- related declines in visual function,” *Journal of American Optometric Association*, vol. 64.1, pp. 71–79, 1993.
- [25] A. J. Larner, *Cognitive Screening Instruments: A Practical Approach*. London; New York: Springer-Verlag London, 2013.
- [26] R. A. Stern, L. M. Abularach, D. R. Seichepine, M. L. Alosco, B. E. Gavett, and Y. Tripodis, “Office-Based Assessment of At-Risk Driving in Older Adults with and Without Cognitive Impairment,” *Journal of Geriatric Psychiatry and Neurology*, vol. 29, no. 6, pp. 352–360, 2016.
- [27] M. Bédard, B. Weaver, M. Man-son-hing, S. Classen, and M. Porter, “The SIMARD Screening Tool to Identify Unfit Drivers: Are We There Now?,” *Primary Care & Community Health*, vol. 2, pp. 133–135, 2011.
- [28] M. Bédard, *et al.*, “It is premature to test older drivers with the SIMARD-MD,” *Accident Analysis and Prevention*, vol. 61, pp. 317–321, 2013.
- [29] D. Koh and H. Kang, “Smartphone-Based Modeling and Detection of Aggressiveness Reactions in Senior Drivers,” in *Proc. IEEE Intell. Vehicles Symp. (IV)*, 2015, no. 4.
- [30] T. A. Dingus, *et al.*, “The 100-Car naturalistic driving study phase II – Results of the 100-Car field experiment,” *U.S. Department of Transportation*, no. April. pp. 1–351, 2006.
- [31] K. L. Campbell, “The SHRP 2 Naturalistic Driving Study: Addressing Driver

- Performance and Behavior in Traffic Safety,” *TR News*, no. 282, pp. 30–35, 2012.
- [32] Y. Barnard, F. Utesch, N. van Nes, R. Eenink, and M. Baumann, “The study design of UDRIVE: the naturalistic driving study across Europe for cars, trucks and scooters,” *European Transport Research Review*, vol. 8, no. 2, p. 14, 2016.
- [33] A. Williamson, *et al.*, “The Australian Naturalistic Driving Study: from beginnings to launch,” in *Proc. 2015 Australasian Road Safety Conference*, 2015.
- [34] D. W. Eby, N. M. Silverstein, L. J. Molnar, D. Leblanc, and G. Adler, “Driving behaviors in early stage dementia : A study using in-vehicle technology,” *Accident Anal. and Prevention and Prevention*, vol. 49, pp. 330–337, 2012.
- [35] G. M. Babulal, *et al.*, “Creating a driving profile for older adults using GPS devices and naturalistic driving methodology.,” *F1000Research*, vol. 5, no. 0, p. 2376, 2016.
- [36] N. Lin, C. Zong, M. Tomizuka, P. Song, Z. Zhang, and G. Li, “An overview on study of identification of driver behavior characteristics for automotive control,” *Mathematical Problems in Engineering*, vol. 2014, no. 2, 2014.
- [37] W. Wang, J. Xi, and H. Chen, “Modeling and recognizing driver behavior based on driving data: A survey,” *Mathematical Problems in Engineering*, vol. 2014, 2014.
- [38] S. Al-Sultan, A. H. Al-Bayatti, and H. Zedan, “Context-aware driver behavior detection system in intelligent transportation systems,” *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4264–4275, 2013.
- [39] G. A. M. Meiring and H. C. Myburgh, “A review of intelligent driving style analysis systems and related artificial intelligence algorithms,” *Sensors (Switzerland)*, vol. 15, no. 12, pp. 30653–30682, 2015.
- [40] M. Enev, A. Takakuwa, K. Koscher, and T. Kohno, “Automobile Driver Fingerprinting,” in *Proc. Proceedings on Privacy Enhancing Technologies*, 2016, vol. 2016, no. 1, pp. 34–51.
- [41] D. Hallac, A. Sharang, R. Stahlmann, A. Lamprecht, M. Huber, and M. Roehder, “Driver Identification Using Automobile Sensor Data from a Single Turn,” in *Proc. IEEE International Conference on Intelligent Transportation Systems*, November 2016, pp. 953–958.
- [42] X. Meng, K. K. Lee, and Y. Xu, “Human Driving Behavior Recognition Based on Hidden Markov Models,” in *Proc. IEEE International Conference on Robotics and Biomimetics*, 2006, pp. 274–279.
- [43] W. Dong, J. Li, R. Yao, C. Li, T. Yuan, and L. Wang, “Characterizing Driving Styles with Deep Learning,” *arXiv:1607.03611v2*, 2016.
- [44] M. Kuderer, S. Gulati, and W. Burgard, “Learning Driving Styles for Autonomous Vehicles from Demonstration,” in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, May 2015, pp. 2641–2646.
- [45] J. Grengs, X. Wang, and L. Kostyniuk, “Using GPS data to understand driving behavior,” *Journal of Urban Technology*, vol. 15, no. 2, pp. 33–53, 2008.
- [46] J. M. Scanlon, K. D. Kusano, and H. C. Gabler, “Analysis of Driver Evasive Maneuvering Prior to Intersection Crashes Using Event Data Recorders,” *Traffic Injury Prevention*, vol. 16, no. November, pp. 182–189, 2015.
- [47] Insurance Bureau of Canada, “How Auto Insurance Premiums are Calculated: Telematics,” 2017. [Online]. Available: <http://www.ibc.ca/on/auto/buying-auto->

- insurance/how-auto-insurance-premiums/telematics/. [Accessed: 18-Oct-2017].
- [48] H. Peter, I. Skog, M. Ohlsson, and J. Ohlsson, "Smartphone Instrumentation for Insurance Telematics," in *Proc. IEEE Instrum. Meas. Tech. Conf. (I2MTC)*, 2015, pp. 1613–1616.
- [49] F. Guo and Y. Fang, "Individual driver risk assessment using naturalistic driving data," *Accident Analysis and Prevention*, vol. 61, pp. 3–9, 2013.
- [50] A. B. Ellison, S. P. Greaves, and M. C. J. Bliemer, "Driver behaviour profiles for road safety analysis," *Accident Analysis and Prevention*, vol. 76, pp. 118–132, 2015.
- [51] W. K. M. Alhajyaseen, M. Asano, and H. Nakamura, "Estimation of left-turning vehicle maneuvers for the assessment of pedestrian safety at intersections," *IATSS Research*, vol. 36, no. 1, pp. 66–74, 2012.
- [52] D. A. Johnson and M. M. Trivedi, "Driving Style Recognition Using a Smartphone as a Sensor Platform," in *Proc. International IEEE Conference on Intelligent Transportation Systems*, 2011, pp. 1609–1615.
- [53] J.-H. Hong, B. Margines, and A. K. Dey, "A smartphone-based sensing platform to model aggressive driving behaviors," in *Proc. ACM conference on Human factors in computing systems*, 2014, pp. 4047–4056.
- [54] J. Howcroft, B. Wallace, R. Goubran, S. Marshall, M. M. Porter, and F. Knoefel, "Variation in acceleration driving patterns as a measure of older adult health status," *IEEE International Symposium on Medical Measurements and Applications*, vol. 1, pp. 298–302, 2017.
- [55] R. B. Wallace, "An Integrated Approach for Episodic Cognition Assessment," Carleton University, 2016.
- [56] B. Wallace, R. Goubran, and F. Knoefel, "Measurement of signal use and vehicle turns as indication of driver cognition," in *Proc. IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2014, vol. 2014, pp. 3747–3750.
- [57] A. M. Myers, A. Trang, and A. M. Crizzle, "Naturalistic study of winter driving practices by older men and women: examination of weather, road conditions, trip purposes, and comfort," *Canadian Journal on Aging*, vol. 30, no. 4, pp. 577–89, 2011.
- [58] M. Kilpeläinen and H. Summala, "Effects of weather and weather forecasts on driver behaviour," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 10, no. 4, pp. 288–299, 2007.
- [59] H. Tuokko, *et al.*, "Associations between age, gender, psychosocial and health characteristics in the Candrive II study cohort," *Accident Analysis and Prevention*, vol. 61, pp. 267–271, 2013.
- [60] J. Langford, *et al.*, "Findings from the Candrive / Ozcandrive study : Low mileage older drivers , crash risk and reduced fitness to drive," *Accident Analysis and Prevention*, vol. 61, pp. 304–310, 2013.
- [61] M. J. Rapoport, *et al.*, "The relationship between cognitive performance, perceptions of driving comfort and abilities, and self-reported driving restrictions among healthy older drivers," *Accident Analysis and Prevention*, vol. 61, pp. 288–295, 2013.
- [62] G. Anne and P. Smith, "Seasonal Variation in Older Adults' Driving Trip

- Distances,” University of Manitoba, 2014.
- [63] A. Smith, *et al.*, “Stability of physical assessment of older drivers over 1 year,” *Accident Anal. and Prevention*, vol. 61, pp. 261–266, 2013.
- [64] B. Wallace, *et al.*, “Automation of the Validation, Anonymization, and Augmentation of Big Data from a Multi-year Driving Study,” in *Proc. IEEE Int. Congr. Big Data*, 2015, vol. 1, pp. 608–614.
- [65] B. Wallace, R. Goubran, F. Knoefel, S. Marshall, and M. Porter, “Measuring variation in driving habits between drivers,” in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2014, pp. 3–8.
- [66] B. Wallace, R. Goubran, and F. Knoefel, “Measurement of Driving Routes and Correlation to Optimal Navigation Paths,” in *Proc. IEEE Int. Inst. Meas. Tech. Conf. (I2MTC)*, pp. 1465–1470.
- [67] B. Wallace and F. Knoefel, “Scoring Cognitive Change Through Sensing and Analysis of Changing Driving Ability,” in *Proc. CMBES*, 2013, vol. 36, pp. 7–10.
- [68] B. Wallace, R. A. Goubran, F. Knoefel, S. Marshall, and M. Porter, “Driver Unique Acceleration Behaviours and Stability over Two Years,” in *Proc. IEEE Int. Congr. Big Data*, 2016, pp. 230–235.
- [69] B. Wallace, *et al.*, “Big data analytics to identify deceleration characteristics of an older driver,” in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, 2015, vol. 1, pp. 89–94.
- [70] PERSENTECH, “OttoView-CD User Manual and Instruction Guide,” Winnipeg, Manitoba, 2010.
- [71] Government of Canada, “Historical Data - Climate - Environment and Climate Change Canada,” *Past weather and climate*, 2017. [Online]. Available: [http://climate.weather.gc.ca/historical\\_data/search\\_historic\\_data\\_e.html](http://climate.weather.gc.ca/historical_data/search_historic_data_e.html). [Accessed: 11-Sep-2016].
- [72] National Research Council Canada, “Sunrise and Sunset Calculator,” 2017. [Online]. Available: <https://www.nrc-cnrc.gc.ca/eng/services/sunrise/index.html>. [Accessed: 11-Sep-2017].
- [73] B. Wallace, *et al.*, “Measurement of Vehicle Acceleration in Studies of Older Drivers from GPS Position and OBDII Velocity Sensors,” in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, 2015, vol. 1, pp. 391–396.
- [74] E. W. Grafarend and F. W. Krumm, “Ellipsoid-of-revolution to cylinder: transverse aspect,” in *Map Projections: Cartographic Information Systems*, 1st ed., no. 1922, Berlin: Springer-Verlag, 2006.
- [75] U. Ramer, “An iterative procedure for the polygonal approximation of plane curves,” *Computer Graphics and Image Processing*, vol. 1, no. 3, pp. 244–256, 1972.
- [76] D. H. Douglas and T. K. Peucker, “Algorithms for the Reduction of the Number of Points Required To Represent a Digitized Line or Its Caricature,” *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 10, no. 2, pp. 112–122, 1973.
- [77] F. Fritsch and R. Carlson, “Monotone Piecewise Cubic Interpolation,” *Society for Industrial and Applied Mathematics*, vol. 17, no. 2, pp. 238–246, 1980.
- [78] V. Pratt, “Direct Least-Squares Fitting of Algebraic Surfaces,” *Computer Graphics*, vol. 21, no. 4, pp. 145–152, 1987.

- [79] K. O. McGraw and S. P. Wong, “Forming inferences about some intraclass correlation coefficients.,” *Psychological Methods*, vol. 1, no. 1, pp. 30–46, 1996.
- [80] N. C. Fung, *et al.*, “Driver Identification Using Vehicle Acceleration and Deceleration Events from Naturalistic Driving of Older Drivers,” in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, 2017, pp. 33–38.
- [81] J. C. Heckelman, “Probabilistic Borda rule voting,” *Social Choice and Welfare*, vol. 21, no. 3, pp. 455–468, 2003.
- [82] A. Zhang, B. Yang, and L. Huang, “Feature Extraction of EEG Signals Using Power Spectral Entropy,” in *Proc. International Conference on BioMedical Engineering and Informatics*, 2008, pp. 435–439.
- [83] MATLAB 2017b, “OOBPermutedVarDeltaError property: Variable importance for prediction error.” MATLAB 2017b, 2017.
- [84] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information Processing and Management*, vol. 45, no. 4, pp. 427–437, 2009.