

A Route-Choice Model for Predicting Pedestrian Behaviour and Violations

by

Christopher Lehmann Skelton

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Master of Applied Science

in

Civil Engineering

Waterloo, Ontario, Canada, 2024

© Christopher Lehmann Skelton 2024

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis contains sections that have been previously incorporated in conference proceedings and, at the time of writing, are included in journal article submissions, as follows:

“A Pathfinding Approach to Predicting Pedestrian Behaviour and Violations”

To be submitted for journal publication in 2024. This paper has been co-authored by myself, and my supervisors: Dr. Bruce Hellinga and Dr Chris Bachmann. I am the primary author of this article. This paper includes elements from all chapters of the thesis, although content from Chapter 1: Introduction, Chapter 2: Literature Review, Chapter 3: Model Development and Formulation, are substantially abridged in the paper compared to this thesis.

“Modelling Pedestrian Crossing and Violation Behaviour as a Pathfinding Process”

Included in the conference proceedings of the 59th Annual Canadian Transportation Research Forum. This conference paper was co-authored by myself, and my supervisors: Dr Bruce Hellinga and Dr Chris Bachmann. I am the primary author of this article. This paper includes content from Chapters 1 and 3 of this thesis, although this paper does not cover the complete model formulation.

Abstract

Pedestrians exhibit diverse behaviours, including crossing violations. Traditionally, development of behavioural models has been divided into route choice and crossing behaviour. Route choice models are stochastic and focused on crowd dynamics, while crossing behaviour models are probabilistic or deterministic and focused on local-level behaviours. Route choice and crossing behaviour are often addressed separately, but they are inherently related.

This research proposes a new pedestrian simulation model where pedestrians navigate through an intersection or mid-block environment, modelled as a grid. Each cell is assigned a cost that varies over time based on the presence of nearby vehicle traffic or changes to signal indications. Each pedestrian perceives the costs in the environment uniquely depending on their own personal preferences, like desired crossing gap or comfort committing a violation and seeks to minimize their total path cost. Pedestrians who are more comfortable committing violations perceive a lower cost for committing a violation. This approach integrates crossing behaviour with route choice and models the trade-offs of engaging in a particular behaviour.

The proposed model is calibrated using video data. The model was applied to three case-studies: a stop-controlled intersection, mid-block crossing, and two crosswalks along the minor approach of a signalized intersection. The model simulates the trade-offs between walking on different surfaces, as well as the trade-off between waiting for a gap in traffic to cross, versus diverting to the nearest designated crosswalk. In the third case study, the model successfully reproduced the proportion of pedestrians crossing against the signal for the north leg crosswalk but did not reproduce the proportion of violations for the south leg crosswalk, which is across a private access. Further investigation should be undertaken into the causes of this and other differences.

Acknowledgements

First and foremost, I would like to thank my supervisors, Dr Bruce Hellinga & Dr Chris Bachmann. This research would not have been possible without their support, assistance, and mentorship. I would like to thank Dr Liping Fu and Dr Adil Al-Mayah for reading this thesis and thank Dr Liping Fu for providing input on implementing the A* pathfinding algorithm.

I would like to especially thank Mark Hummel for piloting the drone used to capture the video footage in this research, allowing a much better angle than a handheld camera and saving me the awkwardness of standing at street corners videotaping passerby.

I would also like to thank my friends, fellow transportation graduate students and the ITE student chapter team for their support, advice, and for providing opportunities to direct my energies elsewhere when the research became tiring.

Lastly, I would like to thank my parents for their unconditional love and support throughout my years at UW, both as an undergraduate and graduate student.

Table of Contents

List of Figures.....	ix
List of Tables	xii
List of Variables Included in Model	xiii
1 Introduction.....	1
1.1 Background.....	1
1.2 Problem Statement	3
1.3 Research Objectives.....	5
1.4 Research Scope	5
1.5 Thesis Structure	6
2 Literature Review.....	7
2.1 Pedestrian Behaviour	7
2.1.1 The Pedestrian Behaviour Scale	7
2.1.2 The Theory of Planned Behaviour	10
2.1.3 Desire Paths	12
2.1.4 Legal Definitions of Pedestrian Violations.....	16
2.2 Models of Individual Pedestrians.....	18
2.2.1 Strategic Models	19

2.2.2	Tactical Models.....	20
2.2.3	Operational Models.....	21
2.3	Aggregate Pedestrian Models	23
2.3.1	Fundamental Flow Models	23
2.3.2	Aggregate Behaviour Models	27
2.4	Summary	29
3	Model Development and Formulation	31
3.1	Defining the Environment.....	32
3.2	Pathfinding.....	38
3.3	Cost Function	42
3.3.1	Base Costs.....	44
3.3.2	Violation Costs.....	46
3.3.3	Agent Costs.....	53
3.4	Pedestrian Awareness and Anticipation.....	58
3.5	Complete Model Formulation and Simulation Structure.....	60
4	Model Calibration	63
4.1	Base and Violation Cost Calibration.....	65
4.2	Simulated Sample Size Sensitivity	67
5	Case Studies	69

5.1	Case Study #1: Stop Controlled Intersection Crosswalk	69
5.2	Case Study #2: Midblock Crosswalk	70
5.3	Case Study #3	72
5.4	Discussion	76
5.5	Limitations	81
6	Conclusions	83
6.1	Future Work	84
	References	86
	Appendix A – Python Code	100
	Appendix B Fisher’s Test R Code	126

List of Figures

Figure 1.1 a) Planned Path & b) Actual Path of Pedestrian.....	4
Figure 2.1 Theory of Planned Behaviour (Ajzen, 1991)	11
Figure 2.2 Example of a Desire Path (Johnson, 2023)	13
Figure 2.3 Flow-Density Relationship for Pedestrians (Daamen & Transport, 2004)	24
Figure 2.4 Unidirectional vs. Bidirectional Pedestrian Flow Relationships (Zhang & Seyfried, 2013)	25
Figure 2.5 Unidirectional vs. Bidirectional Pedestrian Flow Relationships (Fruin, 1987).....	26
Figure 3.1 a) Aerial Image & b) Model Environment	37
Figure 3.2 a) Simple Graph & b) Multigraph (Wilson, 2009)	38
Figure 3.3 Dijkstra's Algorithm.....	39
Figure 3.4 A* Algorithm.....	40
Figure 3.5 Network Representation	41
Figure 3.6 Model Concept	43
Figure 3.7 a) Model Environment & b) Base Cost Heatmap ($\alpha_{base} = 0.5$)	46
Figure 3.8 Concept of Legal Violations as a Function of Distance from the Crosswalk	48
Figure 3.9 Legal Violation Cost for a) Midblock and b) Signalized Intersection	49
Figure 3.10 Crosswalk Legal Violation Cost (CCW, tLV) as a Function of Time	50
Figure 3.11 a) Model Environment & b) Agent Cost Heatmap ($\alpha_{LV} = 0.5$).....	51

Figure 3.12 Examples of Social Contract Considerations	53
Figure 3.13 Conflict Zone Concept	54
Figure 3.14 Probability of Not Yielding for a) Midblock and b) Signalized Intersection.....	56
Figure 3.15 a) Model Environment & b) Agent Cost Heatmap.....	58
Figure 3.16 Live Costs Over Time Compared to Expected Costs.....	59
Figure 3.17 Spatial Awareness of a Pedestrian.....	60
Figure 3.18 Total Cost Visualization	61
Figure 3.19 Simulation Workflow	62
Figure 4.1 Calibration Study Area	64
Figure 4.2 <i>Cgrass</i> Calibration Sub-Area	65
Figure 4.3 <i>CmaxLV</i> Calibration Sub-Area.....	66
Figure 4.4 Absolute Error Between Simulated & Observed Violation Proportions vs. a) Grass Base Cost & b) Max Legal Violation Cost.....	67
Figure 4.5 Proportion of Violating Pedestrians vs. Sample Size for a) Grass Base Cost & b) Max Legal Violation Cost.....	68
Figure 5.1 Case Study #2 Study Area.....	71
Figure 5.2 Case Study #3 Study Area.....	73
Figure 5.3 Illustration of Vehicle Conflict Zone Propagation	75
Figure 5.4 Calibration results for Pedestrians Crossing Against Do Not Walk (DNW) Signal ...	76

Figure 5.5 Sensitivity of Simulated Pedestrian Violations to Traffic Volume and Pedestrian
Delay 79

Figure 5.6 Spatial Differences Between a Common Observed Path and Simulated Path 80

List of Tables

Table 2.1 Items and Filters of the Pedestrian Behaviour Scale	7
Table 2.2 HCM LOS for Mid-Block Crossings (<i>Highway Capacity Manual 7th Edition, 2022</i>)	27
Table 2.3 Contributing Factors to Pedestrian Violations (Ghomi & Hussein, 2022).....	28
Table 3.1 Surface Types	32
Table 3.2 Classification of Social Contract Violations.....	53
Table 5.1 Case Study #1 Results.....	70
Table 5.2 Case Study #2 Results.....	72

List of Variables Included in Model

Variable	Description
$C_{p,n,t}$	Total cost of node n at time t perceived by pedestrian p .
$C_{p,n}^{base}$	Base cost of node n perceived by pedestrian p .
$C_{p,n,t}^{violation}$	Cost of a violation at node n at time t perceived by pedestrian p .
$C_{p,n,t}^{agent}$	Agent cost at node n and time t perceived by pedestrian p .
α_p^{base}	Sensitivity of the pedestrian p to higher cost of unpaved surfaces.
$C_n^{surface}$	Cost of the surface at node n .
$C_{pavement}$	Cost of a paved surface.
$C_{n,t}^{LV}$	Legal violation cost at node n , and time t .
d_{CW}	Distance to the nearest designated crosswalk, in metres.
d_{viol}	Maximum violation distance, in metres.
$C_{CW,t}^{LV}$	Cost of a legal violation at the crosswalk at time t .
C_{max}^{LV}	Maximum value of legal violation cost.
C_{MB}^{LV}	Cost of a legal violation at a mid-block location.
α_p^{LV}	Sensitivity of pedestrian p to legal violations.
$\beta^{Ped-Veh}$	Binary variable, controls conflict zone.
TTC	Time to collision.
$\gamma_{n,p}$	Probability of a vehicle <i>not</i> yielding to a pedestrian at node n as perceived by pedestrian p .
$P_{yield,CW,p}$	Probability of vehicles yielding at the crosswalk perceived by pedestrian p .
$P_{yield,MB,p}$	Probability of vehicles yielding midblock perceived by pedestrian p .
$C^{Ped-Veh}$	Cost of a risky interaction between pedestrians and vehicles.
$C_{n,p}^{Expected}$	Expected cost for node n and pedestrian p .
$C_{n,p,t}^{Live}$	Live cost at node n for pedestrian p at time t .
d_A	Spatial awareness of pedestrian.
t_A	Temporal anticipation of pedestrian.

1 Introduction

1.1 Background

Promoting walking as a mode of transportation is an important priority for governments at all levels. In 2021, the Canadian government approved the Active Transportation Fund. This Fund provides \$400 million over five years to active transportation projects. The importance of active transportation is also reflected at the municipal level. Municipalities across Canada have identified a need to increase the mode share of walking. The Region of Waterloo is targeting an increase to 9% from 7.2% (Region of Waterloo & IBI Group, 2019), and the City of Hamilton is seeking to triple walking mode share from 5% to 15% (City of Hamilton, 2018). American municipalities are also interested in increasing walking mode share. Boston is targeting an increase from 14% to 21% in walking mode share (Boston Transportation Department, 2017) and Minneapolis is targeting an increase from 16% to 25% (City of Minneapolis, 2020).

To support this shift in priorities, agencies have begun to develop dedicated strategies and manuals to support the design of safe and efficient pedestrian facilities. These documents are numerous, but some notable examples include the City of Toronto's Walking Strategy, Book 15: Pedestrian Crossing Treatments of the Ontario Traffic Manual (Ministry of Transportation, 2016a) and the British Columbia Active Transportation Guide (Ministry of Transportation and Infrastructure, 2022).

Pedestrians are heterogeneous system users (Tong & Bode, 2022), with diverse attributes that govern their behaviours. For example, attitudes and age are shown to affect behaviours, particularly regarding violations (Ghomi & Hussein, 2022). Violations can be broadly defined as pedestrian behaviours which are not aligned with how a pedestrian facility is intended to be used.

With diverse attributes, come diverse behaviours and violations. Pedestrians exhibit a diverse range of violation behaviours including spatial and temporal violations (e.g., jaywalking) along roadways (Chen et al., 2022), as well as the creation of desire paths (Richards & Ingold, 2022). An understanding of pedestrian behaviour is also key to designing safe and efficient pedestrian facilities. A pedestrian's behaviour and preferences affect how they want to use the facility. If a pedestrian does not use the facility in the way it was intended to be used, then the facility may not be used as efficiently as it could be, or it may be used in a manner which is unsafe to pedestrians, or other transportation system users. This difference between the intended design of a facility, and its use is sometimes referred to as pedestrian non-compliance or pedestrian violations. For example, poorly placed crosswalks, or crosswalks which are too far apart can lead to pedestrians crossing at unmarked locations, outside of the crosswalk. Pathway or trail designs which do not follow the desired routes of pedestrians are less likely to be used. Instead, pedestrians will select what they deem to be a more efficient route, like cutting across a grassy field to save the time and effort required to follow a more circuitous route along a paved path. These behaviours can also occur in a temporal context as well. Long cycle lengths at traffic signals are associated with poor pedestrian compliance. If the wait time for pedestrians is too long, pedestrians may make a trade-off between the risks of an unprotected crossing, and the delay caused by waiting for the next pedestrian phase (Yang & Sun, 2013).

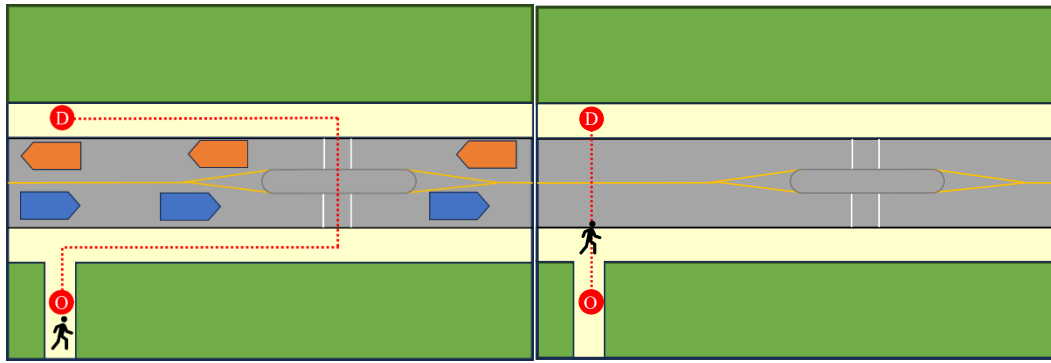
Besides design standards and manuals, one of the ways pedestrian behaviours can be examined is through the use of pedestrian behaviour models. Models are used across the transportation field to assess the performance of the system at different scales and levels of detail. Pedestrian behaviours are usually modelled in two main contexts: route-choice and road-crossing behaviour of pedestrians. Route choice models are usually stochastic and focused on crowd dynamics,

while crossing behaviour models are probabilistic or deterministic and are focused on local-level behaviours and attributes. Crossing behaviour models are most commonly aggregate models, although recently some simulation models have been developed (Rakotoarivelo et al., 2023). For example, the Highway Capacity Manual presents models to estimate pedestrian LOS at crossings and along road segments, but it does not attempt to simulate the behaviour of groups or populations of pedestrians (*Highway Capacity Manual*, 2022).

1.2 Problem Statement

Both route-choice and crossing behaviour are well studied, but often examined separately, despite the fact they are fundamentally related (Papadimitriou et al., 2009). To demonstrate the relationship between route choice and crossing behaviour, consider a scenario where a pedestrian is expecting to encounter busy traffic along a road they must cross to reach their destination. At the crossing location of the shortest path, there is no designated crossing, but there is a signalized crossing upstream along the road. Their expectation is that the road will be busy and thus they plan to divert from the shortest path and cross at the signalized crossing up the road. The route planned based on expected costs is shown in **Figure 1a**.

However, as the pedestrian proceeds along their planned path, they obtain more information about the surrounding environment. If the road is clear of traffic when they reach it and they are comfortable jaywalking across an empty road, they can simply cross at the location of the shortest path as shown in **Figure 1b**.



a)

b)

Figure 1.1 a) Planned Path & b) Actual Path of Pedestrian

This behaviour is a combination of both route choice and crossing behaviour. The pedestrian's route affected their crossing behaviour, and vice-versa, something which is not captured in existing models (Papadimitriou et al., 2009; Tong & Bode, 2022).

These behaviours can have implications for facility design, pedestrian safety, and roadway/intersection operations. A new model is required which can account for the emergent behaviour between pedestrians, the environment, and other agents, to model the behaviour of individual pedestrians more realistically than existing models.

There are other behaviours besides jaywalking which are not explicitly addressed in existing models, but which are regularly seen such as cutting across a grassy corner between two paved paths to save time and distance, and walking on the road to avoid snow covered sidewalks,

The above scenarios and others like them are common but are not adequately addressed in existing models. A need exists for a model which accounts for pedestrians' attitudes, social (subjective) norms, and rules; and integrates both route-choice, and crossing levels behaviours of pedestrians.

1.3 Research Objectives

The primary objective of this research is to improve the modelling of pedestrian behaviours compared to current models and capture the emergent route-choice and crossing behaviour of pedestrians as they interact with, and react to, the spatial and temporal characteristics of their environment and the behaviours of other agents.

The research objective will be achieved by proposing, calibrating and applying a new conceptual pedestrian model. Through a series of three test-cases at intersection and mid-block environments in the Region of Waterloo, Ontario, Canada, this research seeks to demonstrate both the capabilities and limitations of the new model. By developing this novel pedestrian model, this research aims to improve how pedestrians are modelled.

1.4 Research Scope

This research focuses on pedestrian modelling for a similar scope that might be seen in a typical assessment of intersection operations for vehicle traffic. The study area is an intersection between roadways or a mid-block crosswalk. The selected locations for field data collection are in the Region of Waterloo. The transferability of results between many different intersections, mid-block locations and municipalities are not examined. The spatial scope of the model is relatively small ($\leq 200 \times 200$ metres), representing a typical intersection study area for a vehicle traffic study. Larger pedestrian networks are not considered. Furthermore, the temporal period of analysis is relatively short (≤ 1 h) and attempts to capture peak pedestrian traffic conditions.

The research does not investigate long term changes to the environment. For example, the gradually trampling of vegetation by pedestrians, or the effect of snowfall on pedestrian

behaviour. Additional limitations that emerged from research will be discussed subsequently in this thesis.

1.5 Thesis Structure

This thesis is structured into seven chapters, including this chapter. Chapter 2 reviews the existing literature related to pedestrian behaviour and modelling approaches. The benefits and shortcomings of current approaches are noted and discussed.

Chapter 3 discusses the development and formulation of the proposed model. The rational and assumptions of the model are discussed along with the logic and flow of the simulation.

Chapter 4 presents the calibration of the model. Calibration was performed with data collected on Ring Road on the University of Waterloo Campus using a handheld video device and a video recording drone. The measure of effectiveness is the proportion of pedestrians committing a violation between a given O-D pair.

Chapter 5 explores the capabilities and weaknesses of the model through three case studies. The first case study is of the same site location used for calibration, but the simulation results are compared to a different observed dataset. The second case study is for a mid-block crossing near 256 Phillip Street in Waterloo, ON. The third case study is for the signalized intersection at University Avenue West & Phillip Street in Waterloo.

Chapter 6 summarizes the findings and conclusions of the research. This includes notable model results, the limitations of the current model and how those limitations may be addressed in future research.

2 Literature Review

This literature review provides a high-level overview of the behaviours people are known to exhibit in their role as pedestrians. It then discusses the existing methods and models for predicting pedestrian behaviours. Behaviour prediction methods can be separated into individual-level approaches, and aggregate-level mathematical models. Each of these categories are further subdivided and examined in subsequent sections.

2.1 Pedestrian Behaviour

2.1.1 The Pedestrian Behaviour Scale

The Pedestrian Behaviour Scale (PBS) is a self-report scale developed to measure injury risk behaviours across all ages of pedestrians (Granié et al., 2013). The PBS was based on six other existing questionnaires: the Scale of Pedestrian Behaviour (Moyano Díaz, 1997), the Driver Behaviour Questionnaire (Reason et al., 1990), Lawton’s aggressive driver behaviour scale (Lawton et al., 1997), the Positive Driver Behaviour Scale (Özkan & Lajunen, 2005), the Adolescent Road User Behaviour Questionnaire (Elliott & Baughan, 2004), and Granié’s road user behaviour questionnaire (Granié, 2008)

Initially the PBS consisted of 40 items (behaviours) with seven filters, however the PBS was further filtered to include only 23 items and three filters, categorized as follows in Table 2.1:

Table 2.1 Items and Filters of the Pedestrian Behaviour Scale (Granié et al., 2013)

Category	Items
Violation	<ul style="list-style-type: none">- I cross diagonally to save time- I cross outside the pedestrian crossing even if there is one less than 50 m away

	<ul style="list-style-type: none"> - I cross the street even though the pedestrian light is red - I cross even though the light is still green for vehicles
Error	<ul style="list-style-type: none"> - I cross the street between parked cars - I start to cross on a pedestrian crossing and I finish crossing diagonally to save time - I cross between vehicles stopped on the roadway in traffic jams - I walk on the roadway to be next to my friends on the sidewalk or to overtake someone who is walking slower than I am
Lapse	<ul style="list-style-type: none"> - I forget to look before crossing because I am thinking about something else - I forget to look before crossing because I want to join someone on the sidewalk on the other side - I cross without looking because I am talking with someone - I realize that I have crossed several streets and intersections without paying attention to traffic
Aggressive	<ul style="list-style-type: none"> - I get angry with another user and insult him - I get angry with another user (pedestrian, driver, cyclist, etc.) and I yell at him - I get angry with another user (pedestrian, driver, cyclist, etc.) and I make a hand gesture - I get angry with a driver and hit his vehicle

Positive	<ul style="list-style-type: none"> - I let a car go by, even if I have the right-of-way, if there is no other vehicle behind it - When I am accompanied by other pedestrians, I walk in single file on narrow sidewalks so as not to bother the pedestrians I meet - I stop to let the pedestrians I meet by - I walk on the right-hand side of the sidewalk so as not to bother the pedestrians I meet
Filter	<ul style="list-style-type: none"> - I walk for the pleasure of it - I take public transportation (buses, metro, tramway, etc.) - I walk because I have no other choice

The definitions of each category have their roots in previous literature reviewed to create the PBS. Violations are intentional deviations from formal or informal social practices, while errors and lapses are unintentional deviations (Reason et al., 1990). Errors occur when a pedestrian's action is not appropriate for the circumstance, while lapses occur when the pedestrian's action is not aligned with their normal intent, typically due to inattention (Rasmussen, 1980). Positive and aggressive behaviours are descended from similar items examined in the Driver Behaviour Questionnaires (Lawton et al., 1997; Özkan & Lajunen, 2005).

Pedestrians were asked to scale each item in the PBS from 1 ("never") to 6 ("very often") based on how often the pedestrians engage in a behaviour. Validation of the PBS involved a principal component analysis which identified four factors as explaining 55% of the variance: transgressions (errors and violations), lapses, aggressive behaviour and positive behaviour

(Granić et al., 2013). Cronbach's alphas were also calculated and indicated that the PBS has good internal reliability.

Since its original introduction and validation by Granić et al. (2013), PBS has been validated many times across the world (Vandroux et al., 2022). In Europe and the Eastern Mediterranean, versions of the PBS have been validated in Turkey, Greece, Hungary, Estonia, Russia (Solmazer et al., 2020), Serbia (Antić et al., 2016), as well as Pakistan, and Iran (Nordfjærn & Zavareh, 2016). In Asia and the Western Pacific, the PBS has been validated in China (Qu et al., 2016; Xu et al., 2018), Vietnam (Dinh et al., 2020), and Australia (O'Hern et al., 2020). In the Americas the original PBS was validated in the United States and alternative questionnaires have also been developed in the United States and other countries (Deb et al., 2017)

These applications have shown several conclusions relevant to this thesis, including that pedestrian risk-taking behaviour can depend on social and cultural values (Nordfjærn & Zavareh, 2016), and pedestrian behaviour can depend on age, gender, and walking attitudes (Antić et al., 2016). For example, those who walk by necessity make more violations, but those who walk for pleasure make more lapses.

2.1.2 The Theory of Planned Behaviour

The theory of planned behaviour (TPB) is a theory that suggests the behaviour of individuals are influenced by their intentions (Ajzen, 1991). These intentions are themselves determined by three factors: attitudes, subjective norms, and perceived behavioural controls. Attitudes refer to the individual's own appraisal of whether performing a behaviour is favourable or unfavourable, subjective norm refers to the individual's perception social pressure to perform (or not perform) an action, and perceived behavioural control refers to the individual's perceived ability for

themselves to perform a behaviour. The combination of these three factors creates an intention, which may then result in a behaviour, as shown in **Figure 2.1**.

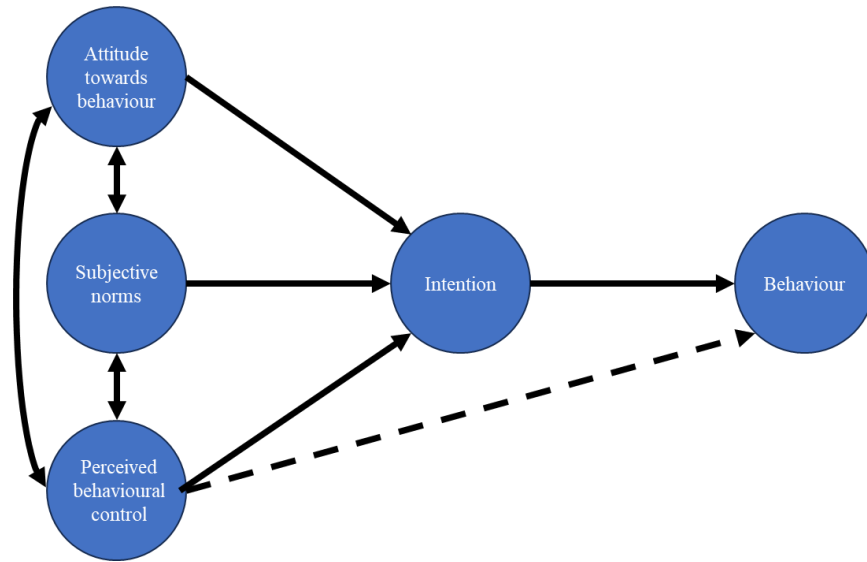


Figure 2.1 Theory of Planned Behaviour (Ajzen, 1991)

Since its original introduction, the TPB has been used for diverse applications in transportation.

In China, the TPB was used to examine individual's intentions to switch modes from private vehicles to public transit (Chen & Chao, 2011). In Sweden the TPB was used to predict speeding and dangerous overtaking violations for motor vehicles(Forward, 2009; Warner & Åberg, 2006). In Denmark, the TPB has been used to estimate departure time choices of travellers (Thorhauge et al., 2016)

Notably for this thesis, the TPB has also been used for research into pedestrian behaviour. The effect of the TPB on adult pedestrian behaviour in different crossing and pedestrian violation contexts has been examined (Barton et al., 2016; X. Zhao et al., 2022; Zhou et al., 2016), and has also been used to describe the behaviour of adolescent pedestrians (Evans & Norman, 2003; Meir et al., 2023; Xiao et al., 2021). Across reviewed studies, the pedestrian's, demographics

attitudes, subjective norm, and perceived behaviour control, were all highlighted as significant predictors of pedestrian behaviour.

2.1.3 Desire Paths

Desire paths, sometimes referred to as social trails, are unplanned paths which are created by the repeated movements and behaviours of people, vehicles or animals (Richards & Ingold, 2022). A common example of a desire path is an area of grass which has been trampled repeatedly by pedestrians (see **Figure 2.2**). Over time, the grass is worn down, leaving only dirt. The mechanics of how desire paths emerge are intuitive. Desire paths are a result of a discrepancy between land use/activities and path system (Ma et al., 2024). Pedestrians desire to take the shortest path. In a homogeneous environment, such as an open empty field, a pedestrians desired path will be a straight line from origin to destination (Helbing, Schweitzer, et al., 1997). However, the environment a pedestrian must traverse is almost always heterogeneous, with different types of surfaces which make be more or less attractive to a pedestrian (Ma et al., 2024).



Figure 2.2 Example of a Desire Path (Johnson, 2023)

For many surfaces repeated movements of people lead to long-term changes in the environment, such as in the previous example, where the grassy surfaces were gradually eroded leaving a narrow dirt path in its place.

Helbing, Keltsch, et al. (1997) proposed a model for such environmental changes shown in

Equation 2.1:

$$\frac{dG(r,t)}{dt} = \frac{1}{T(r)} [G_0(r) - G(r,t)] + I(r) \left[1 - \frac{G(r,t)}{G_{max}(r)} \sum_{\alpha} \delta(r - r_{\alpha}(t)) \right] \quad (2.1)$$

Where $G(r,t)$ represents the walking comfort of a surface at place r and time t . A surface has some base and maximum comfort, G_0 and G_{max} , respectively, as well as a surface durability, T . α denotes the simulated pedestrian (Helbing, Keltsch, et al., 1997).

Helbing, Keltsch, et al. (1997) further defined a simple model for pedestrian movement considering environmental changes. The attractiveness V of a trail segment at place r from the

perspective of pedestrian α is given by **Equation 2.2** and walking direction e_α is given by **Equation 2.3**.

$$V_{tr}(\mathbf{r}_\alpha, \mathbf{t}) = \int d^2 \mathbf{r} e^{-\frac{|\mathbf{r}-\mathbf{r}_\alpha|}{\sigma(\mathbf{r}_\alpha)}} G(\mathbf{d}, \mathbf{t}) \quad (2.2)$$

$$\mathbf{e}_\alpha(\mathbf{r}_\alpha, \mathbf{t}) = \frac{d_\alpha - \mathbf{r}_\alpha + \nabla_{\mathbf{r}_\alpha} V_{tr}(\mathbf{r}_\alpha, \mathbf{t})}{|d_\alpha - \mathbf{r}_\alpha + \nabla_{\mathbf{r}_\alpha} V_{tr}(\mathbf{r}_\alpha, \mathbf{t})|} \quad (2.3)$$

Where d_α is the position of the pedestrian.

The application of the model considered the simple real-world case of a trail system on the University Campus of Brasilia and showed good agreement with empirical real-world observations. However, the model application only considered a grassy surface which is eroded to gravel and did not consider other surface types which pedestrians may commonly walk on (e.g., asphalt, concrete). Furthermore, the model did not consider the heterogeneous perceptions of pedestrians.

Kudinov et al. (2018) developed a model for desire paths based on the A* pathfinding algorithm, where pedestrians take the lowest weight (least cost) path. The environment is modelled as a hexagonal grid with nodes at each grid cell that are connected by edges. If an edge exists between two nodes, the pedestrian can traverse the edge and travel between the nodes. The weight W of traversing an edge e is based on a constant term, which is a property of the surface, and a time variable component which presents the “trampledness” of a surface (Kudinov et al., 2018). This relationship is given by **Equation 2.4**.

$$W(e) = W_{const}(e) - W_{var}(e) \quad (2.4)$$

Note: W_{var} gradually increases up to some maximum value W_{max} , as in (Helbing, Keltsch, et al., 1997).

Based on empirical findings from Kudinov et al. (2018), the weight of a paved surface is 1, and an untrampled lawn is 2.7, implying that the lawn is 2.7 times less comfortable to traverse than a paved surface. Well-trampled lawns have a weight of 1.1.

Kudinov et al. (2018) also introduces the notion of different types of pedestrians, defining two groups: decent and indecent. Indecent pedestrians take the shortest path regardless of the “trampledness” of the surface. In other words, indecent pedestrians always perceive $W_{var} = W_{max}$. Decent pedestrians perceive W_{var} accurately, so their behaviour is affected by how much the surface has been trampled.

Ma et al. (2024) also developed a model which considers the heterogeneity of the environment, and the emergence of desire paths. Their model considers two guiding principles of behaviour: global conception and local adaptation. The global conception refers to the pedestrian’s goal to “Go Straight” and minimize the detoured distance. The local adaptation refers to the pedestrian seeking the easiest (maximally affording) surface to walk on, which is based on the type of ground, and the number of previous footprints (Ma et al., 2024). Ma et al. defined seven landcover surfaces (buildings, side-roads, roads, barriers, lawns, parking areas, and sidewalks). The affordance of each surface, is based on previous studies in ecology and wildlife migration (Savary et al., 2022; Shirk et al., 2010).

The Ma et al. (2024) model of surface affordance is shown in **Equation 2.5**.

$$c_i = Integer \left[\left(\frac{r_i - 1}{R_{max} - 1} \right)^n \times C_{max} \right] \quad (2.5)$$

Where c_i is the affordance of the surface, r_i is the rank of the surface relative to other surfaces.

R_{max} and C_{max} are the rank and affordances of sidewalks (the highest ranked surface).

2.1.4 Legal Definitions of Pedestrian Violations

Pedestrian violations may have legal consequences (e.g., tickets, reprimands, charges). The legal implications of pedestrian violations vary from jurisdiction to jurisdiction.

In Ontario, where this research is conducted, the Highway Traffic Act is the principal piece of legislation regulating “highways”, which includes all public highways, streets, avenues and other ways which are used by the general public for the passage of vehicles (Highway Traffic Act, R.S.O. 1990, c. H.8, 1990).

The Highway Traffic Act has three relevant Sections regarding pedestrian violations:

Section 140: Pedestrian Crossovers

(4) “No pedestrian shall leave the curb or other place of safety at a pedestrian crossover and walk, run or move into the path of a vehicle that is so close that it is impracticable for the driver of the vehicle to [stop before entering the crosswalk and not proceed until the pedestrian is no longer in the roadway].”

Section 144: Traffic Signals

(22) “Where portions of a roadway are marked for pedestrian use, no pedestrian shall cross the roadway except within a portion so marked.”

(23) “Subject to subsections (24) and (27), a pedestrian approaching a traffic control signal showing a circular green indication or a straight-ahead green arrow indication and facing the indication may cross the roadway”

(24) *“No pedestrian approaching a traffic control signal and facing a flashing circular green indication or a solid or a flashing left turn arrow indication in conjunction with a circular green indication shall enter the roadway.”*

(25) *“No pedestrian approaching a traffic control signal and facing a red or amber indication shall enter the roadway.”*

(26) *“Where pedestrian control signals are installed and show a “walk” indication, every pedestrian facing the indication may cross the roadway in the direction of the indication despite subsections (24) and (25).”*

(27) *“No pedestrian approaching pedestrian control signals and facing a solid or flashing “don’t walk” indication shall enter the roadway.”*

(28) *“Every pedestrian who lawfully enters a roadway in order to cross may continue the crossing as quickly as reasonably possible despite a change in the indication he or she is facing and, for purposes of the crossing, has the right of way over vehicles”*

Section 179: Duties of pedestrian when walking along highway

(1) *“Where sidewalks are not provided on a highway, a pedestrian walking along the highway shall walk on the left side thereof facing oncoming traffic and, when walking along the roadway, shall walk as close to the left edge thereof as possible.”*

In summary, pedestrians are required to use and obey the pedestrian infrastructure (facilities, signals, signage, and pavement markings) provided for them. In the absence of pedestrian infrastructure, it is not clear what constitutes a pedestrian violation. Notably, this has implications for crossings made outside of marked pedestrian crossings, sometimes referred to as “jaywalking”. Jaywalking is not explicitly prohibited in the absence of a marked crosswalk, but

this raises the question: How far from a crosswalk does a pedestrian need to be before it can be said there is “no marked crosswalk” and jaywalking is considered legal?

Instead of codifying a specific distance in legislation, Ontario court cases involving jaywalking have relied on “rules-of-thumb” and past precedents. For example, the Toronto Police Service has been quoted as advising pedestrians that jaywalking is prohibited within 30 metres of a marked crossing and the cases *R. v. Tablate* and *R. v. Dorian* have both determined ~100 metres to be a reasonable distance at which jaywalking is permitted, although this does not represent the minimum possible distance (Todorovic & Pedersen, 2020).

Similar provisions to those in the *Highway Traffic Act* can be found in other province’s highway legislation including Quebec (Highway Safety Code, 1986), Nova Scotia (Public Highways Act, 1989), Saskatchewan (Traffic Safety Act, T-18.1, 2006).

2.2 Models of Individual Pedestrians

At an individual scale, pedestrian behavioural modelling can be categorized into three levels: the strategic level, the tactical level, and the operational level (S. P. Hoogendoorn & Bovy, 2004). At the strategic level, the pedestrian selects their departure time and activities (which activities will be performed, in which order and where); at the tactical level, the pedestrian selects a route to accomplish their activities; and at the operational level, the pedestrian makes small-scale choices relating to walking behaviour (e.g., crossing behaviour)

This review of pedestrian behavioural models will roughly follow this categorization, acknowledging that the available literature on pedestrian behaviour is diverse and not all models fall neatly into these categories.

Furthermore, it is necessary to briefly define some common terms related to pedestrian modelling. Microsimulation models simulate the behaviour of individual decision-making units (Nigel & Klaus, 2005). In microsimulation modelling, the emphasis is placed on interactions between the individuals or policies of a system and the individual. For example, what is the effect of adding ramp-metering to a congested highway on the motorists who travel that highway.

Although microsimulation modelling simulates the behaviour of individuals, it is not to be confused with agent-based modelling (ABM). In ABM the emphasis is placed on local interactions and the emergent properties that appear as a result of those interactions (Nigel & Klaus, 2005). Agents interact with each other and the environment and can alter their behaviour based on changes in their surroundings and the heterogeneous preferences of the agents themselves (Bastariento et al., 2023).

2.2.1 Strategic Models

As summarized by Timmermans et al. (1992), early approaches to modelling strategic level behaviour were focused on examining trip generation and trip chains for retail areas in city centers and shopping centres.

One common generic approach is an Origin-Destination (OD) matrix based on the Markov Chain model, which estimates the probability of a pedestrian visiting a destination, given some origin (Timmermans et al., 1992).

Other approaches have also been examined, including estimating visitors to different locations within a large exhibition event (Canca et al., 2013), modelling planned and unplanned store visits in city centers (J. Dijkstra et al., 2014), and developed a needs-based model where the pedestrian

activities derive their utility from needs which can change over time based on past events (Arentze & Timmermans, 2009).

2.2.2 Tactical Models

The four principles of pedestrian route choice are: information perception, information integration, response to information, and a decision-making mechanism (Tong & Bode, 2022).

When perceiving information, pedestrians will selectively choose to focus on certain information, ignoring information which is not useful for their route choice decisions (Stea, 1973). Pedestrian's attention are commonly said to be driven by either a bottom-up or top-down approach (Katsuki & Constantinidis, 2014). In a bottom-up or stimulus approach, pedestrian attention is driven by external forces (e.g., distracted by a billboard). On the other hand, in a top-down or goal-oriented approach, attention is driven by prior knowledge and tasks relating to the goal (e.g., looking for navigational signage).

Furthermore, pedestrians have diverse but limited abilities to perceive their environment. The ability to constantly perceive their environment is important for route choice as information may change over time (Yang et al., 2009). While vision is the primary information source for pedestrians, olfactory and auditory stimuli are also valuable information sources (Porter et al., 2005; Wightman & Kistler, 1990).

The pedestrian must then integrate their knowledge and construct a mental map of the environment which they can then navigate.

In pedestrian route choice models, a common approach is to represent the environment as a series of links and nodes (Bhowmick et al., 2019), or as a grid with each cell represented as a node (Kudinov et al., 2018). The route is then often determined using a shortest-path algorithm

such as Dijkstra's or A*. After perceiving and integrating information, the pedestrian must then respond to the information and make decisions about their route. The pedestrians response may be governed in part by social conventions (Tong & Bode, 2022), or by positive or negative feedback loops, like in the earlier examples of desire paths (Helbing, et al., 1997). Then, pedestrians may make decision based on the utility of the available options, a heuristic, or a combination of both

These basic principles of pedestrian route choice can be found in existing route choice models. Often, route-choice models will focus on a specific factor related to route choice, such as pedestrian route-choice with a focus on temporally changing ground conditions (Helbing, et al., 1997; Kudinov et al., 2018); route-choice based on physical effort (Al-Widyan et al., 2017; Greenberg, n.d.), or the impacts of spatiotemporal variations on route choice (Cooper & Cowlagi, 2018). The Cooper & Cowlagi (2018) implementation also included the possibility of waiting behaviour, where pedestrians can minimize risk by waiting for hazards to pass. However, it did not consider the effects of social attitudes or ground conditions on route-choice.

A review of existing route-choice and crossing behaviour models identified a need for an integrated route-choice and crossing behaviour approach based on flexibility, (Papadimitriou et al., 2009). This would allow for flexibility, which addresses the pedestrian's decision-making process along the entire trip in relation to the pedestrians' own attributes, surrounding traffic and route characteristics.

2.2.3 Operational Models

The subject of pedestrian operational models has been of interest to researchers for some time. Early studies attempted to compared pedestrian movements to fluids or gases with some success

(Henderson, 1971; S. Hoogendoorn & Bovy, 2000). In the modern day, where simulations have become popular tools for pedestrian modelling, the social force model has become popular.

The social force model of pedestrian dynamics was first proposed by Helbing & Molnar. The premise of the social force model is that the pedestrian has an internal motivation to act which is affected by the surrounding environment (e.g., obstacles, other pedestrians) as forces. These forces are not exerted by the environment onto the pedestrian, but rather reflect the pedestrian's internal motivations to be attracted to, or repelled from something (Helbing & Molnar, 1995).

The core equations of the social force model are given in **Equation 2.6**.

The change in a pedestrian α velocity \vec{w}_α is given by:

$$\frac{d\vec{w}_\alpha}{dt} = \vec{F}_\alpha(t) + \textit{fluctuations} \quad (2.6)$$

Where $\vec{F}_\alpha(t)$ is the sum of the forces acting on the pedestrian, and is comprised of a driving force from origin to destination \vec{F}_α^0 ; repulsive forces due to other pedestrians β , $\vec{F}_{\alpha\beta}$; repulsive forces from physical barriers B , $\vec{F}_{\alpha B}$; and attractive forces to other people or objects i , $\vec{F}_{\alpha i}$.

In general, pedestrians avoid other pedestrians who are strangers, but may be attracted by other pedestrians such as friends or street artists. The same is true for objects. Pedestrians avoid obstacles but may be attracted by displays or advertisements along their route. Since its introduction, the social force model has become a popular modelling method in industry, in part due to its implementation in PTV's Viswalk software. Because of its popularity, the social force model has been the subject of much research including investigations into modelling pedestrian flows under bottleneck conditions (Kretz et al., 2008), calibration (Kretz et al., 2017), and travel time integration (Kretz, 2012).

2.3 Aggregate Pedestrian Models

Aggregate models of pedestrian behaviour are models which describe the behaviour of groups or populations of pedestrians. This section focuses on two types of aggregate models: pedestrian flow at the macroscopic level, and aggregate behavioural model, which estimate the likelihood of a particular behaviour occurring among a group of pedestrians.

2.3.1 Fundamental Flow Models

The fundamental relationship for pedestrians is similar to that of vehicle traffic, where a relationship between speed (u), flow (q), and density (k) is established (Vanumu et al., 2017).

The fundamental relationship between these parameters is given by: $q = ku$ (Treiber & Kesting, 2013).

These characteristics are considered to be macroscopic, since they describe the behaviour of a pedestrian traffic stream (Daamen & Transport, 2004). However, the definitions are slightly different from those used to describe vehicle traffic, since pedestrians can move freely in two dimensions, as opposed to vehicles that move in one dimension along a road. Pedestrians have desired directions as well as a desired walking speeds, which means that pedestrian speeds must be computed as vectors (Treiber & Kesting, 2013). Pedestrians flows also include a direction. Additionally, unlike vehicle traffic, pedestrian density is computed in two dimensions instead of one (i.e., P/m^2).

As with vehicle traffic, fundamental diagrams are often shown to illustrate the relationships between these macroscopic characteristics. Three diagrams are commonly shown in the literature: flow – density, speed – density, and speed – flow. **Figure 2.3** shows the conceptual relationship between flow (q) and density (k), and highlights several key points, including

(Daamen & Transport, 2004): Free speed, u_0 , the speed at which flow and density are zero; capacity, q_c , the maximum flow; and jam density, k_j , the density at which speed and flow are zero.

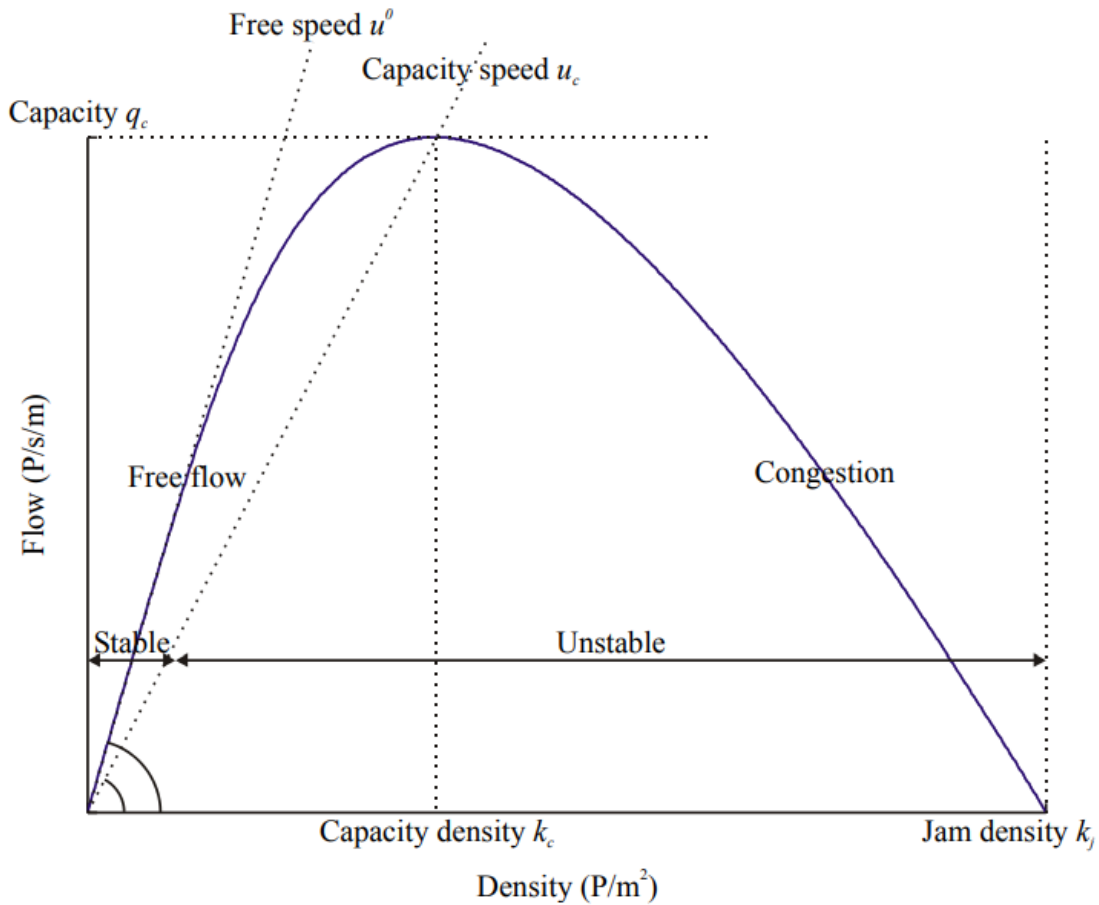


Figure 2.3 Flow-Density Relationship for Pedestrians (Daamen & Transport, 2004)

The shape of the fundamental diagrams depends on the type of pedestrian flow (uni-directional vs bi-directional) and the characteristics of the pedestrian facility (e.g., staircases vs. hallways).

Zhang & Seyfried (2013) conducted a series of experiments to compare uni-directional and bi-directional flows, as shown in **Figure 2.4**. Fruin illustrates the differences between these flows even more clearly in **Figure 2.5**. Several studies have also examined pedestrian flows in different

facilities and found that while straight corridors have the highest capacity, the specific flow does not depend on the width of the corridor (Hankin & Wright, 1958; Zhang et al., 2011; Zhang & Seyfried, 2013). In bottlenecks, where capacity is reduced significantly, flow depends on the width of the bottleneck (Seyfried et al., 2007; Tian et al., 2012). On stairs, capacity and speed are affected by the gradient of the stairs (Graat et al., 1999) .

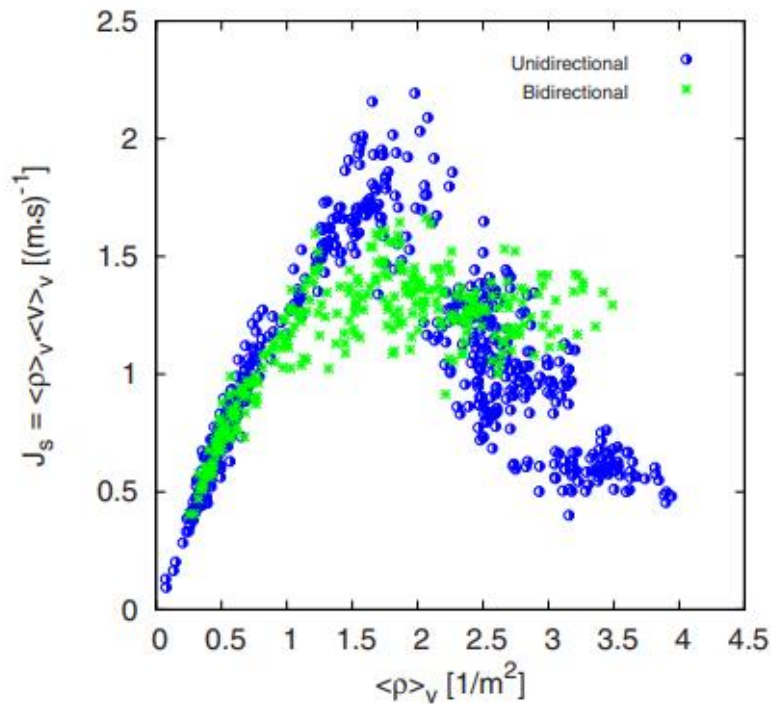


Figure 2.4 Unidirectional vs. Bidirectional Pedestrian Flow Relationships (Zhang & Seyfried, 2013)

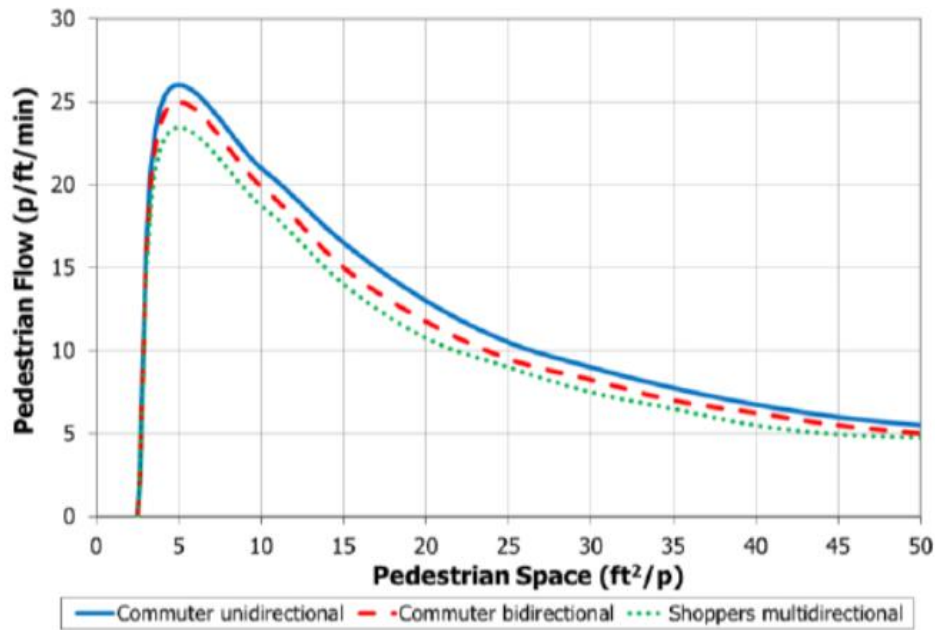


Figure 2.5 Unidirectional vs. Bidirectional Pedestrian Flow Relationships (Fruin, 1987)

When evaluating the relative performance of a system, the pedestrians Level of Service for a facility is an important consideration. Level of Service (LOS) is a qualitative measure to assess the performance of a facility for specific users, such as pedestrian (*Highway Capacity Manual*, 2022). The Highway Capacity Manual (HCM) 7th edition sets out the latest guidelines for pedestrian crossing LOS, summarized in **Table 2.2**.

Pedestrian crossing LOS determination for a mid-block or two-way stop controlled case is based on the proportion of pedestrians who are dissatisfied with the crossing experience, which is itself a function of pedestrian delay, annual average daily vehicle traffic (AADT), the odds that a pedestrian would be satisfied with the crossing, and the presence of pedestrian infrastructure such as median refuges, marked crosswalks, or flashing beacons (*Highway Capacity Manual*, 2022). The HCM 6th edition follows a very similar methodology to HCM 7th edition, but LOS is classified by delay, rather than pedestrian dissatisfaction (*Highway Capacity Manual*, 2016). The HCM 6th edition LOS criteria are also shown in **Table 2.2**.

Table 2.2 HCM LOS for Mid-Block Crossings

LOS	Proportion of Dissatisfied Pedestrians P_D (HCM 7)	Average Delay d_p (s/pedestrian) (HCM 6)	Comments (HCM 7)
A	$P_D < 0.05$	0-5	Nearly all pedestrians would be satisfied
B	$0.05 \leq P_D < 0.15$	5-10	At least 85% of pedestrians would be satisfied
C	$0.15 \leq P_D < 0.25$	10-20	Fewer than one-quarter of pedestrians would be dissatisfied
D	$0.25 \leq P_D < 0.33$	20-30	Fewer than one-third of pedestrians would be dissatisfied
E	$0.33 \leq P_D < 0.50$	30-45	Fewer than one-half of pedestrians would be dissatisfied
F	$P_D \geq 0.50$	>45	The majority of pedestrians would be dissatisfied
<p>Note: P_D = proportion of pedestrians giving a “dissatisfied” rating or worse.</p>			

2.3.2 Aggregate Behaviour Models

Aggregate models of pedestrian behaviour attempt to estimate the mathematical relationship between the number of pedestrians exhibiting a behaviour/likelihood of a behaviour occurring, and one or more explanatory variables. In this review, the focus is given to those studies which identify factors influencing pedestrian compliance and violation behaviours. The common

violations examined include spatial violations, such as jaywalking, or temporal violations, signal crossing against a “Do Not Walk” signal indication.

Ghomi & Hussein (2022) reviewed over 200 studies from 2000-2021 related to pedestrian violations. Using a Latent Dirichlet Allocation text mining method, they were able to summarize the impacts of different contributing factors on pedestrian violations. The contributing factors related to the frequency of pedestrian violations are summarized in **Table 2.3**.

Table 2.3 Contributing Factors to Pedestrian Violations (Ghomi & Hussein, 2022)

Category	Factor	Effect on Violations Frequency
Traffic	Higher traffic volume	Increased
	Higher vehicle speed	Increased
	Presence of heavy vehicles	Increased
	Presence of on street parking	Increased
Location	Higher number of lanes	Increased
	Longer and wider crosswalk	Increased
	Presence of central refuge islands	Decreased
	Presence of countdown signals	Decreased
	Presence of traffic signals	Decreased
Pedestrian Attributes	Being young	Increased*

	Being male	Inconclusive
Environmental Conditions	Adverse weather conditions	Inconclusive
	Lack of illumination	Increased
Built-Environment	Land use that attracts pedestrian activities	Inconclusive
	Presence of schools and bus stops	Increased

*some studies showed otherwise

Ghomi & Hussein (2022) are overall quite comprehensive, however, there are other factors in the literature which deserve attention. Dhoke & Choudhary (2023) completed a similar review of 73 studies, determining that one of the most common reasons for pedestrian non-compliance in the mid-block environment is caused by time pressure on the pedestrian (e.g., running late to work or attempting to catch a bus)(Dhoke & Choudhary, 2023). Another study also identified mobile phone use as a contributor to cognitive distraction that causes pedestrians to be inattentive and non-compliant (Hatfield & Murphy, 2007).

2.4 Summary

The literature surrounding pedestrian behaviour and pedestrian modelling is extensive and provides this research with important context for the model development. First, the pedestrian attributes such as attitudes and age have a significant effect on behaviours, particularly regarding violations. These pedestrian attributes are heterogeneous and vary from pedestrian to pedestrian.

Existing tactical modelling literature does not consider the connections between route-choice and crossing behaviour, nor the impacts of societal expectations on the pedestrians.

Furthermore, pedestrians exhibit a diverse range of violation behaviours including spatial and temporal violations along roadways, as well as the creation of desire paths. While there are a number of studies which have examined these behaviours separately, there are no studies that have examined these behaviours over the course of a pedestrian's entire trip (Papadimitriou et al., 2009). A pedestrian's decisions are also affected by the surrounding environment, which can change over time, requiring the pedestrian to process the new information, and decide whether to update their decisions accordingly.

3 Model Development and Formulation

This section presents the formulation of the proposed model and the rationale behind the formulation. In developing the model, a similar approach is taken to Kudinov et al. (2018). Namely, pedestrian decisions are based on an empirically defined cost function, where some pedestrians perceive different costs than others. The environment is represented as a grid structure, where pedestrians can travel between cells of the grid. Travelling between cells or waiting in a cell incurs some cost. The least cost path is found using an existing pathfinding algorithm.

From the review of pedestrian behaviours and existing models, several concepts must be integrated into the model and evolved beyond the work done by Kudinov et al. (2018). Firstly, the pedestrians in the model must be heterogeneous, with stochastically generated attributes. Furthermore, differing pedestrians' attitudes towards certain surface types or behaviours must be considered. For example, some pedestrians will consider a certain behaviour to be a severe violation, while others may consider it to be a moderate or minor violation. Additionally, specific consideration is given to pedestrian behaviours at mid-block and intersection locations, which were not addressed by Kudinov et al. (2018) and only partially addressed by Smirnov et al. (2020). Unlike previous work by Kudinov et al. (2018) or Helbing, Keltsch et al. (1997), the long-term changes in surface quality (e.g., trampledness), are not considered. The current focus of this model is for short-term temporal scopes ($\leq 1h$), where these properties are unlikely to change.

The development and rationale of the model are discussed in this section using the intersection of Ring Road and the access to the Science Teaching Complex within the University of Waterloo campus in Waterloo ON, as a real-world example to help visualize each aspect of the model.


This intersection is an all-way, three leg, stop controlled intersection. The model is implemented as a simulation in the Python programming language. Pedestrians are agents in the simulation, each with heterogeneous attributes, as discussed in subsequent sections.




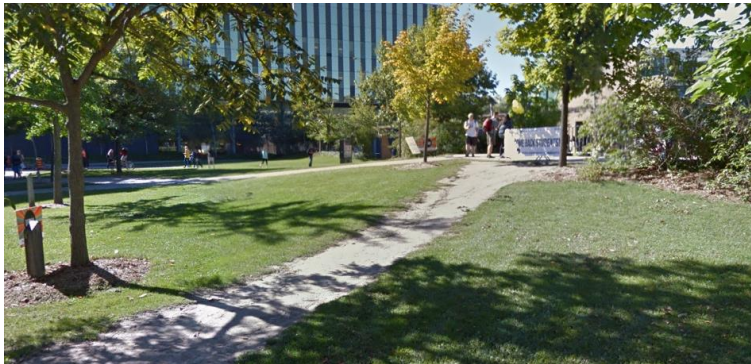
3.1 Defining the Environment



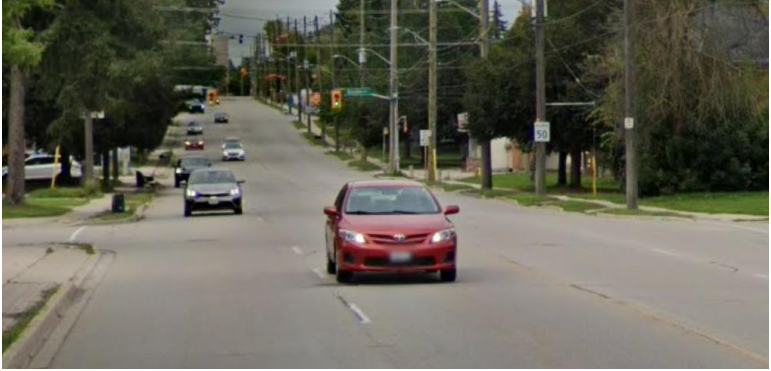
The model environment is represented as a grid of 1m x 1m cells, where each cell is assigned cost. The cost of a cell depends on attributes of the real-world surface it represents. So, each cell is assigned to be a specific surface type based on its physical surface (e.g., pavement, light vegetation, dense vegetation, gravel, etc.) and use restrictions where applicable. For example, paved surfaces are subcategorized based on the intended road user (bicycle, pedestrian, motor vehicle, shared, etc.). No distinction between weather conditions of the same surface (e.g., wet vs dry pavement) is made in this work.




Table 3.1 summarizes the defined surface types and provides photographic examples taken from around the Region of Waterloo using Google Streetview Imagery.



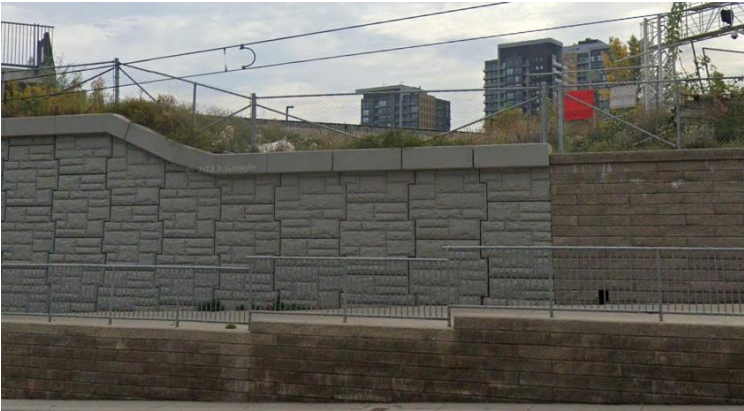
Table 3.1 Surface Types


Code	Description	Photographic Examples (Google, n.d.)
CW	Crosswalk	

<p>GV- MU</p>	<p>Gravel, Multi-Use Trail</p>	
<p>GV- MV</p>	<p>Gravel, Motor Vehicle</p>	
<p>GV- NT</p>	<p>Gravel area, not intended for traffic</p>	
<p>GV-PE</p>	<p>Gravel, Pedestrian</p>	

<p>PA-BI</p>	<p>Pavement, Bicycle</p>	
<p>PA-MU</p>	<p>Pavement, Multi-Use</p>	
<p>PA-MV</p>	<p>Pavement, Motor Vehicle</p>	

<p>PA-NT</p>	<p>Paved area, not intended for traffic</p>	 <p>Note: the median is the PA-NT area</p>
<p>PA-PE</p>	<p>Pavement, Pedestrian</p>	
<p>VE-LO</p>	<p>Light vegetation</p>	

<p>VE- ME</p>	<p>Medium vegetation</p>	
<p>VE-HI</p>	<p>Dense vegetation</p>	
<p>OB</p>	<p>Obstacle</p>	
<p>OT-PR</p>	<p>Other Prohibited Area</p>	<p>Miscellaneous category for areas prohibited to pedestrians which are not captured by other surface types/</p>

WT	Water	
----	-------	--

Using this classification system, each real-world surface can be assigned a surface type based on information gathered during a site visit or through satellite imagery. **Figure 3.1a** shows the real-world satellite imagery of the Ring Road at STC example location, compared to the assigned surface types in the model environment (**Figure 3.1b**). The different colours in **Figure 3.1b** denote the different surfaces (e.g., all orange surfaces are CW (Crosswalks)). Unless otherwise noted, all aerial images in this thesis were obtained using a drone during the data collection process, which is discussed later in this thesis.

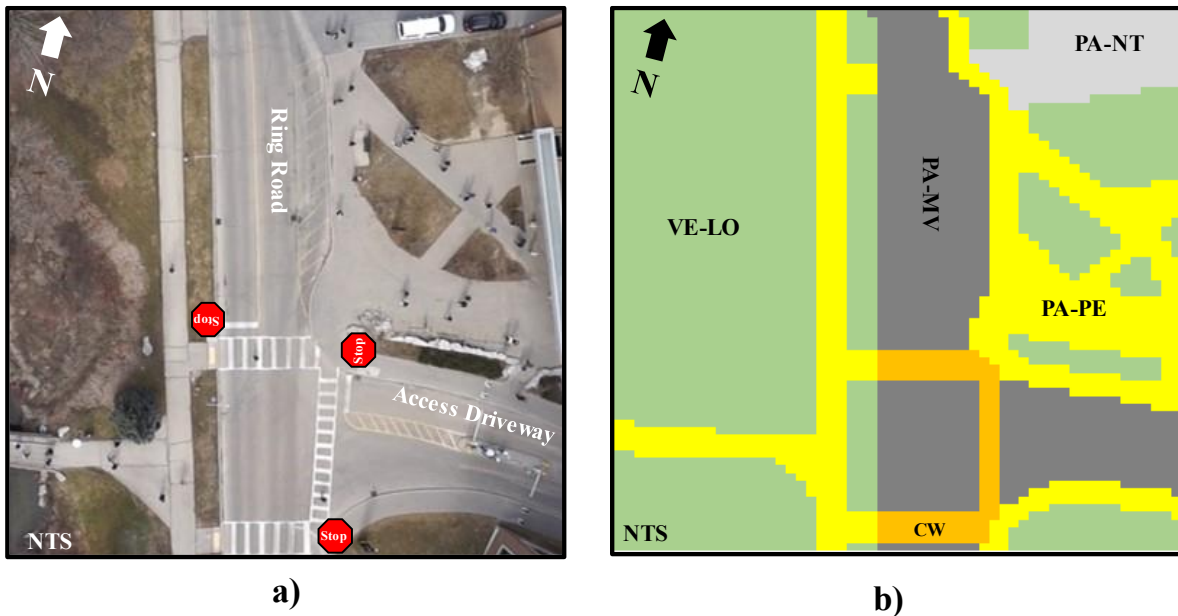


Figure 3.1 a) Aerial Image & b) Model Environment

3.2 Pathfinding

Pathfinding algorithms are a form of graph search algorithm which finds a route between a given origin and destination through a network. This network is represented by a graph which is made up of *vertices* (also referred to as *nodes*), which can be connected to each other by *edges* (also referred to as *links*) (Wilson, 2009). Edges are assigned a numerical weight which in the simplest cases represents the distance between two vertices, but can also represent other “costs” incurred for travelling between vertices such as toll fees, travel time, physical effort, etc. *Simple graphs* have a maximum of one edge connecting a pair of vertices and no vertices are connected to themselves, as shown in **Figure 3.2a**. Graphs with more than one edge connecting two vertices or vertices connected to themselves are referred to as *multigraphs*, as shown in **Figure 3.2b**. In this example, we can imagine a road network with multiple highway connections between S and T or Q and S, and a parking lot at node P (Wilson, 2009). The edges of a graph may also be considered undirected (analogous to a two-way street) or directed (one-way street).

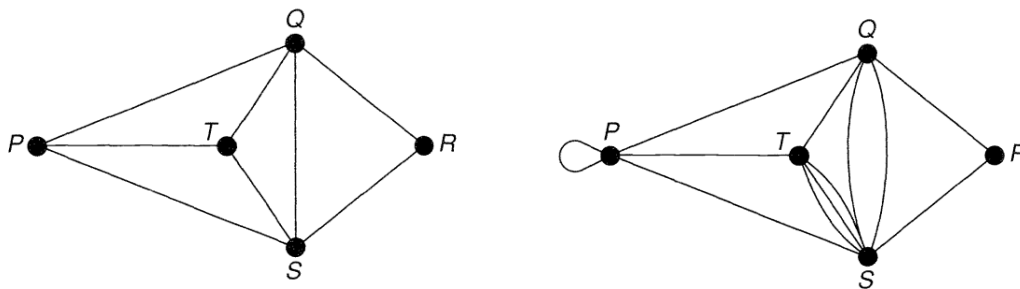


Figure 3.2 a) Simple Graph & b) Multigraph (Wilson, 2009)

The subject of pathfinding is an ongoing area of research, but many commonly used algorithms can trace their routes back to Dijkstra's Algorithm (E. W. Dijkstra, 1959). The algorithm starts at the origin and examines the next least-cost unexamined vertex available for travel. The algorithm tracks which vertices have been examined and updates the cost to travel to each vertex from the

origin if a lower cost path to that vertex is found. Dijkstra’s algorithm repeats this process and expands outwards from the origin until the destination is reached (E. W. Dijkstra, 1959). The algorithm is guaranteed to find an optimal path if one exists. The phrasing “an optimal path” is used since there may exist more than one path with the same lowest cost.

While Dijkstra’s algorithm is guaranteed to find an optimal path, it often expands the search unnecessarily to nodes in the opposite direction of the destination, adding computational time. Consider the example of a simple undirected graph shown in **Figure 3.3**. Relying on only the shorter edge weights, Dijkstra’s algorithm would start at O and first examine vertices A, B and C, before realizing it cannot reach D, thus restarting and examining E

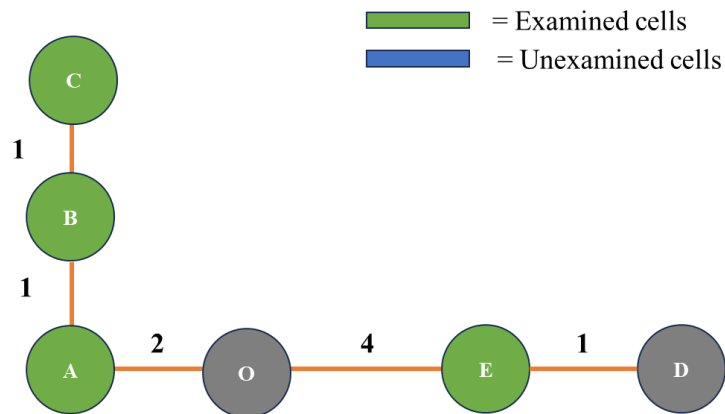


Figure 3.3 Dijkstra’s Algorithm

A modification of Dijkstra’s algorithm, known as A*, addresses this limitation by informing the search using a heuristic h (Hart et al., 1968). Most commonly h represents the distance from the location of a vertex to the destination vertex. When deciding which vertices to examine next, A*, considers the sum f of the edge cost g and the heuristic h : $f = g + h$. This guides the search towards the destination more efficiently by favouring vertices which are closer to the destination vertex. A* examines fewer vertices, but still achieves an optimal path, as shown in **Figure 3.4**.

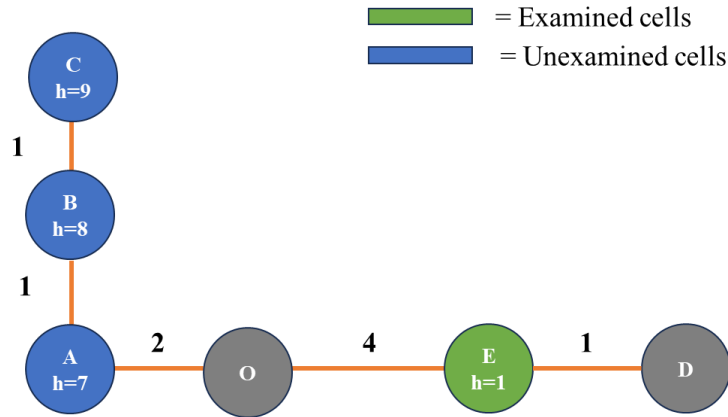


Figure 3.4 A* Algorithm

In this research, A* is used as the pathfinding algorithm because of its efficiency and because it has been used in related studies such as (Kudinov et al., 2018) and (Smirnov et al., 2020). The environment grid is represented as a graph of vertices where pedestrians can move between neighbouring vertices along edges.

This work uses a 1m x 1m square grid. At the centre of each cell is a node, which is connected to neighbouring nodes by links. Pedestrian travel along links between nodes. The cost of travelling along a link is associated with the destination node, and so the remainder of the thesis will discuss costs in terms of nodes.

Square grids are conceptually simple to understand, however diagonal movements on square grids pose a challenge since a diagonal movement is slightly longer than a straight movement and therefore incur more cost and take more time. In this simulation diagonal costs are multiplied by $\sqrt{2}$ but the time step for diagonal movements uniform for all directions. Therefore, diagonal movements can be susceptible to a “diagonal speed boost” where pedestrians arrive at the next node slightly earlier than they otherwise would if they were moving forwards, backwards, left or right. Travel time is not the measure of effectiveness for this work, so this speed boost in

unlikely to be a significant factor. This “speed boost” does mean, that, for long diagonal movements across temporally varying surfaces, the pedestrian may perceive a slightly different cost, because they have arrived earlier than they ought to have.

Diagonal movements in the XY plane are permitted, subject to the caveats listed above. Because the costs in the model can change over time, time exists as a third dimension of the graph where edges are directed forwards in time (i.e., the pedestrian cannot travel backwards in time). The time step (Δt) is the time it takes for a pedestrian to walk one metre, as defined by their walking speed. A 3D representation for the case of a single node to its neighbouring nodes at the next time step is shown in **Figure 3.5**.

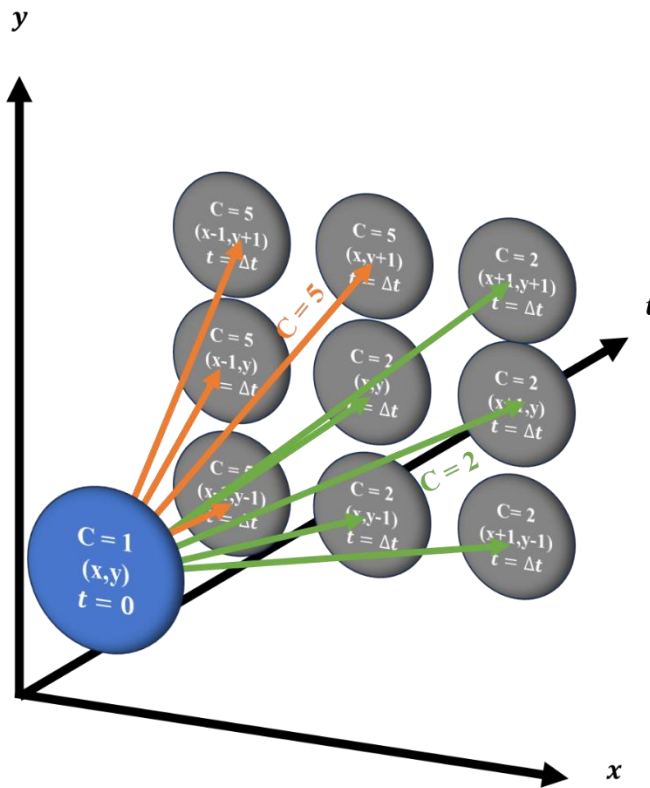


Figure 3.5 Network Representation

3.3 Cost Function

The unattractiveness of a route is expressed as a cost. Many route options are possible between a given origin and destination and the pedestrian selects what they perceive to be lowest cost (i.e., most attractive) route. Any behaviours the pedestrian engages in (jaywalking, corner cutting, etc.) are because they perceive those behaviours as being more cost-effective than not engaging in those behaviours. The pedestrian behaviour produced by the model is a result of the pedestrian's preferred route through the surrounding environment. The environment includes three main components the pedestrian must consider: 1) the attributes of the surfaces, which impact the pedestrians ease of travel; 2), applicable rules and social norms (usually prescribed by signage or pavement markings), which affect the degree to which a pedestrian will perceive an action to be a violation; and 3), other agents such as vehicles or other pedestrians. This concept is shown in **Figure 3.6**.

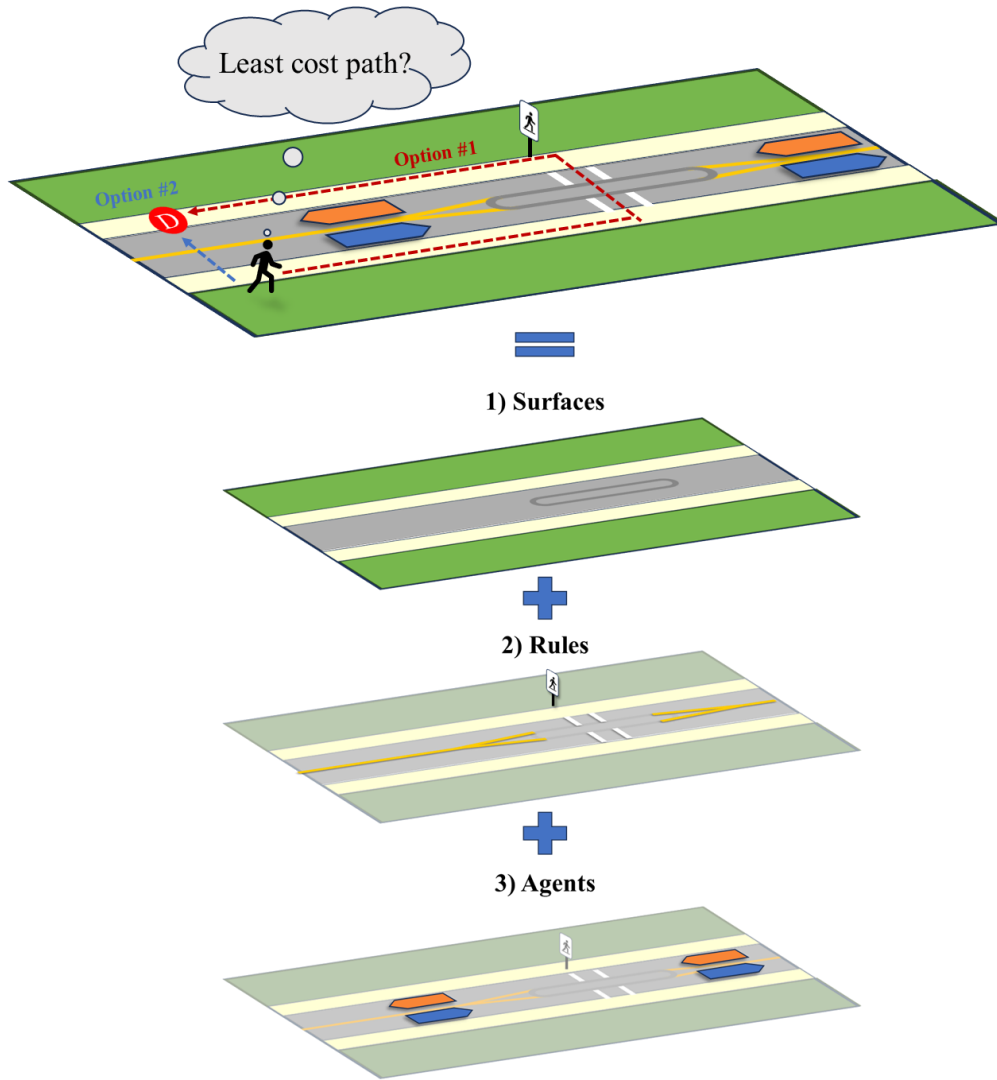


Figure 3.6 Model Concept

So, the total cost (C) of a node n at time t perceived by pedestrian p is expressed in **Equation 3.1**

below:

$$C_{p,n,t} = C_{p,n}^{base} + C_{p,n,t}^{violation} + C_{p,n,t}^{agent} \quad (3.1)$$

Where:

- $C_{p,n}^{base}$ is the base cost of node n perceived by pedestrian p
- $C_{p,n,t}^{violation}$ is the cost of a violation at node n at time t perceived by pedestrian p

- $C_{p,n,t}^{agent}$ is the agent cost at node n and time t perceived by pedestrian p

The individual terms of the cost function are discussed in more detail in subsequent sections.

3.3.1 Base Costs

Base costs reflect the physical attributes of each surface, as it pertains to a pedestrian's ability to traverse it. In essence, the base cost of a node reflects the relative difficulty of traversing that node. In this way, the base cost could be thought of as a proxy for the physical effort required to traverse a node, however this is not always an appropriate analogy.

Different pedestrians have different perceptions of the quality of a surface. A pedestrian's physical mobility has a significant impact on this perception. If a pedestrian requires a wheelchair, they may be unable to traverse a dense vegetation surface, thus perceiving a very high cost. This may be due to pre-existing attitudes. For example, a pedestrian who cares deeply about maintaining pristine grass, may choose not to risk damaging the grass by walking on it. Pedestrians may also have different perceptions, depending on how well equipped they are for certain environmental conditions. A hiker in a raincoat and waterproof boots is likely much more accepting of wet, muddy conditions than a businessperson in dress shoes and a suit.

The model defines a paved surface as the reference for the costs of all other surfaces. A paved surface (concrete or asphalt) is considered to be the ideal surface for pedestrians. For the reasons mentioned above, pedestrians may have different perceptions of the relative differences between a paved surface and unpaved surfaces. So, a sensitivity parameter α_p^{base} is defined as an attribute of pedestrian p which reflects the sensitivity of the pedestrian p to higher cost of unpaved surfaces. α_p^{base} is drawn from a triangular distribution ranging from 0 to 1 with mode 0.5. A

pedestrian with $\alpha_p^{base} = 0$ perceives all surfaces to be just as desirable (have the same cost) as pavement, while a pedestrian with $\alpha_p^{base} = 1$ is the most sensitive to unpaved surfaces.

The triangular distribution is used for now as a simple alternative to the normal distribution, since no information is known about the standard deviation of pedestrian surface sensitivity. Future work could investigate other distributions.

The base cost of node n perceived by pedestrian p is given by **Equation 3.2**:

$$C_{p,n}^{base} = \alpha_p^{base} (C_n^{surface} - C_{pavement}) + C_{pavement} \quad (3.2)$$

Where:

- $C_n^{surface}$ is the cost of the surface at node n
- $C_{pavement}$ is the cost of a paved surface

For most practical calibration purposes, $C_{pavement} = 1$. A value of one is considered as the lowest cost possible since the A* algorithm is intended for graphs with edge costs greater than zero (Hart et al., 1968).

The costs of the grid cells can be visualized as a heatmap, as shown in **Figure 3.7b** for $\alpha_p^{base} = 0.5$. Higher costs are shown in red, while lower costs are shown in green. **Figure 3.7a** shows the model environment for reference. Since both the sidewalk, crosswalk, road, and parking area are all paved, there is no difference between them in terms of base costs. Light vegetation (VE-LO), however, is not paved and has a higher cost.

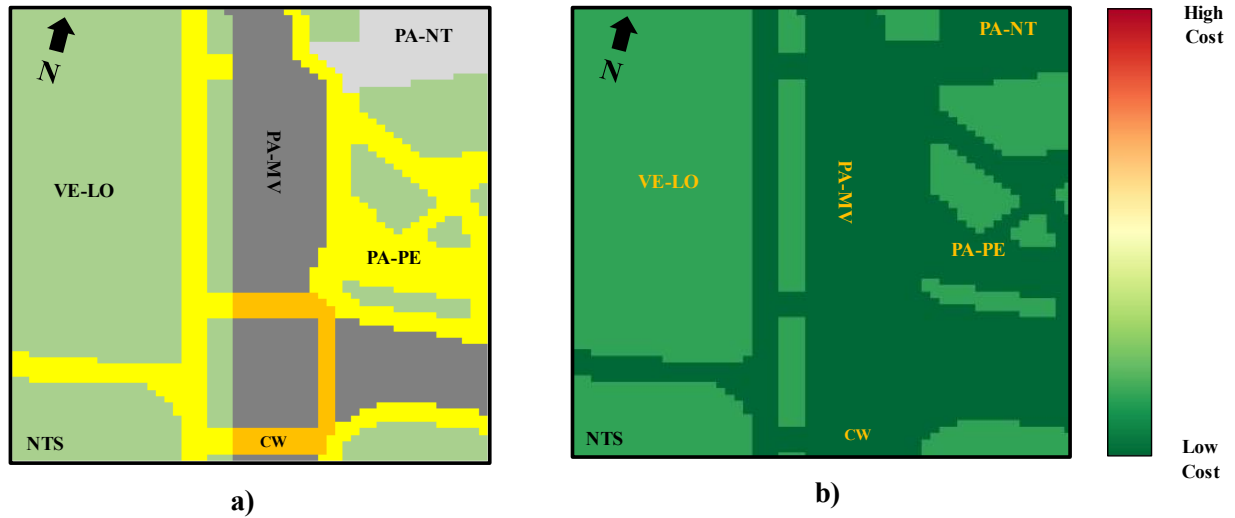


Figure 3.7 a) Model Environment & b) Base Cost Heatmap ($\alpha_p^{base} = 0.5$)

3.3.2 Violation Costs

Violation costs reflect a pedestrian’s perceived penalty for entering a node that they are prohibited from entering (i.e., engaging in a violation behaviour). As outlined by (Ajzen, 1991) and (Granié et al., 2013), a pedestrian’s perception of whether or not an action is a violation depends on personal attitudes and social norms/rules. Social norms or rules may be formalized in written documents such as legislation or by-laws, or they may be informal and reflect the social expectations of others in a society or social group. This thesis refers to violations of formalized rules as *legal violations* and violations of informal rules as *social contract violations*. In this research, legal violations are the focus as they are more relevant to the context of an intersection or mid-block environment.

Social contract violations are discussed but are not implemented in the simulation; examples of social contract violations are discussed in Section 3.3.2.2.

3.3.2.1 *Legal Violations*

In the context of this thesis, which is focused on the intersection and mid-block context, legal violations are violations of established legislation and by-laws related to traffic and public roadways. In Ontario, the Highway Traffic Act is the principal piece of legislation regulating traffic on public roadways. As established in Section 2.1.4, the Highway Traffic Act does not provide guidance on how far from a crossing area a pedestrian must be before crossing is no longer considered a violation (Highway Traffic Act, R.S.O. 1990, c. H.8, 1990). Similarly vague guidelines are found in other provinces. In the absence of guidance from legislation, court cases in Ontario have used 30 metres as a guideline (Todorovic & Pedersen, 2020).

Given the vagueness of the precise definition of a legal violation, it is hypothesized that the perceived cost for making a violation varies depending on a pedestrian's distance from a crosswalk. To illustrate this relationship, consider the real-world example of Phillip Street just north of University Avenue in Waterloo. Phillip Street is a two-lane road with the University of Waterloo Campus along its west side, and several student residences along the east side. Naturally, this section of Phillip Street is a popular crossing point for students, and a pedestrian refuge island has been built to accommodate this movement. However, this crossing does not align with the preferred route for many students, so jaywalking is common.

Instead, what is observed is that many students cross at the island, some cross near the island, and some cross far away from the island. In between, however, very few students cross, as shown in **Figure 3.8**. This suggests that near and far away from the island, most pedestrians are not perceiving a high cost for crossing, but in between the cost is much higher.

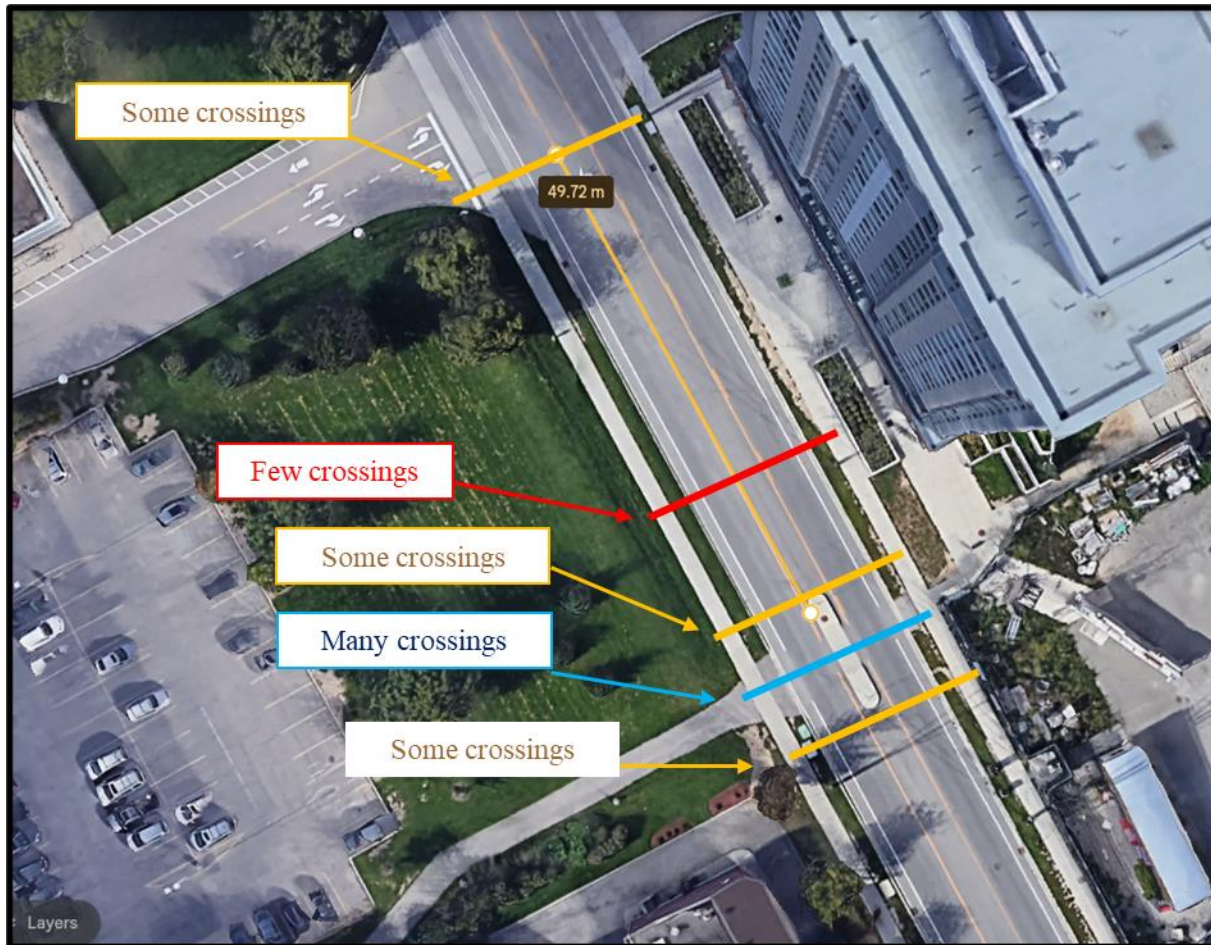


Figure 3.8 Concept of Legal Violations as a Function of Distance from the Crosswalk

The legal violation cost at node n , and time t , $C_{n,t}^{LV}$, is formulated as a function of distance to the nearest designated crosswalk, d_{CW} , with a maximum violation distance d_{viol} , which is equal to 30 metres in this research. This relationship is shown in **Figure 3.9a** for the case of a mid-block crossing and an additional example is shown in **Figures 3.9b** for the case of a signaled intersection.

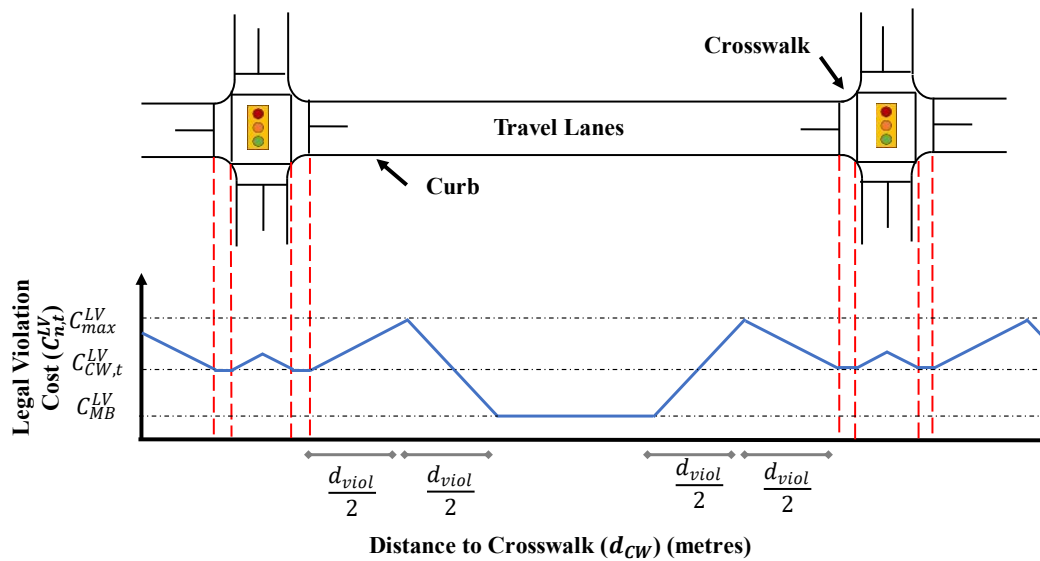
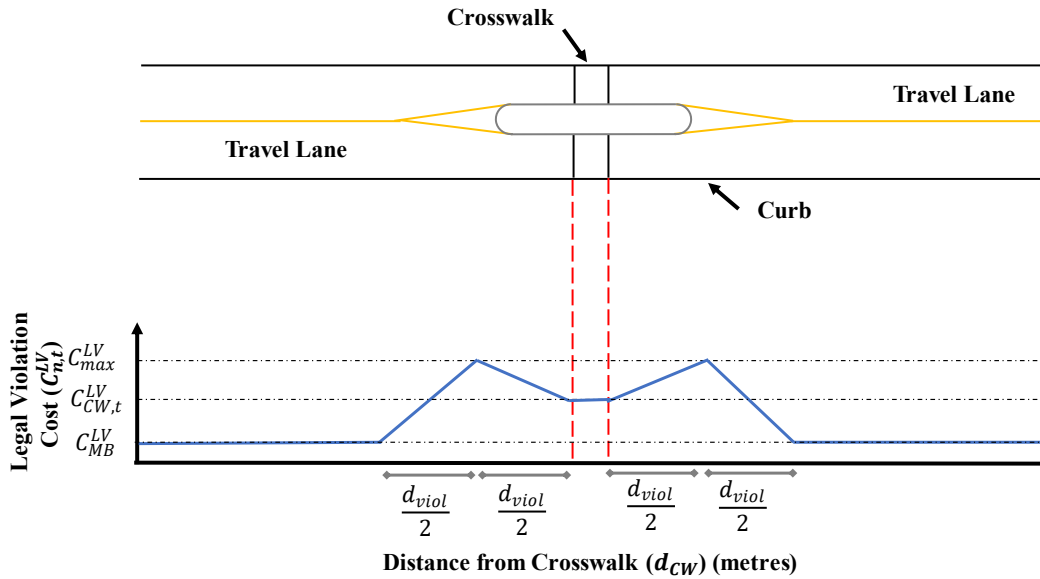


Figure 3.9 Legal Violation Cost for a) Midblock and b) Signalized Intersection

At 0 metres from the nearest crosswalk, a pedestrian perceives $C_{CW,t}^{LV}$, the cost of a legal violation at the crosswalk at time t . $C_{CW,t}^{LV} = 0$, except for when the pedestrian is crossing against a Do Not

Walk signal indication, thus committing a temporal violation. This conceptual relationship of $C_{CW,t}^{LV}$ in a crosswalk over time is shown in **Figure 3.10**.

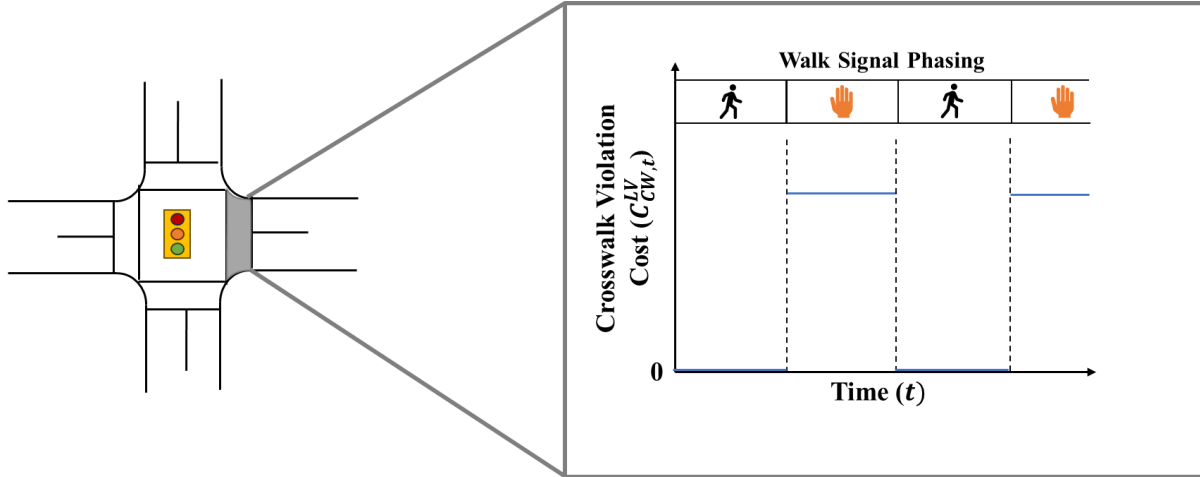


Figure 3.10 Crosswalk Legal Violation Cost ($C_{CW,t}^{LV}$) as a Function of Time

From 0 metres to $\frac{d_{viol}}{2}$, $C_{n,t}^{LV}$ increases linearly until it reaches its maximum value, C_{max}^{LV} , at $\frac{d_{viol}}{2}$.

From $\frac{d_{viol}}{2}$ to d_{viol} metres, the $C_{n,t}^{LV}$ decreases linearly, and beyond 30 metres, the pedestrian

perceives C_{MB}^{LV} , the cost of a legal violation at a mid-block location. C_{max}^{LV} can be calibrated

separately but cannot be lower than either of $C_{CW,t}^{LV}$ or C_{MB}^{LV} . This ensures that, based on violation

costs, the pedestrian does not experience a lower cost for avoiding the crosswalk than for

crossing at the crosswalk, so pedestrians are never discouraged from using the crosswalk. Just

like the base cost, a sensitivity parameter, α_p^{LV} , is also introduced for each pedestrian p and is

drawn from a triangular distribution ranging from 0 to 1 with mode 0.5. This represents the

pedestrians' attitudes towards legal violations. C_{MB}^{LV} and $C_{CW,t}^{LV}$ are assumed to be zero in this

initial work, which is consistent with the Ontario guidelines that crossing in the crosswalk or

beyond a certain distance from the crosswalk is not considered to be a violation. Future work

should study this assumption further.

The legal violation costs of the grid cells can be visualized as a heatmap, as shown in **Figure 3.11b** for $\alpha_p^{LV} = 0.5$. **Figure 3.11a** shows the model environment for reference. C^{LV} applies to road cells only. The function of $C_{n,t}^{LV}$ is expressed mathematically as a linear piecewise function, shown in **Equation 3.3**.

$$C_{n,t}^{LV} = \begin{cases} \frac{C_{max}^{LV} - C_{CW,t}^{LV}}{\frac{d_{viol}}{2}} d_{CW} & \text{if } 0 < d_{CW} \leq \frac{d_{viol}}{2} \\ \frac{C_{MB}^{LV} - C_{max}^{LV}}{\frac{d_{viol}}{2}} d_{CW} + 2C_{max}^{LV} & \text{if } \frac{d_{viol}}{2} < d_{CW} < d_{viol} \\ 0 & \text{otherwise} \end{cases} \quad (3.3)$$

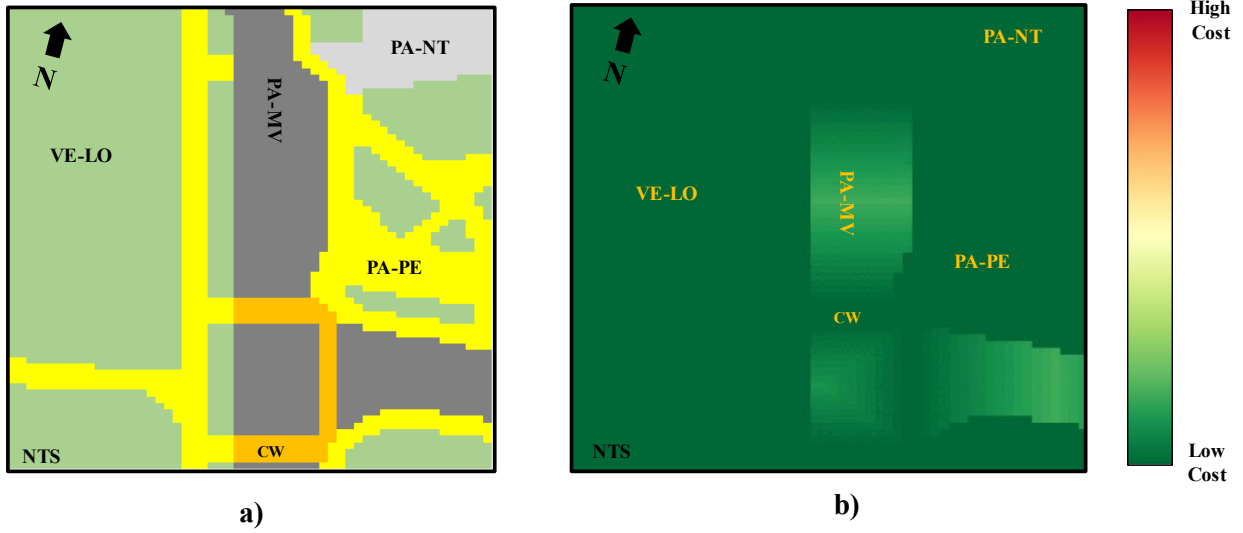


Figure 3.11 a) Model Environment & b) Agent Cost Heatmap ($\alpha_p^{LV} = 0.5$).

An individual pedestrian's perceived violation cost, $C_{p,n,t}^{violation}$ is shown in **Equation 3.4**:

$$C_{p,n,t}^{violation} = \alpha_p^{LV} C_{n,t}^{LV} \quad (3.4)$$

3.3.2.2 *Social Contract Violations*

The concept of the social contract has existed for millennia and describes the notion that there are certain informally agreed upon and understood rules which are collectively agreed upon in order to live in a morally just society (Friend, n.d.).

An easy way to illustrate social contract violations is to consider the distinction between cutting through a commercial parking lot to same time, versus cutting through a homeowner's garden. In principle, the action is the same (i.e., trespassing on private property), but in practice, pedestrians are likely to perceive the two actions very differently.

In this thesis, it is hypothesized that there are two principal factors which influence the severity of a violation of the social contract for pedestrians: access and durability. With access, whether the public commonly has access to the area is the key consideration. A commercial parking lot is openly accessible to patrons (usually without any need to prove that you are actually a patron of the business), while a homeowner's garden is only open upon express invite of the property owner.

With durability, the likelihood that an individual pedestrian will cause damage to the surface is the key consideration. If a surface is fragile, like a garden flower bed, walking over it could cause substantial damage and be considered an unacceptable behaviour. Walking on pavement, however, is unlikely to cause meaningful damage. Durability and access are not mutually exclusive. If a surface is both publicly accessible and durable, it is likely considered acceptable to walk on, and the pedestrian would perceive a low cost. If the surface is both fragile and private accessed, it is likely considered unacceptable to walk on, and the pedestrian would perceive a high cost. **Figure 3.12** and **Table 3.2** demonstrate several possible outcomes, given a simple binary classification for each factor.

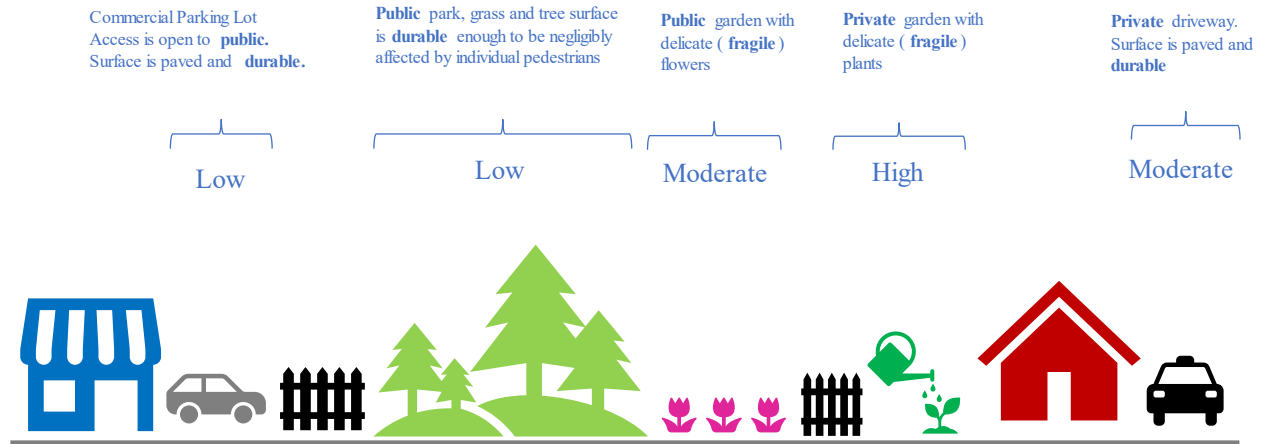


Figure 3.12 Examples of Social Contract Considerations

Table 3.2 Classification of Social Contract Violations

	Public Access	Private Access
Durable	Low	Moderate
Fragile	Moderate	High

The above scenarios which describe social contract violations are not the focus of this research and are not implemented in the current model.

3.3.3 Agent Costs

So far, both the base costs and violation costs are independent of the behaviour of other agents in the model and would not change over time during a simulation. There are very few practical applications where one would be interested in modelling a pedestrian's with absolutely no interaction with other agents, so agent costs are needed. Other agents in the model who might be expected to affect pedestrian behaviour include drivers, other pedestrians, cyclists, transit vehicles, and even animals, especially if those animals are perceived to be dangerous

Depending on the context, different agents may have different impacts on pedestrians route-choice. For example, pedestrians generally wish to avoid interactions or collisions with vehicles

or unfamiliar pedestrians, increasing perceived costs. On the other hand, pedestrians who are friendly with each other may gravitate towards each other, effectively lowering the cost perceived by the pedestrian. Lastly, the effect of time pressure on pedestrians trying to catch transit vehicles is well documented (Dhoke & Choudhary, 2023). The pedestrian may perceive a lower cost at a node if traversing the node allows them to make a transit connection. For this research, the selected focus is the interactions between pedestrian and vehicles. Specifically, the consequences of risky interactions like collisions with vehicles and the pedestrian’s desire to avoid these interactions.

A method which is conceptually like gap acceptance is proposed. A walking speed is defined as an attribute of each pedestrian, drawn from a normal distribution found in the literature (e.g., where $\mu = 1.34 \text{ m/s}$ and $\sigma = 0.37 \text{ m/s}$ (Sahaleh et al., 2012)). For each road, the time required for the pedestrian to traverse the road is calculated (i.e., the required gap), and the time until the next vehicle will pass that point along the road is calculated (i.e., a “modified” time to collision (TTC)).

For the region along the road where $TTC < gap$, there exists a conflict zone for which the binary variable, $\beta^{Ped-Veh} = 1$. This conflict zone extends in front of the vehicle and includes the vehicles length. Within this zone, the cost of risky pedestrian vehicle interaction is perceived by the pedestrian. The conflict zone concept is shown in **Figure 3.13**.

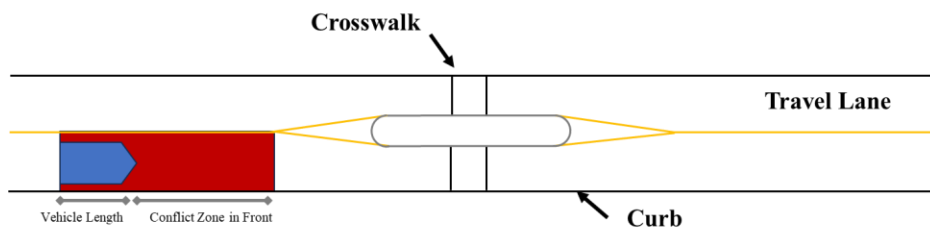
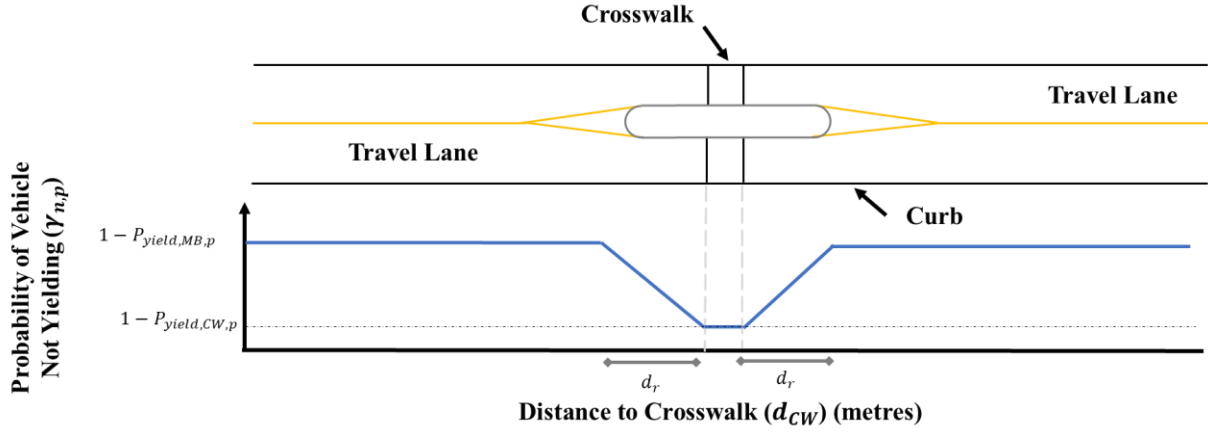


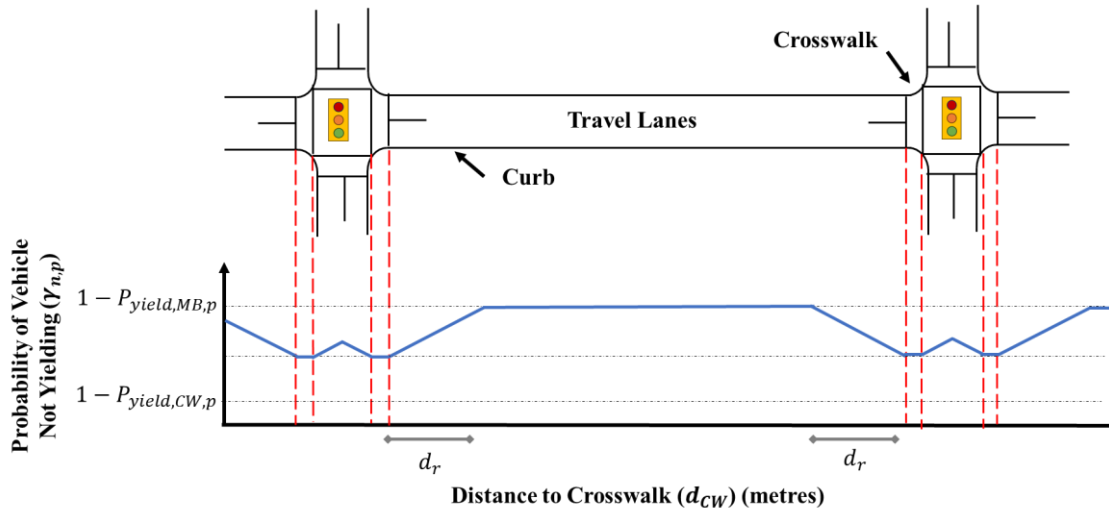
Figure 3.13 Conflict Zone Concept

Near crosswalks, pedestrians have some expectation that vehicles will yield to pedestrians at node n if node n is within a distance d_r of a crosswalk. $\gamma_{n,p}$ is the probability of a vehicle *not* yielding to a pedestrian at node n as perceived by pedestrian p .

At the crosswalk ($d_{CW} = 0$), $\gamma_{n,p}$ is equal to the perceived probability of vehicles not yielding at the crosswalk $\gamma_{n,p} = 1 - P_{yield,CW,p}$, where $P_{yield,CW,p}$ is the perceived probability of vehicles yielding at the crosswalk. Like the legal violation cost, $\gamma_{n,p}$ increases linearly until $d_{CW} > d_r$, after which point it is the perceived probability of vehicles not yielding mid-block, $\gamma_{n,p} = 1 - P_{yield,MB,p}$, where $P_{yield,MB,p}$ is the perceived probability of vehicles yielding at a midblock location. This relationship is visualized in **Figure 3.14a** for a mid-block example and **Figure 3.14b** for a signalized intersection example. $\beta^{Ped-Veh}$ and γ_n are multiplied by the cost of a risky interaction between pedestrians and vehicles $C^{Ped-Veh}$.



a) Midblock Location



b) Signalized Intersection

Figure 3.14 Probability of Not Yielding for a) Midblock and b) Signalized Intersection

The final formulation of the agent costs is shown in **Equation 3.5**:

$$C_{p,n,t}^{agent} = \gamma_n \beta^{Ped-Veh} C^{Ped-Veh} \quad (3.5)$$

Where:

$$\gamma_n = \begin{cases} \frac{(1 - P_{yield,MB}) - (1 - P_{yield,CW})}{d_r} d_{CW} + (1 - P_{yield,CW}) & \text{if } 0 \leq d_{CW} \leq d_r \\ 1 - P_{yield,MB} & \text{otherwise} \end{cases}$$

In the current implementation, vehicle trajectories and traffic signal statuses are extracted from the VISSIM microscopic traffic simulation software. The connection between the simulation and VISSIM is presently one-way, which means that vehicles cannot react to the behaviour of pedestrians simulated using the proposed model. The VISSIM vehicles cannot react to pedestrians to yield, so, $P_{yield,cw,p} = 1$ at stop-controlled intersections, and at signalized intersection crosswalks with walk indications. Midblock, $P_{yield,MB,p} = 0$ is assumed for all pedestrians. For uncontrolled midblock crossings, the probability of yielding depends on the level of control present at the crossing and is assigned on a case-by-case basis. These assumptions should be investigated further in future work. This assumption should be investigated in future work.

$C^{Ped-Veh}$ represents a significant risk to the pedestrian (a likely collision with a vehicle). It is assumed that pedestrians will want to avoid this outcome “at all costs” and are willing to detour as far as required (i.e., the edge of the map) to avoid this and is therefore a function of the map dimensions. The largest detour a pedestrian could take is twice the longest map dimension.

$C^{Ped-Veh}$ used in this research is 141, which reflects twice the longest dimension of the study areas considered in this work. This represents the distance (along a paved surface with a base cost of 1) the pedestrian would be willing to detour. More detail on the study areas is included in subsequent sections.

The agent costs of the grid cells can be visualized as a heatmap, as shown in **Figure 3.15b** for a single time increment. **Figure 3.15a** shows the model environment for reference. At this stage of research, $C_{p,n,t}^{agent}$ only reflect pedestrians’ interaction with vehicles (not other pedestrians) and therefore it is applied to road cells only. Additionally, the one-way connection with VISSIM

means that, $P_{yield,cw,p} = 1$ is assumed at the crosswalk, which causes a gap in the conflict zone at the crosswalk (as shown in **Figure 3.15b**).

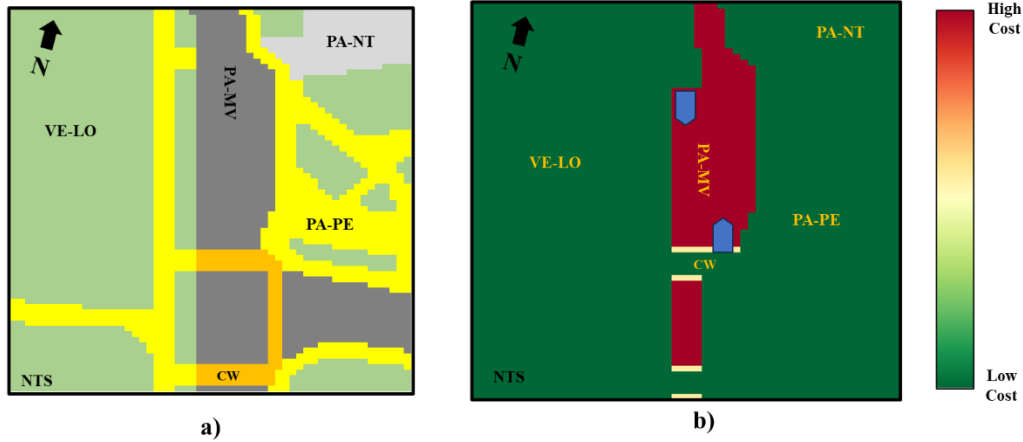


Figure 3.15 a) Model Environment & b) Agent Cost Heatmap

3.4 Pedestrian Awareness and Anticipation

The formulation of costs has been introduced with no limitations on pedestrian perception. Of course, the notion that pedestrian perception is limited is likely familiar to the reader and is documented in the literature.

Obstructions or poor eyesight can limit the spatial area which pedestrians can accurately perceive (Tong & Bode, 2022). This is referred to in this thesis as the *spatial awareness* of pedestrians.

Future events cannot be known with certainty either, referred to as *temporal anticipation*. Indeed, the example used in Chapter 1 implicitly assumes that pedestrian perception has limitations. The pedestrian would not need to alter their path if they had perfect perception of the road for all time into the future, because they would already know that the road would be clear of vehicles and would not first plan different path.

Where (and when) pedestrians have limited perception of the surrounding environment, they must make assumptions based on expectations of future events. This cost for node n and

pedestrian p is referred to as the expected cost, $C_{n,p}^{Expected}$, which is constant over the duration of the simulation. This is distinguished from the live cost, $C_{n,p,t}^{Live}$, which represents the previously formulated cost at node n for pedestrian p at time t . $C_{n,p}^{Expected}$ is calculated as the average of $C_{n,p,t}^{Live}$ over the runtime of the simulation, visualized in **Figure 3.16**. In this work, only the costs of roads change temporally and therefore expected costs only differ from live costs for road cells.

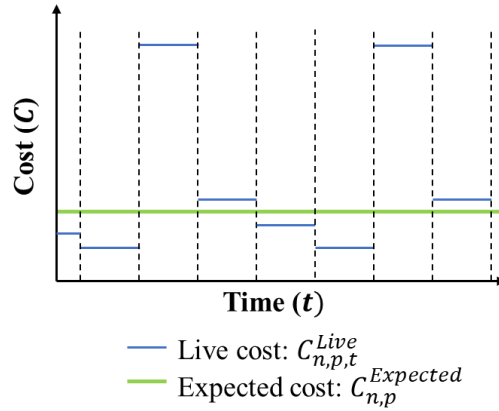


Figure 3.16 Live Costs Over Time Compared to Expected Costs

All pedestrians have perfect spatial awareness of events within radius d_A , in metres, *and* have spatial awareness of events along a road, if any part of that road is within radius d_A . The spatial awareness of a pedestrian is visualized in **Figure 3.17**. d_A is suggested to equal 5 metres, based on existing research into pedestrian gaze patterns (Q. Zhao et al., 2023). In this work, the study area locations are along straight roads with good visibility. However, the presence of obstacles may limit the distance along the road which the pedestrian is able to perceive. Therefore, the distance a pedestrian can see along the road may be restricted on a site-by-site basis. The pedestrian also has perfect anticipation of events for some time, t_A , in seconds, into the future. At a minimum, t_A is the time the pedestrian requires to cross the road. In this research it is assumed that the boundary between expected and live costs is clearly denoted, but it is also possible that

the boundary is much less clear and that pedestrians may consider a weighting between live and expected costs. This concept should be investigated in future studies.

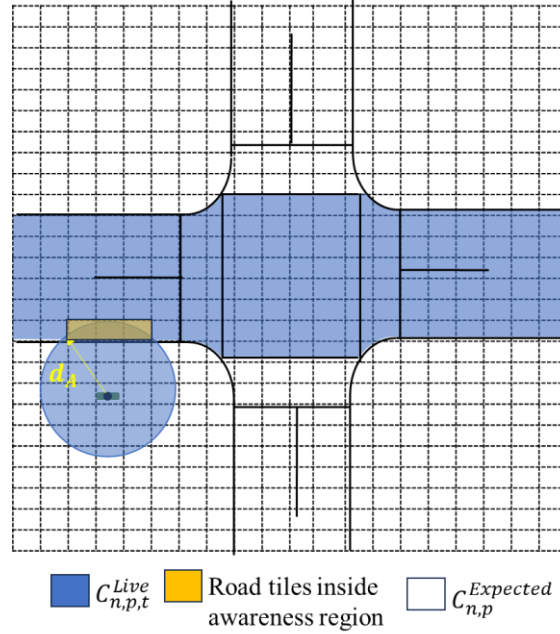


Figure 3.17 Spatial Awareness of a Pedestrian

3.5 Complete Model Formulation and Simulation Structure

The final formulation of the live costs, $C_{n,p,t}^{Live}$, in the current model is given in **Equation 3.6**. The formulation of the total costs can also be visualized by combining the previous heatmaps, as shown in **Figure 3.18**.

$$C_{p,n,t}^{Live} = C_{p,n}^{base} + C_{p,n,t}^{violation} + C_{p,n,t}^{agent} \quad (3.6)$$

Where:

- $C_{p,n}^{base} = \alpha_p^{base} (C_n^{surface} - C_{pavement}) + C_{pavement}$
- $C_{p,n}^{violation} = \alpha_p^{LV} C_{n,t}^{LV}$

$$\begin{aligned}
- C_{n,t}^{LV} &= \begin{cases} \frac{C_{max}^{LV} - C_{CW,t}^{LV}}{\frac{d_{viol}}{2}} d_{CW} & \text{if } 0 < d_{CW} \leq \frac{d_{viol}}{2} \\ \frac{C_{MB}^{LV} - C_{max}^{LV}}{\frac{d_{viol}}{2}} d_{CW} + 2C_{max}^{LV} & \text{if } \frac{d_{viol}}{2} < d_{CW} < d_{viol} \\ 0 & \text{otherwise} \end{cases} \\
- C_{p,n,t}^{agent} &= \gamma_n \beta^{Ped-Veh} C^{Ped-Veh} \\
- \gamma_n &= \begin{cases} \frac{(1 - P_{yield,MB}) - (1 - P_{yield,CW})}{d_r} d_{CW} + (1 - P_{yield,CW}) & \text{if } 0 \leq d_{CW} \leq d_r \\ 1 - P_{yield,MB} & \text{otherwise} \end{cases}
\end{aligned}$$

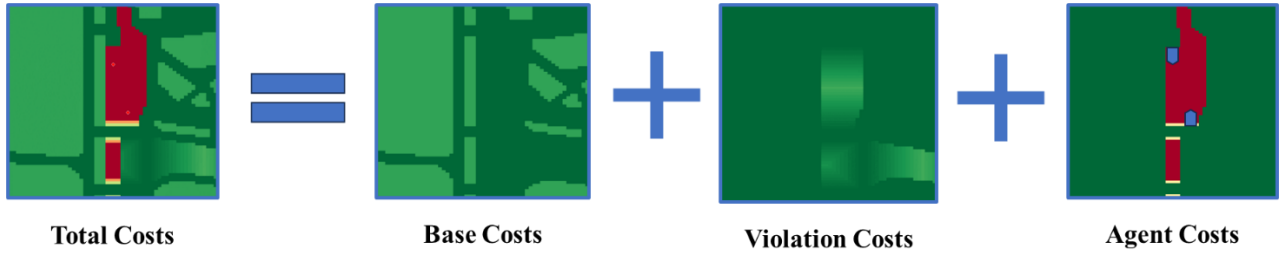


Figure 3.18 Total Cost Visualization

Expected costs, $C_{n,p}^{Expected}$, are calculated as the average of $C_{n,p,t}^{Live}$ over the runtime of the simulation. The model is implemented as a simulation in Python and the code used to implement the model is included in **Appendix A**. **Figure 3.19** shows the workflow of the model for a single pedestrian. In this work, the VISSIM traffic simulation model is used to model the movement of vehicles through the calibration and validation case study networks. For each case study, the network is coded to reflect the real-world geometry, road network configuration, and traffic controls, and vehicle demands are specified to correspond to the observed real world vehicle volumes. The simulation is executed and the vehicle positions and traffic signal status in each time interval are recorded and stored for input to the pedestrian simulation model.

The pedestrian simulation model is executed once for each pedestrian being simulated. The model is executed n times using the same set of vehicle position and traffic signal status input

files to simulate n different pedestrians for a given case study. The A* pathfinding algorithm (with Euclidian distance as the heuristic) is used to find the shortest path based on the live costs (Equation 3.6) and expected costs.

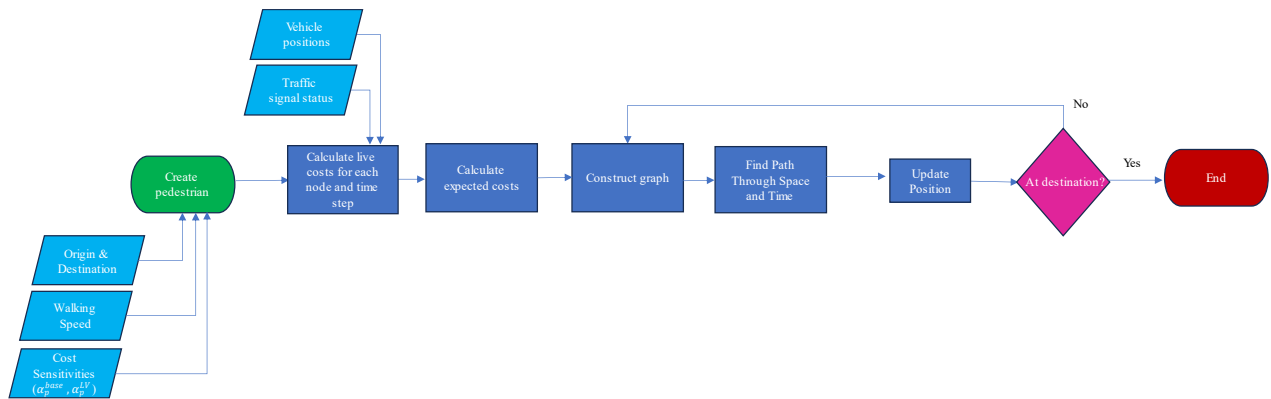


Figure 3.19 Simulation Workflow

4 Model Calibration

Before the model can be used for real-world applications, calibration of the cost parameters is required. For this research, the measure of effectiveness is the proportion of pedestrians performing a given behaviour. $C_{surface}$, and C_{max}^{LV} , are calibrated for the surface costs, and violation costs respectively.

It would be challenging to calibrate both $C_{surface}$ and C_{max}^{LV} simultaneously. Specific scenarios where the other cost terms are negligible and have little impact on the selected route can be used to isolate behaviours related to specific costs. If they are not negligible, values for the other cost terms must already be known.

Calibrating $C_{surface}$ requires investigating a portion of the pedestrian route where pedestrians are forced to select between a more direct route over a more costly surface, or a longer route over a less costly surface. Similarly, calibrating C_{max}^{LV} requires a portion of the pedestrian route where vehicles are not present (i.e., $C^{Ped-Veh} = 0$), and there is no difference between surfaces.

A three-leg all-way stop intersection (along with its surrounding features) on Ring Road near the Science Teaching Complex on the University of Waterloo Campus was selected as the study area for calibration (**Figure 4.1**). The intersection experiences significant pedestrian volumes during the lunchtime period as students walk between student residences and classes on the campus. This location also has only two surface types for pedestrians to traverse (pavement and grass), simplifying the calibration procedure for $C_{surface}$.

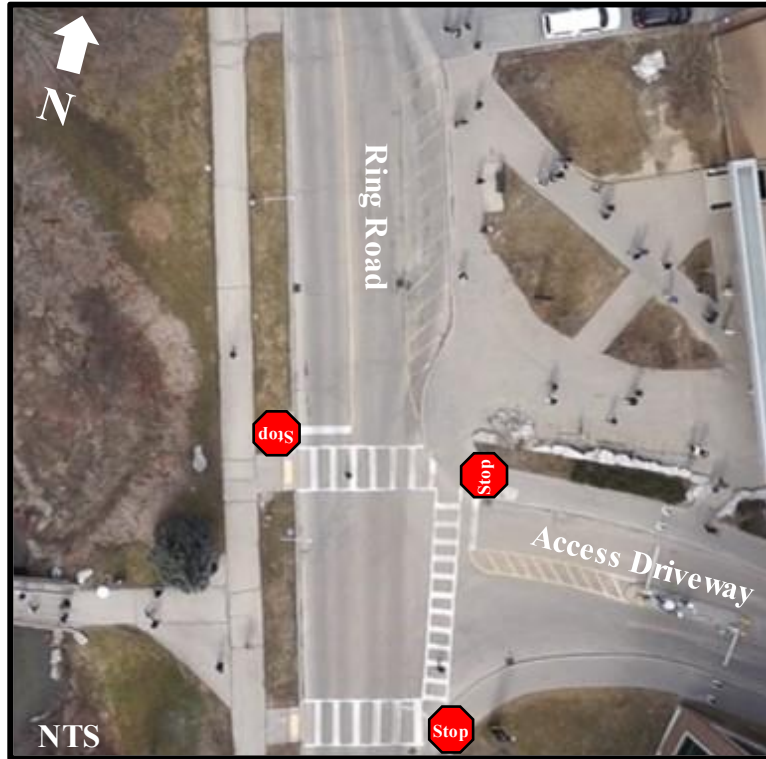


Figure 4.1 Calibration Study Area

Two videos were collected as follows:

- Dataset 1 (15 minutes): Tuesday January 30th, 2024, 12:00pm – captured using handheld video recording device
- Dataset 2 (23 minutes): Wednesday February 7th, 2024, 11:45am – captured using overhead video drone

The handheld video recording device was an iPhone 13 Mini, and the overhead video was captured using a DJI M300 RTK drone with a Sony camera attachment. The aerial photos shown in this thesis are screenshots from the drone video.

Datasets 1 and 2 were combined and used for calibration. For each cost calibration, a different sub-area of the study area was examined.

4.1 Base and Violation Cost Calibration

For $C_{surface}$, the sub-area selected was the intersection of sidewalks, located just west of the intersection. The violation behaviour of interest is pedestrians cutting across the grass instead of walking across pavement, so the surface is grass and the specific cost being determined is C_{grass} . Five possible origins/destinations were also defined. **Figure 4.2** shows the sub-area for C_{grass} , the origins/destinations, and an example of the violation behaviour of interest.

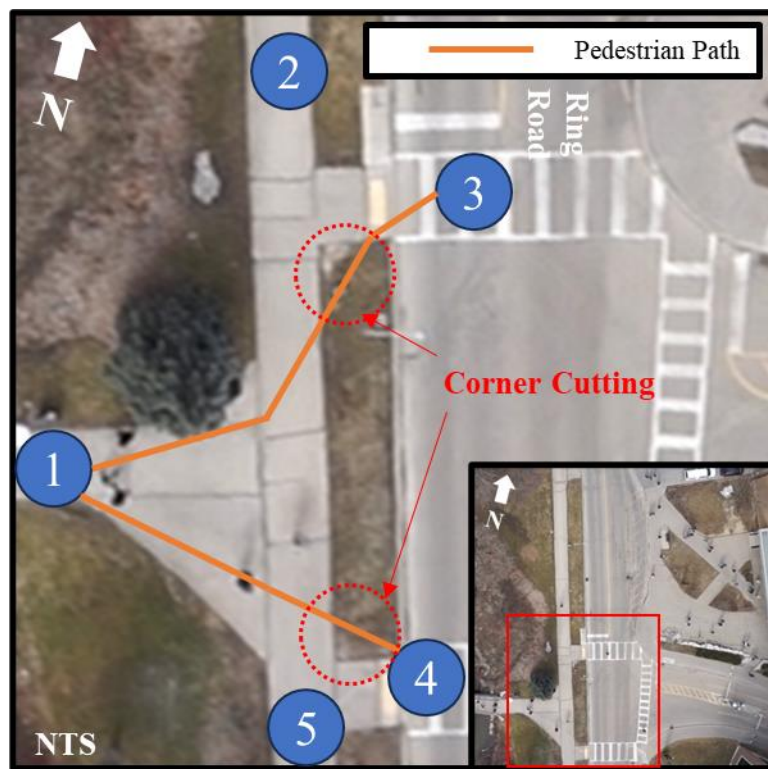


Figure 4.2 C_{grass} Calibration Sub-Area

For C_{max}^{LV} , the sub-area is the north leg crosswalk, where the violation behaviour of interest is pedestrians “corner-cutting” from the crosswalk to the sidewalk, via the road. **Figure 4.3** shows the sub-area, origin/destinations, and an example of the violation behaviour of interest for C_{max}^{LV} .

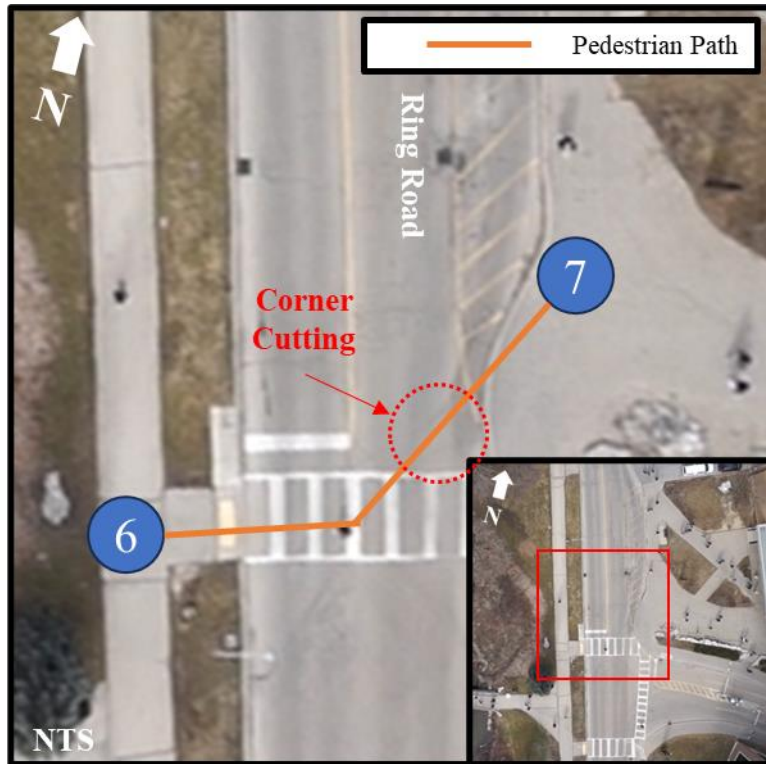


Figure 4.3 C_{max}^{LV} Calibration Sub-Area

For, C_{grass} , many of the O-D pairs have very low sample sizes. Therefore, this discussion will focus on O-D pairs which consistently had volumes greater than 10 pedestrians: O-D pairs 1-3, 1-4, 3-1, and 4-1. 159 pedestrians were observed in total for these four O-D pairs. O-D 3-1 was used for determining the calibrated cost value since it had the highest sample size (62 pedestrians). Likewise, for C_{max}^{LV} , the number of pedestrians crossing in the eastbound direction (O-D 5-6) was relatively low, so the calibrated cost value is determined using only westbound crossings (O-D 6-5) with a sample size of 51 pedestrians. Notably, the observed proportion of pedestrians undertaking a given violation behaviour for a given O-D pair is relatively consistent across both datasets.

Pedestrians were simulated to have origins and destination corresponding to the O-Ds of interest.

The designated origins and destinations for field data collection do not correspond to a precise

map location but rather an approximate location. For example, O-D point 1 in Figure 17 represents the paved path. Since the spatial resolution of the simulation involves 1m x 1m grid cells, pedestrian origins and destination locations were modelled in the simulation as the O-D position centroid plus a random location error between plus/minus two metres.

Different cost values for both C_{grass} and C_{max}^{LV} were selected and simulated 30 times each until the absolute error between the simulated and observed proportions of pedestrians performing violations of interest were minimized.

Figure 4.4a and 4.4b shows the absolute error between simulated and observed proportions at different cost values for C_{grass} and C_{max}^{LV} , respectively. C_{grass} was determined to be 2.5, and C_{max}^{LV} was determined to be 9.5. Notably, the value of 2.5 determined for C_{grass} is similar to the value found in the literature of 2.7 (Kudinov et al., 2018).

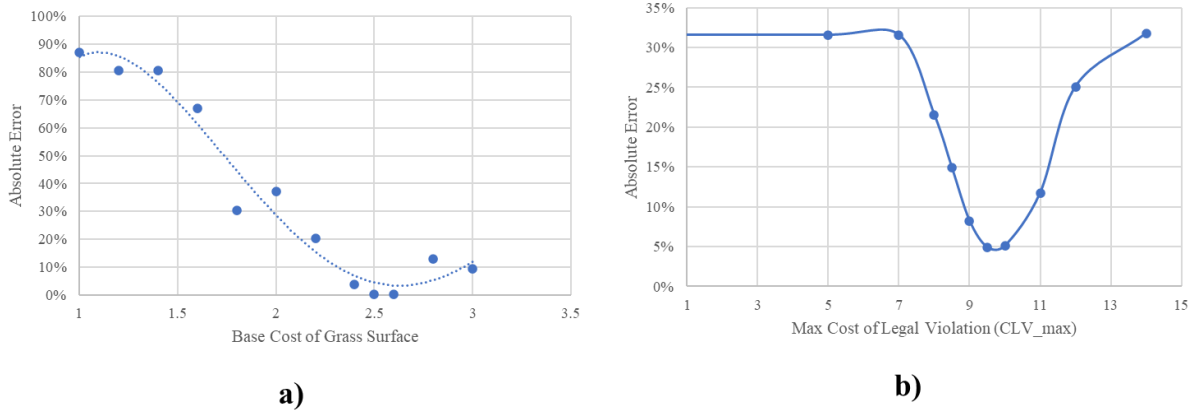
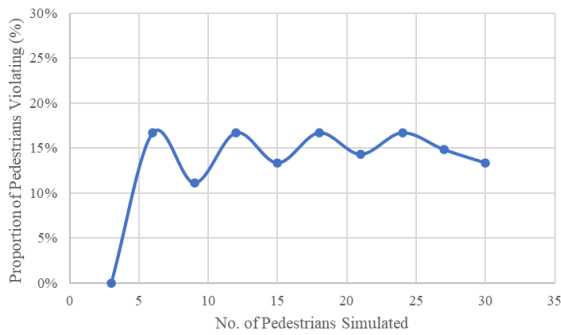


Figure 4.4 Absolute Error Between Simulated & Observed Violation Proportions vs. a) Grass Base Cost & b) Max Legal Violation Cost

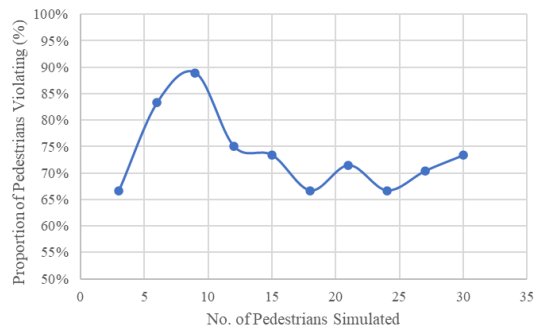
4.2 Simulated Sample Size Sensitivity

During calibration, each cost value was simulated for 30 pedestrians. For the calibration to be valid, it is necessary to verify that 30 pedestrians are sufficient to estimate the effect of the cost

parameters. Therefore, an additional test was performed, where the proportion of pedestrians violating was tested for different sample sizes. This test was performed for both grass costs and the max legal violation cost. The results are shown in **Figure 4.5a and 4.5b** for the grass costs and max legal violation cost, respectively. In both cases, the results obtained fluctuate dramatically at low sample sizes (due to stochasticity) but begin to stabilize around a sample size of 12 and becomes more stable as 30 is approached. This shows that a sample size of 30 is indeed reasonable for this application.



a)



b)

Figure 4.5 Proportion of Violating Pedestrians vs. Sample Size for a) Grass Base Cost & b) Max Legal Violation Cost

5 Case Studies

After initial calibration, three case-study locations were examined for preliminary validation. All locations are in Waterloo, ON. Each location contains a different type of pedestrian facility. The case studies examined are (1) stop controlled intersection crosswalk, (2) midblock crossing of a two-lane collector road, and (3) crosswalk at a signalized intersection.

5.1 Case Study #1: Stop Controlled Intersection Crosswalk

For the first case study, the Ring Road location near the Science Teaching Complex (i.e., same site used for calibration) was used again with identical origins/destinations and behaviours of interest, but a new dataset was collected for validation. This was done to validate that the calibrated parameters were not overly fitted to the calibration dataset. 23 minutes of drone footage was captured on Wednesday February 14th, 2024, beginning at 11:45am.

Table 5.1 summarizes the observed results and the simulation results for the OD Pairs identified above with 10 or more pedestrians. For each O-D pair, 30 pedestrians were simulated, and a hypothesis test (Fischer's Exact Test) was performed between the simulated and observed proportions at the 95% confidence level to determine if the difference between simulated and observed results is statistically significant. Fisher's Test was performed using R (see **Appendix B**). The null hypothesis is that the two proportions are not different. In this case, failing to reject the null hypothesis (H_0) is desirable. The mean absolute error (MAE) is computed as the volume weighted mean of the absolute errors for each O-D pair.

Table 5.1 Case Study #1 Results

O-D Pair	Observed Sample	% of Pedestrians Violating		Absolute Error Simulated - Observed	Ho (p_sim = p_obs)
		Observed	Simulated		
3-1	100	12.00%	13.3%	1.3%	Do not reject
1-4	20	10.00%	6.7%	3.3%	Do not reject
4-1	25	12.00%	3.3%	8.7%	Do not reject
6-7 (EB)	25	60.00%	66.7%	6.7%	Do not reject
Mean Absolute Error (MAE) %				3.4%	

5.2 Case Study #2: Midblock Crosswalk

The selected midblock crosswalk is located at 256 Phillip Street in Waterloo. Phillip Street is a two-lane, north-south collector road with sidewalks located on both sides. A pedestrian island is located between the buildings at 250 and 256 Phillip Street. To the west of Phillip Street is the University of Waterloo campus, and to the east there are several high-rise residential buildings with ground floor commercial space. Consequently, there are a substantial number of crossings made on Phillip Street, many of which involve crossing outside of the crosswalk (i.e., “jaywalking”, a violation behaviour). In the middle of Phillip Street, there is a two-way left-turn lane for vehicle access to nearby businesses.

Along this stretch of Phillip Street, a crosswalk is provided which connects a walking path to the University of Waterloo campus with the buildings at 250 and 256 Phillip Street. The crosswalk does not have pavement marks, but a pedestrian refuge island is provided in the two-way left-turn lane and both the island, and adjacent curbs include curb ramps and tactile ground surface indicator plates. Vehicles are not required to yield to pedestrians. However, the crossing lacks the required “Pedestrians Yield to Vehicles” Signage required by the Ontario Traffic Manual for uncontrolled crossings (Ministry of Transportation, 2016b). Consequently, it is common for

vehicles to yield to pedestrians even though it is not required. Therefore, absent a more detailed investigation into pedestrians' perceived probability of yielding, $P_{yield,CW,p} = 1$ is assumed.

Six possible origins/destinations were defined. The behaviours of interest for this location are crossings which do not use the crosswalk for the duration of the crossing. All O-D pairs that require crossing of Phillip Street, except O-D 2-5 and 5-2, produce behaviours of interest for modelling.

Figure 5.1 shows the study area, the origins/destinations, and examples of the behaviours of interest. For clarity, not all O-D pairs of interest have a behaviour (pedestrian path) shown between them.

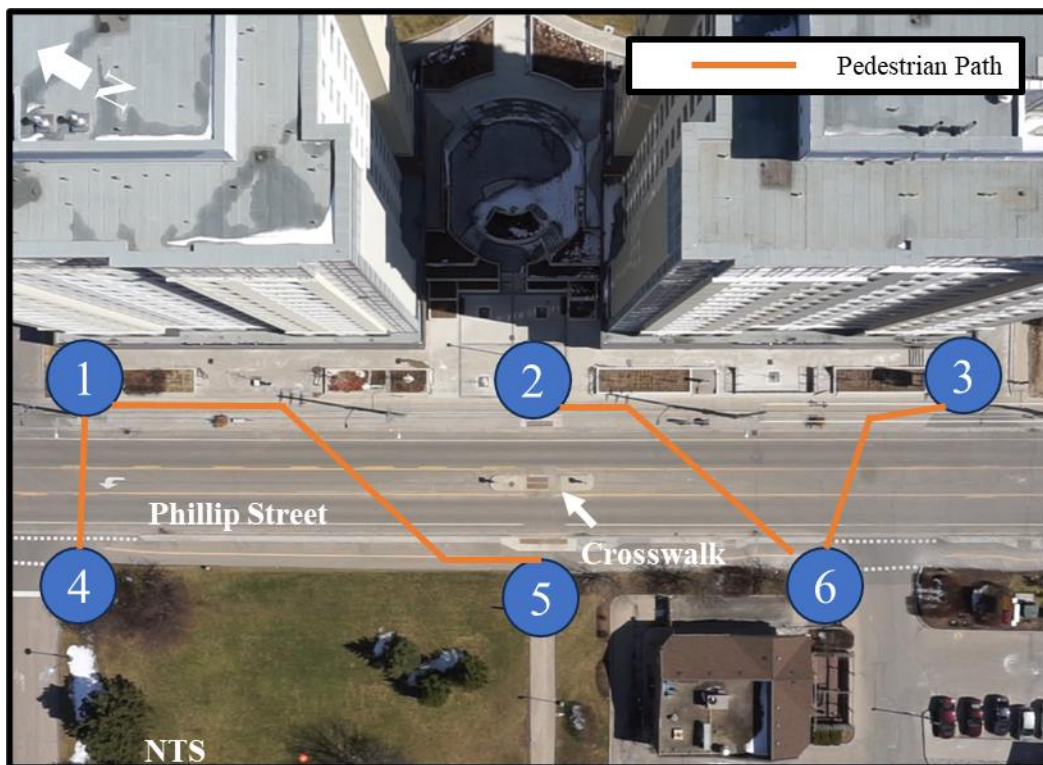


Figure 5.1 Case Study #2 Study Area

Approximately 52 minutes of drone video footage of the study area was collected on March 25, 2024, beginning at 11:40am. The bidirectional traffic volume on Phillip Street was 462 vph. For each O-D pair where the pedestrian volume was 10 or more, 30 pedestrians were simulated. The observed and simulated proportions of pedestrians engaging in violation behaviour (i.e., crossing outside the crosswalk) are provided in **Table 5.2**.

Table 5.2 Case Study #2 Results

O-D Pair	Observed Sample	% of Pedestrians Violating		Absolute Error Simulated - Observed	Ho (p_sim = p_obs)
		Observed	Simulated		
1-4	40	100.0%	100.0%	0.0%	Do not reject
4-1	14	100.0%	100.0%	0.0%	Do not reject
1-5	11	72.7%	80.0%	7.3%	Do not reject
5-1	12	58.3%	70.0%	11.7%	Do not reject
3-5	66	75.8%	66.7%	9.1%	Do not reject
5-3	27	66.7%	63.3%	3.3%	Do not reject
3-6	36	100.0%	100.0%	0.0%	Do not reject
6-3	10	100.0%	100.0%	0.0%	Do not reject
Mean Absolute Error (MAE) %				4.2%	

5.3 Case Study #3

University Avenue West and Phillip Street was selected as the location for Case Study #3. The intersection of University Avenue West & Phillip Street is a signalized intersection located approximately 150 metres south of 256 Phillip Street, and therefore the attributes of Phillip Street are largely identical to those described at 256 Phillip Street. University Avenue West is a four-lane, east-west arterial roadway with sidewalks located on both sides. At the intersection of University Avenue and Phillip Street, there are dedicated left-turn lanes on all approaches. There are no dedicated right-turn lanes. The south approach of the intersection provides access to a

university student residence complex and is essentially a driveway. The study area of Case Study #3 is shown in **Figure 5.2**.

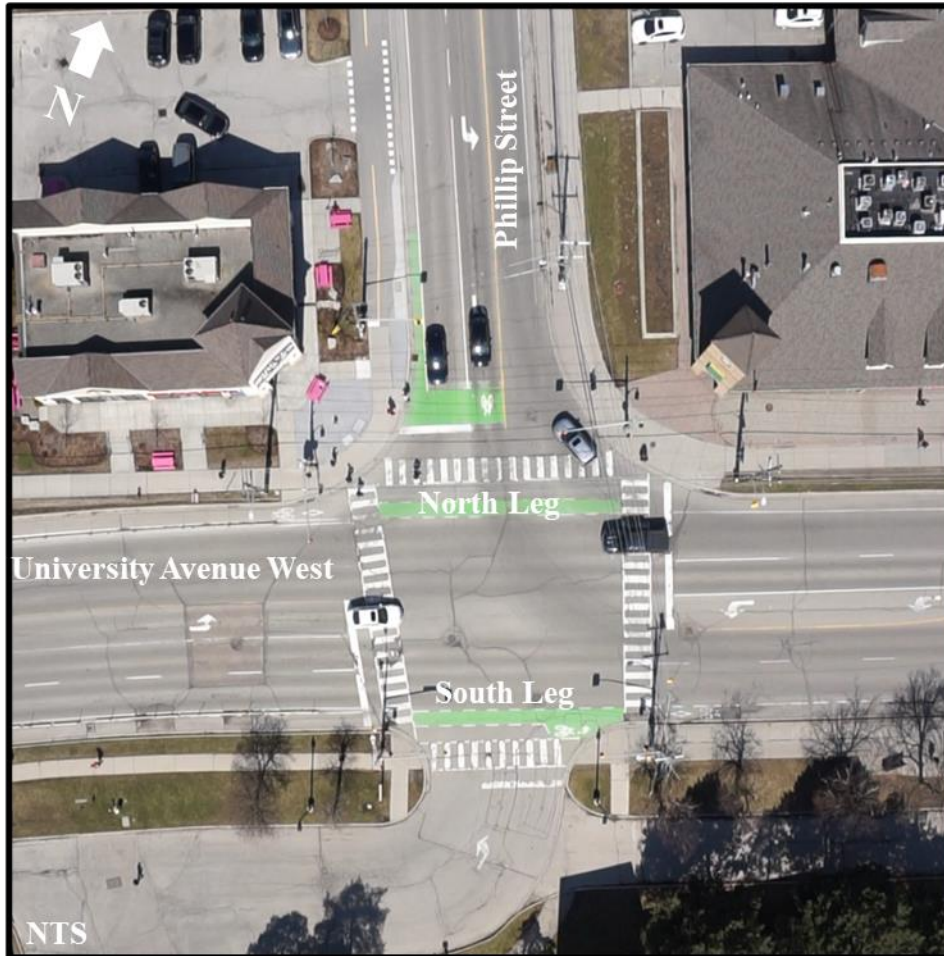


Figure 5.2 Case Study #3 Study Area

Data for Case Study #3 were collected on March 28, 2024, for 50 minutes beginning at approximately 11:50am. In total, 112 pedestrians were observed crossing the north and south legs.

The application of the model to a signalized intersection brings two new challenges which must be discussed first. The first is the pedestrian's anticipation time (t_A). At a signalized intersection, the pedestrian is aware that the signal will change at some time in the future. If they have been to

the intersection before, they may even know the durations of each signal phase. Therefore, t_A for a signalized intersection must be extended to include the maximum waiting time (i.e., the “do not walk interval” for the crosswalk) plus some additional time to allow for crossing. The signal timings were recorded in the field and t_A was set to 50 seconds. If t_A is set shorter than maximum waiting time, the pedestrian will not know the signal phase will change to “walk” in the future. Consequently, the pedestrian will never choose to wait.

The second challenge relates to vehicle conflict zones when a vehicle is approaching an intersection at which the vehicle may make a turning movement. Pedestrians cannot always anticipate the movement a vehicle will make at an intersection. Drivers may signal late or not at all, forcing the pedestrian to assume the vehicle may make any turning movement. In this thesis, this is referred to as propagating the conflict zone beyond the vehicle’s current road section. The propagation affects all lanes that are downstream of the vehicle’s path (i.e., lanes the vehicle could potentially occupy in the future). **Figure 5.3** illustrates the concept of the conflict zone propagation for a vehicle approaching an intersection.

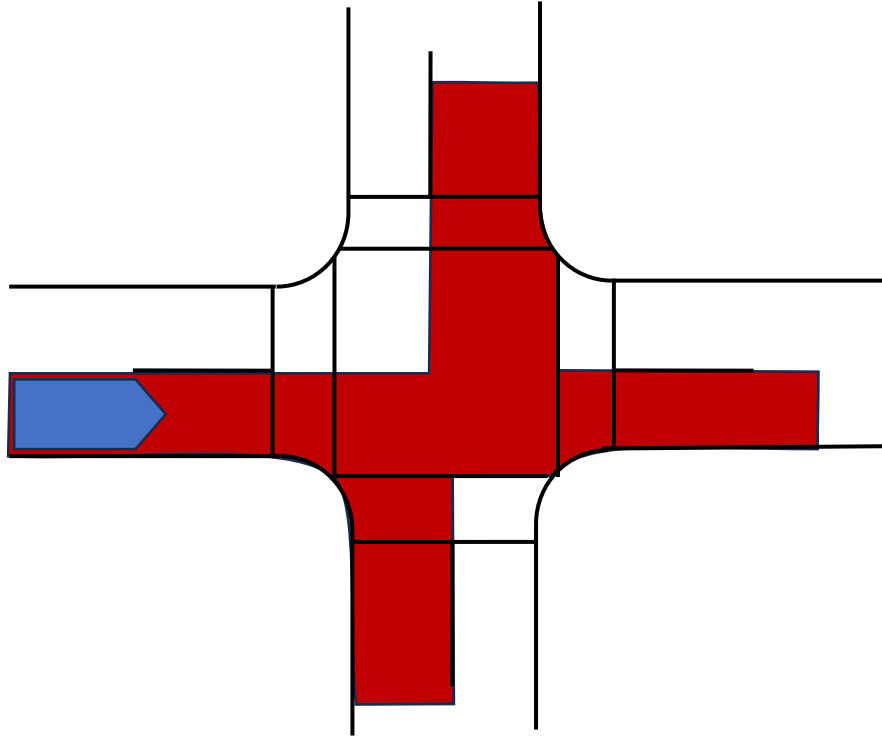


Figure 5.3 Illustration of Vehicle Conflict Zone Propagation

The effects of propagation are related to expected costs. It is possible for an approaching vehicle to make any turn, however some turning movements are more common than others. In principle, this could reduce the perceived cost of the conflict zone for low volume turning movements. In this research, all possible turning movements are considered equally likely and therefore the propagated conflict zone is the same for all turning movements, as shown in **Figure 5.3**.

Signalized intersections permit us to examine a pedestrian behaviour not previously investigated in the other case studies: pedestrians entering the intersection against the Flashing Do Not Walk (FDW) or Do Not Walk (DNW) signal indications. In the drone video recorded at the site, signal indications are inferred from the movement of vehicles. The FDW indication occurs during the green phase for vehicle traffic and cannot be distinguished from the Walk indication in the video. Therefore, only the DNW phase is examined in this case study.

When a pedestrian enters the crosswalk during the Do Not Walk phase, they would experience a cost of $C_{CW,t}^{LV}$. During the Walk indication, there is no violation so $C_{CW,t}^{LV} = 0$. The value of $C_{CW,t}^{LV}$ during the Do Not Walk phase must be considered carefully. $C_{CW,t}^{LV} = 0$ (no violation) and $C_{CW,t}^{LV} = 9.5$ (violation) were taken from initial calibration (**Figure 4.4b**). $C_{CW,t}^{LV} = 5$ was also examined to provide an approximate midpoint between the two values. 30 pedestrians were simulated crossing each of the North and South leg crosswalks. **Figure 5.4** summarizes the results.

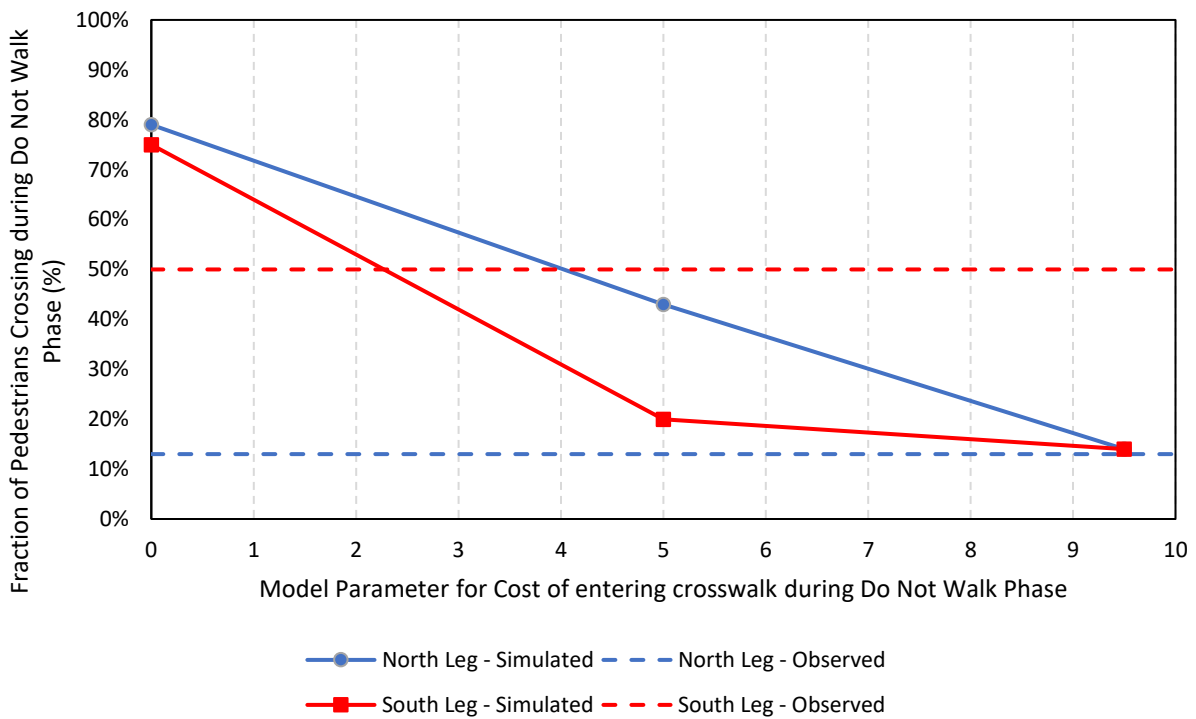


Figure 5.4 Calibration results for Pedestrians Crossing Against Do Not Walk (DNW) Signal

5.4 Discussion

The results show that the model can predict the proportion of pedestrians performing a violation behaviour on their way from a given origin to destination. For Case Study #1, the data collected

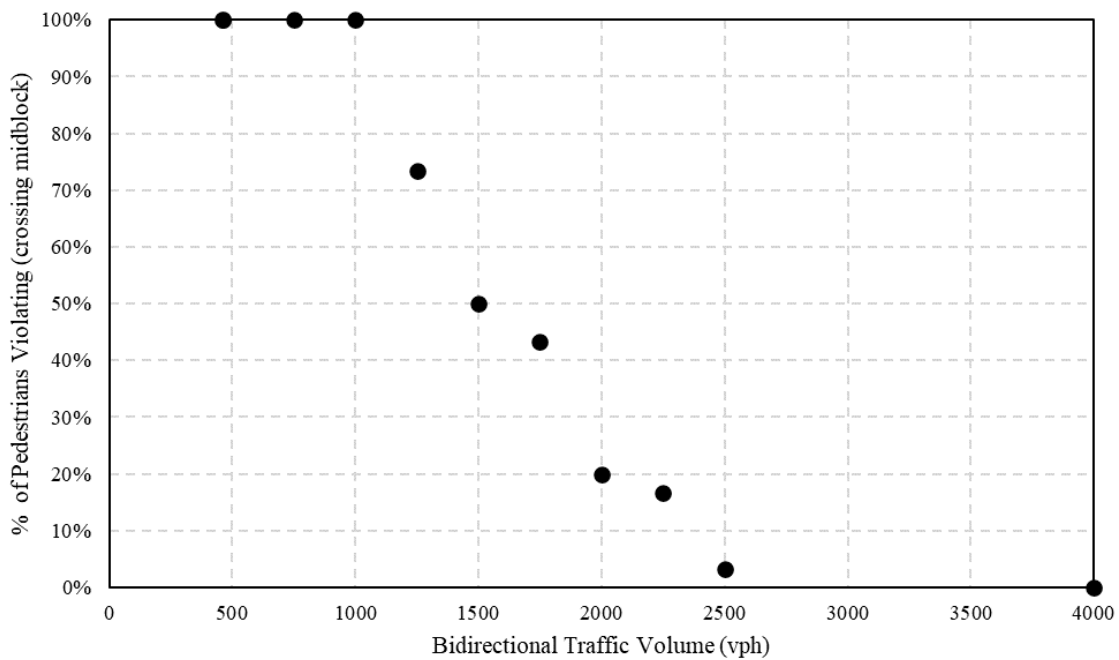
for validation were similar to the data used for calibration (same site), so it is not surprising that the simulation model was able to reproduce the proportions of violations across all O-D pairs.

The MAE for Case Study #1 was 3.4%

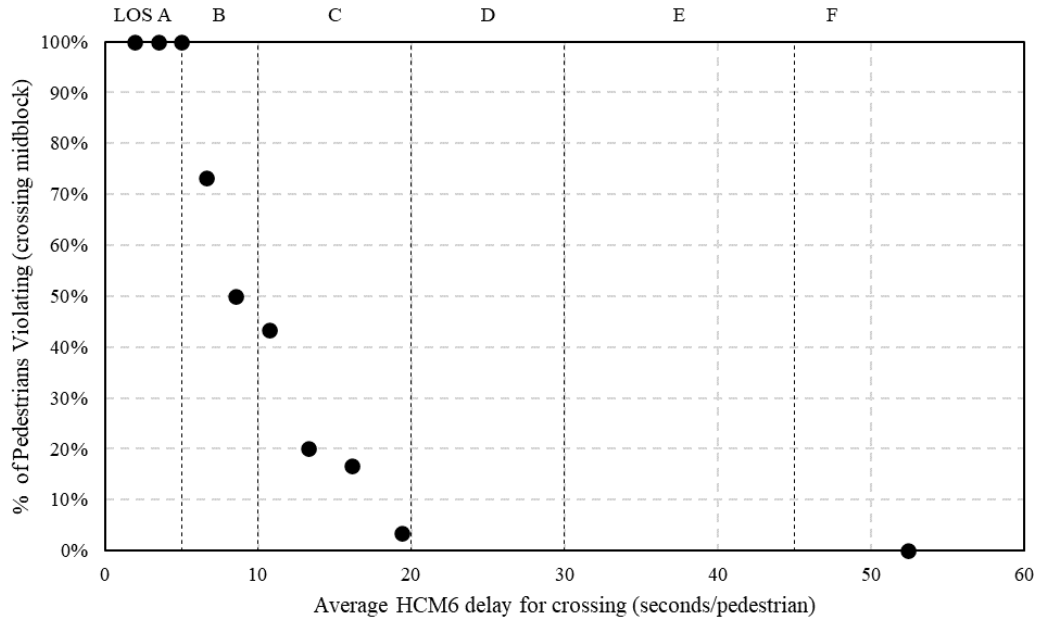
For Case Study #2, all the pedestrians observed to walk between OD pairs 1-4, 4-1, 3-6, and 6-3, chose to cross midblock rather than at the crosswalk (i.e., 100% of pedestrians engaged in violation behaviour). The simulation model also predicted 100% violation behaviour for these O-D pairs. Though this result is encouraging, it is possible the simulation model will predict 100% violation behaviour regardless of traffic volumes.

We expect that as traffic volumes increase, the time headways between vehicles become smaller, and pedestrians would need to wait longer before they could cross midblock. Although not obligated, some vehicles do yield to pedestrians on Phillip Street. So, as the wait time increases, we expect more pedestrians to cross at the crosswalk instead, where a vehicle is likely to yield, permitting the pedestrian to cross. This behaviour was investigated by simulating 30 pedestrians travelling from O-D 1-4 for a range of different traffic volumes on Phillip Street. These results are shown in **Figure 5.5**. **Figure 5.5a** shows that as the bidirectional traffic volume increases, the fraction of pedestrians engaging in violations (i.e., crossing at midblock) decreases as pedestrians increasingly choose to use the crossing facilities rather than risk colliding with a vehicle. This result is consistent with expectations. **Figure 5.5b** shows the simulated fraction of pedestrians crossing at midblock as a function of the average delay pedestrians are estimated to experience to cross at midblock. The average delay was estimated using the Highway Capacity Manual Version 6 methodology assuming that pedestrians crossing at midblock treat the crossing as a two-stage crossing and use the central dual left turn lane as a temporary refuge (behaviour that was observed in the field). The reported average delays are the sum of the delay from each stage

of the two-stage crossing. The traffic volumes evaluated represent a range in Level of Service for pedestrians from A to F. The correlation coefficient between the average HCM6 pedestrian delay and the simulated fraction of pedestrians crossing midblock is -0.74. These results confirm that the proposed pedestrian simulation model produces pedestrian violation behaviour that is expected. As the length of time pedestrians are expected to wait to cross midblock increases, they increasingly chose to cross at the designated crossing.



(a) Simulated Fraction of Pedestrians Crossing at Midblock (violations) as a Function of Traffic Volume



(b) Simulated Fraction of Pedestrians Crossing at Midblock (violations) as a Function of Average HCM6 Pedestrian Delay

Figure 5.5 Sensitivity of Simulated Pedestrian Violations to Traffic Volume and Pedestrian Delay

The results for Case Study #2 for O-D pairs 1-5, 5-1, 3-5 & 5-3 are also encouraging but warrant further discussion. The measure of effectiveness for these tests is not the geometrical similarity between observed and simulated paths and so the spatial differences between observations and simulations are not quantified in this work. However, it was noted during testing that simulated pedestrians tend to cross further from the crosswalk than they were observed to in the case study.

Figure 5.6 illustrates a typical example of this difference.

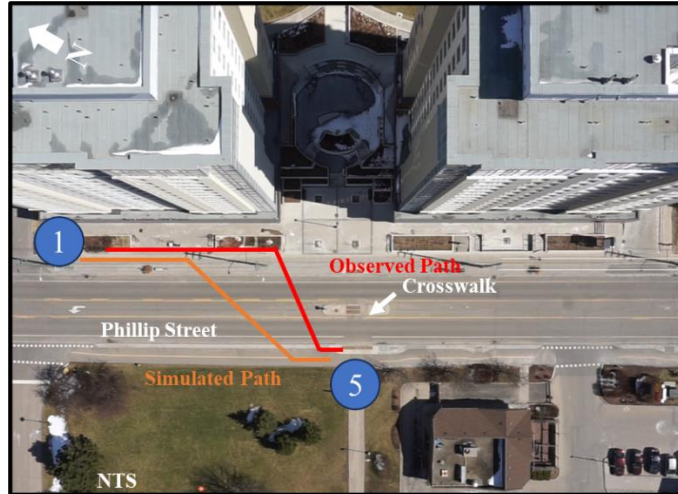


Figure 5.6 Spatial Differences Between a Common Observed Path and Simulated Path

This discrepancy suggests that, while the violation behaviour is correctly predicted by the simulation model (i.e., the pedestrian crosses diagonally outside of the crosswalk), the finer spatial details, such as where specifically pedestrians depart from the curb, may not be accurately reproduced with current parameter estimates and assumptions. In this specific case, pedestrians appear to be more likely to cross closer to the crosswalk than the simulation suggests. This may be because the maximum violation distance d_{viol} was not calibrated but was instead assumed based on legal precedents instead of observed values. Increasing d_{viol} would extend the distance from the crosswalk where crossing is considered a violation, thus incentivizing pedestrians to cross closer to the crosswalk. Further study is needed to examine the accuracy of simulated path trajectories.

The results for Case Study #3 are more mixed. We were unable to get accurate model results for both the north and south crosswalk with a single value of $C_{CW,t}^{LV}$ (the legal violation costs at a crosswalk). As shown in **Figure 5.4**, $C_{CW,t}^{LV}$ of 9.5 from initial calibration produces very good results for the North Leg crosswalk, while none of the tested values produce good results for the south leg of the crosswalk. The south leg of University Avenue & Phillip Street is a low-volume

private access to a university residence complex. Since it is not a public roadway, this may mean that this access is a unique case with attributes not fully captured in the model. On private roadways, the calibrated parameter values may change. Further investigation with a larger dataset of public and private roadways is recommended.

5.5 Limitations

To the author's knowledge, this is a first attempt at explicitly modelling the trade-offs of pedestrian behaviours and integrating the tactical and operational behaviours of pedestrian into a single model. As such, there are many limitations of the model in its current form, as well as limitations with the implementation.

This model has limited scope and focuses on legal violations and vehicles as agents that pedestrians must avoid. While this is a good starting point, there are other factors which can influence pedestrian behaviour including but not limited to other pedestrians, time pressure (e.g., trying to catch a bus), long term surface changes such as the trampling of grass and many more. The pedestrian kinematics included in the model are also limited. Pedestrians can only walk at a constant speed or wait. There is also no consideration for the slope of the ground, or the weather conditions.

This model also assumes pedestrians have an initial knowledge of the features of the environment (aside from vehicles). That is, pedestrians know the locations of all roadways, crosswalks, vegetation areas and other features in advance; the only thing pedestrians "discover" during the simulation is the location and speeds of vehicles. Blocked visibility due to obstacles is also not considered. Another assumption is that pedestrians can accurately perceive the

environmental factors that impact route choice. In other words, there are no situations where a pedestrian “misjudges” the position or speed of a vehicle, leading to a near-miss or collision.

There are a couple limitations related to the implementation. The first is the one-way interaction between VISSIM and the pedestrian simulation, which prevents vehicles from yielding to pedestrians, limiting the types of interactions that can be modelled. Future work should develop a two-way interaction between the pedestrian model and vehicle simulation model. Another limitation relates to the 1m x 1m square grid model environment. Diagonal movements on square grids pose a challenge since a diagonal movement is slightly longer than a straight movement. In this simulation, the costs are adjusted for the diagonal movements by multiplying by $\sqrt{2}$, but the time step for diagonal movements is not adjusted. Therefore, diagonal movements are subject to a small speed error causing pedestrians to arrive at the next node slightly earlier than they would otherwise. Most of the routes explored in this work are predominantly straight lines and pedestrian travel time is not the measure of effectiveness for this work, so this error is unlikely to be a significant factor, but future work should address this issue, either by dynamically adjusting the time steps during the simulation, or by considering a different form of grid where all possible movements are equal distance (e.g., a hexagonal grid).

Lastly, this model introduces many different parameters that do not have well established values in the literature. This thesis described the calibration of several of these parameters, but future work should expand the calibration efforts to capture a wider range of traffic and pedestrian environments.

6 Conclusions

This thesis proposes a new pedestrian simulation model integrating tactical level route-choice with operational level crossing decisions to predict pedestrian violation behaviours. Each pedestrian attempts to obtain the least cost path between a given origin and destination and continues to update their optimal path while enroute to their destination. The perceived cost is a function of base costs, which reflect the ease of moving over the surface; legal violation costs, which reflect pedestrian's perceptions of what actions are right and wrong according to codified laws and rules; and agent costs, which in this work reflect pedestrians perceived risk of the dangers of having conflicts with vehicles.

The model was calibrated using field data obtained on Ring Road at the University of Waterloo campus in Waterloo, Ontario, Canada. After calibration, the capabilities and weaknesses of the model were explored through three test cases. Case Study #1 involved the same site used for calibration but considered a different observed dataset. Case Study #2 was a mid-block crossing location on a collector roadway. These results showed that the model simulates the trade-offs between waiting for a gap in traffic to cross, versus diverting to the nearest designated crosswalk.

In Case Study #3, the model was applied to the signalized intersection at University Avenue and Phillip Street in Waterloo. Two crosswalks were examined for pedestrians crossing against the Do Not Walk signal. The model successfully reproduced the proportion of pedestrians crossing against the signal for the north leg crosswalk but did not reproduce the proportion of violations for the south leg crosswalk, which is across a private access, and further investigation should be undertaken into the causes of this.

In all, these results suggest that the proposed model accurately reflects spatially aggregate pedestrian path choice behaviour, including violation behaviour, for the conditions evaluated.

6.1 Future Work

Since this model is (to the author's knowledge) a first attempt, there are many opportunities for future research. The first and most obvious avenue is to expand the model to include additional aspects of pedestrian behaviour and kinematics. This includes the ability for the pedestrian to vary their walking speeds, introducing additional agents to the model (e.g., transit vehicles, other pedestrians, etc.), and implementing social contract violations into the simulation. Additionally, only two base costs (pavement and light vegetation) were investigated in this work. Future work should consider additional surfaces (e.g., gravel, dense vegetation, etc.) and additional attributes of the surface such as wear, weather factors (wet vs. snow vs. dry), and surface slope. Additional work to define the boundary between expected and live costs, is also recommended.

Many of the current model parameters are assumed. This includes the sensitivity parameters (assumed triangular distribution), the probability of yielding, the violation distance, and spatial awareness. To the author's knowledge, these parameters (or equivalents) are not widely available in the literature, so surveys could be conducted to support appropriate values. A sensitivity analysis should also be conducted to prioritize which parameters to investigate first.

There are also several implementation changes that should be explored. For example, the one-way interaction between the external VISSIM traffic simulation and the pedestrian simulation limits the capabilities of the model, since vehicles cannot react to pedestrians. This mandated various assumptions for model parameters. Future work should investigate opportunities to either better link VISSIM and the pedestrian model or integrate traffic simulation into the model implementation. Future implementations should also investigate different network representations which would provide a more accurate simulation of pedestrian speeds. This would tie in nicely with future work on variable pedestrian speeds.

Lastly, the sample sizes in this work were limited due to time constraints. Future work should be conducted across a much larger sample of pedestrian environments and investigate the transferability of calibrated values between sites.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Al-Widyan, F., Al-Ani, A., Kirchner, N., & Zeibots, M. (2017). An effort-based evaluation of pedestrian route choice. *Scientific Research and Essays*, 12(4), 42–50. <https://doi.org/10.5897/SRE2017.6482>
- Antić, B., Pešić, D., Milutinović, N., & Maslač, M. (2016). Pedestrian behaviours: Validation of the Serbian version of the pedestrian behaviour scale. *Transportation Research Part F: Traffic Psychology and Behaviour*, 41, 170–178. <https://doi.org/10.1016/j.trf.2016.02.004>
- Arentze, T. A., & Timmermans, H. J. P. (2009). A need-based model of multi-day, multi-person activity generation. *Transportation Research Part B: Methodological*, 43(2), 251–265. <https://doi.org/10.1016/j.trb.2008.05.007>
- Barton, B. K., Kologi, S. M., & Siron, A. (2016). Distracted pedestrians in crosswalks: An application of the Theory of Planned Behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, 37, 129–137. <https://doi.org/10.1016/j.trf.2015.12.012>
- Bastariento, F. F., Hancock, T. O., Choudhury, C. F., & Manley, E. (2023). Agent-based models in urban transportation: Review, challenges, and opportunities. *European Transport Research Review*, 15(1), 19. <https://doi.org/10.1186/s12544-023-00590-5>
- Bhowmick, D., Winter, S., & Stevenson, M. (2019, September 19). *Comparing the costs of pedestrian wayfinding heuristics across different urban network morphologies*. <https://doi.org/10.17608/k6.auckland.9846137.v1>

- Boston Transportation Department. (2017). *Go Boston 2030*.
https://www.boston.gov/sites/default/files/file/document_files/2017/03/go_boston_2030_-_6_boston_in_2030_spreads.pdf
- Canca, D., Zarzo, A., Algaba, E., & Barrena, E. (2013). Macroscopic attraction-based simulation of pedestrian mobility: A dynamic individual route-choice approach. *European Journal of Operational Research*, 231(2), 428–442. <https://doi.org/10.1016/j.ejor.2013.05.039>
- Chen, C. F., & Chao, W. H. (2011). Habitual or reasoned? Using the theory of planned behavior, technology acceptance model, and habit to examine switching intentions toward public transit. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(2), 128–137. <https://doi.org/10.1016/j.trf.2010.11.006>
- Chen, Chen, X., Wang, R., & Gao, M. (2022). Characterization of Pedestrian Crossing Spatial Violations and Safety Impact Analysis in Advance Right-Turn Lane. *International Journal of Environmental Research and Public Health*, 19(15), 9134.
<https://doi.org/10.3390/ijerph19159134>
- City of Hamilton. (2018). *City of Hamilton Transportation Master Plan Review and Update*.
<https://www.hamilton.ca/sites/default/files/2022-08/masterplan-transportation-update-2018.pdf>
- City of Minneapolis. (2020). *Transportation Action Plan*. <https://go.minneapolismn.gov/final-plan>
- Cooper, B. S., & Cowlagi, R. V. (2018). Path-planning with waiting in spatiotemporally-varying threat fields. *PLoS ONE*, 13(8), e0202145. <https://doi.org/10.1371/journal.pone.0202145>
- Daamen, W., & Transport, I. (2004). *Modelling Passenger Flows in Public Transport Facilities*.

- Deb, S., Strawderman, L., DuBien, J., Smith, B., Carruth, D. W., & Garrison, T. M. (2017). Evaluating pedestrian behavior at crosswalks: Validation of a pedestrian behavior questionnaire for the U.S. population. *Accident Analysis & Prevention, 106*, 191–201. <https://doi.org/10.1016/j.aap.2017.05.020>
- Dhoke, A., & Choudhary, P. (2023). Is there a relationship between time pressure and pedestrian non-compliance? A systematic review. *Transportation Research Part F: Traffic Psychology and Behaviour, 93*, 68–89. <https://doi.org/10.1016/j.trf.2023.01.001>
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik, 1*(1), 269–271. <https://doi.org/10.1007/BF01386390>
- Dijkstra, J., Timmermans, H., & Jessurun, J. (2014). Modeling Planned and Unplanned Store Visits within a Framework for Pedestrian Movement Simulation. *Transportation Research Procedia, 2*, 559–566. <https://doi.org/10.1016/j.trpro.2014.09.096>
- Dinh, D. D., Vŭ, N. H., McIlroy, R. C., Plant, K. A., & Stanton, N. A. (2020). Effect of attitudes towards traffic safety and risk perceptions on pedestrian behaviours in Vietnam. *IATSS Research, 44*(3), 238–247. <https://doi.org/10.1016/j.iatsr.2020.01.002>
- Elliott, M. A., & Baughan, C. J. (2004). Developing a self-report method for investigating adolescent road user behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour, 7*(6), 373–393. <https://doi.org/10.1016/j.trf.2004.10.002>
- Evans, D., & Norman, P. (2003). Predicting adolescent pedestrians' road-crossing intentions: An application and extension of the Theory of Planned Behaviour. *Health Education Research, 18*(3), 267–277. <https://doi.org/10.1093/her/cyf023>

- Forward, S. E. (2009). The theory of planned behaviour: The role of descriptive norms and past behaviour in the prediction of drivers' intentions to violate. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(3), 198–207.
<https://doi.org/10.1016/j.trf.2008.12.002>
- Friend, C. (n.d.). Social Contract Theory | Internet Encyclopedia of Philosophy. In *Internet Encyclopedia of Philosophy*. Retrieved March 3, 2024, from <https://iep.utm.edu/soc-cont/>
- Fruin, J. J. (1987). *Pedestrian Planning and Design*. Elevator World.
- Ghomi, H., & Hussein, M. (2022). An integrated text mining, literature review, and meta-analysis approach to investigate pedestrian violation behaviours. *Accident Analysis & Prevention*, *173*, 106712. <https://doi.org/10.1016/j.aap.2022.106712>
- Google. (n.d.). *Google Streetview* [Map]. Alphabet Inc. Retrieved June 19, 2024, from <https://www.google.com/maps>
- Graat, E., Midden, C., & Bockholts, P. (1999). Complex evacuation; effects of motivation level and slope of stairs on emergency egress time in a sports stadium. *Safety Science*, *31*(2), 127–141. [https://doi.org/10.1016/S0925-7535\(98\)00061-7](https://doi.org/10.1016/S0925-7535(98)00061-7)
- Granié, M.-A. (2008, March 4). *Influence de l'adhésion aux stéréotypes de sexe sur la perception des comportements piétons chez l'adulte*. Recherche Transports Sécurité. <https://rts.revuesonline.com/article.jsp?articleId=12904>
- Granié, M.-A., Pannetier, M., & Guého, L. (2013). Developing a self-reporting method to measure pedestrian behaviors at all ages. *Accident Analysis & Prevention*, *50*, 830–839. <https://doi.org/10.1016/j.aap.2012.07.009>

- Greenberg, E. (n.d.). *1 A Physical Effort-Based Model for Pedestrian Movement in 2 Topographic Urban Environment*.
- Hankin, B. D., & Wright, R. A. (1958). Passenger Flow in Subways. *Journal of the Operational Research Society*, 9(2), 81–88. <https://doi.org/10.1057/jors.1958.9>
- Hart, P. E., Nilsson, N. J., & Raphael, B. (1968). A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2), 100–107. *IEEE Transactions on Systems Science and Cybernetics*.
<https://doi.org/10.1109/TSSC.1968.300136>
- Hatfield, J., & Murphy, S. (2007). The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections. *Accident Analysis & Prevention*, 39(1), 197–205. <https://doi.org/10.1016/j.aap.2006.07.001>
- Helbing, D., Keltsch, J., & Molnár, P. (1997). Modelling the evolution of human trail systems. *Nature*, 388(6637), Article 6637. <https://doi.org/10.1038/40353>
- Helbing, D., & Molnar, P. (1995). Social Force Model for Pedestrian Dynamics. *Physical Review E*, 51(5), 4282–4286. <https://doi.org/10.1103/PhysRevE.51.4282>
- Helbing, D., Schweitzer, F., Keltsch, J., & Molnár, P. (1997). Active walker model for the formation of human and animal trail systems. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 56(3), 2527–2539.
<https://doi.org/10.1103/PhysRevE.56.2527>
- Henderson, L. F. (1971). The Statistics of Crowd Fluids. *Nature*, 229(5284), Article 5284.
<https://doi.org/10.1038/229381a0>

- Highway Capacity Manual 6th Edition: A Guide for Multimodal Mobility Analysis* (with Transportation Research Board & National Academies of Sciences, Engineering, and Medicine). (2016). National Academies Press. <https://doi.org/10.17226/24798>
- Highway Capacity Manual 7th Edition: A Guide for Multimodal Mobility Analysis* (with Transportation Research Board & National Academies of Sciences, Engineering, and Medicine). (2022). National Academies Press. <https://doi.org/10.17226/26432>
- Highway Safety Code (1986). <https://www.legisquebec.gouv.qc.ca/en/document/cs/c-24.2>
- Highway Traffic Act, R.S.O. 1990, c. H.8 (1990). <https://www.ontario.ca/laws/view>
- Hoogendoorn, S., & Bovy, P. H. L. (2000). Gas-Kinetic Modeling and Simulation of Pedestrian Flows. *Transportation Research Record*, 1710(1), 28–36. <https://doi.org/10.3141/1710-04>
- Hoogendoorn, S. P., & Bovy, P. H. L. (2004). Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2), 169–190. [https://doi.org/10.1016/S0191-2615\(03\)00007-9](https://doi.org/10.1016/S0191-2615(03)00007-9)
- Johnson, S. (2023, April 22). *What desire paths can tell us about how to design safer, better public spaces*. <https://www.abc.net.au/news/2023-04-23/desire-paths-offer-insight-into-use-of-public-spaces/102191044>
- Katsuki, F., & Constantinidis, C. (2014). Bottom-up and top-down attention: Different processes and overlapping neural systems. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, 20(5), 509–521. <https://doi.org/10.1177/1073858413514136>

- Kretz, T. (2012). The Effect of Integrating Travel Time. In *arXiv e-prints*.
<https://doi.org/10.48550/arXiv.1204.5100>
- Kretz, T., Hengst, S., & Vortisch, P. (2008). *Pedestrian Flow at Bottlenecks—Validation and Calibration of Vissim’s Social Force Model of Pedestrian Traffic and its Empirical Foundations* (arXiv:0805.1788). arXiv. <https://doi.org/10.48550/arXiv.0805.1788>
- Kretz, T., Lohmiller, J., & Sukennik, P. (2017). *Some Indications on how to Calibrate the Social Force Model of Pedestrian Dynamics* (arXiv:1801.00276). arXiv.
<https://doi.org/10.48550/arXiv.1801.00276>
- Kudinov, S., Smirnov, E., Malyshev, G., & Khodnenko, I. (2018). *Planning Optimal Path Networks Using Dynamic Behavioral Modeling* (91). Article 91.
<https://doi.org/10.29007/ptfv>
- Lawton, R., Parker, D., Manstead, A. S. R., & Stradling, S. G. (1997). The Role of Affect in Predicting Social Behaviors: The Case of Road Traffic Violations. *Journal of Applied Social Psychology*, 27(14), 1258–1276. <https://doi.org/10.1111/j.1559-1816.1997.tb01805.x>
- Ma, L., Brandt, S. A., Seipel, S., & Ma, D. (2024). Simple agents – complex emergent path systems: Agent-based modelling of pedestrian movement. *Environment and Planning B: Urban Analytics and City Science*, 51(2), 479–495.
<https://doi.org/10.1177/23998083231184884>
- Meir, A., Tapiro, H., & Oron-Gilad, T. (2023). Towards safer, more walkable urban environments for child-pedestrians—Application of the theory of Planned behavior. *Safety Science*, 164. <https://doi.org/10.1016/j.ssci.2023.106148>

- Ministry of Transportation. (2016a). *Pedestrian Crossing Treatments*. Ministry of Transportation.
- Ministry of Transportation. (2016b, June). *Ontario Traffic Manual (OTM)—Book 15—Pedestrian crossing treatments*.
<https://www.library.mto.gov.on.ca/SydneyPLUS/Sydney/Portal/default.aspx?component=AAAAY&record=fa5caef1-9963-4786-b3c9-4b5e50e70321>
- Ministry of Transportation and Infrastructure. (2022, January 21). *Active Transportation Design Guide—Province of British Columbia*. Ministry of Transportation and Infrastructure.
<https://www2.gov.bc.ca/gov/content/transportation/transportation-infrastructure/engineering-standards-guidelines/traffic-engineering-safety/active-transportation-design-guide>
- Moyano Díaz, E. (1997). Teoría del comportamiento planificado e intención de infringir normas de transito en peatones. *Estudos de Psicologia (Natal)*, 2, 335–348.
<https://doi.org/10.1590/S1413-294X1997000200008>
- Nigel, G., & Klaus, T. (2005). *Simulation For The Social Scientist*. McGraw-Hill Education (UK).
- Nordfjærn, T., & Zavareh, M. F. (2016). Individualism, collectivism and pedestrian safety: A comparative study of young adults from Iran and Pakistan. *Safety Science*, 87, 8–17.
<https://doi.org/10.1016/j.ssci.2016.03.005>
- O’Hern, S., Stephens, A. N., Estgfaeller, N., Moore, V., & Koppel, S. (2020). Self-reported pedestrian behaviour in Australia. *Transportation Research Part F: Traffic Psychology and Behaviour*, 75, 134–144. <https://doi.org/10.1016/j.trf.2020.10.002>

- Özkan, T., & Lajunen, T. (2005). A new addition to DBQ: Positive Driver Behaviours Scale. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(4), 355–368.
<https://doi.org/10.1016/j.trf.2005.04.018>
- Papadimitriou, E., Yannis, G., & Golias, J. (2009). A critical assessment of pedestrian behaviour models. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(3), 242–255. <https://doi.org/10.1016/j.trf.2008.12.004>
- Porter, J., Anand, T., Johnson, B., Khan, R. M., & Sobel, N. (2005). Brain mechanisms for extracting spatial information from smell. *Neuron*, 47(4), 581–592.
<https://doi.org/10.1016/j.neuron.2005.06.028>
- Public Highways Act (1989).
<https://nslegislature.ca/sites/default/files/legc/statutes/public%20highways.pdf>
- Qu, W., Zhang, H., Zhao, W., Zhang, K., & Ge, Y. (2016). The effect of cognitive errors, mindfulness and personality traits on pedestrian behavior in a Chinese sample. *Transportation Research Part F: Traffic Psychology and Behaviour*, 41, 29–37.
<https://doi.org/10.1016/j.trf.2016.06.009>
- Rakotoarivelo, A. N., Auberlet, J.-M., & Brémond, R. (2023). Will I cross when I see you crossing? Towards illegal pedestrian road crossing simulations. *Transportation Research Part F: Traffic Psychology and Behaviour*, 93, 309–325.
<https://doi.org/10.1016/j.trf.2023.01.022>
- Rasmussen, J. (1980). What can be Learned from Human Error Reports?: Changes in Working Life. *Proceedings of an International Conference on Changes in the Nature and Quality of Working Life*, 97–113.

- Reason, J., Manstead, A., Stephen, S., Baxter, J., & Campbell, K. (1990). Errors and violations on the roads: A real distinction? *Ergonomics*, *33*(10–11), 1315–1332.
<https://doi.org/10.1080/00140139008925335>
- Region of Waterloo & IBI Group. (2019, June). *Moving Forward – 2018 Transportation Master Plan*. https://www.regionofwaterloo.ca/en/living-here/resources/Transportation-Master-Plan/DOCS_ADMIN-3030800-v3-TMP_Report_Moving_Forward_Main_Report_FINAL_2019-06-12.pdf
- Richards, B., & Ingold, T. (2022). Forming Paths Within Post-Industrial Landscapes. In D. Svensson, K. Saltzman, & S. Sörlin (Eds.), *Pathways* (pp. 316–331). White Horse Press.
<https://www.jstor.org/stable/j.ctv2p5zn1t.23>
- Sahaleh, S., Bierlaire, M., Farooq, B., Danalet, A., & Hänseler, F. (2012). *Scenario Analysis of Pedestrian Flow in Public Spaces*.
- Savary, P., Foltête, J. C., & Garnier, S. (2022). Cost distances and least cost paths respond differently to cost scenario variations: A sensitivity analysis of ecological connectivity modeling. *International Journal of Geographical Information Science*, *36*(8), 1652–1676.
<https://doi.org/10.1080/13658816.2021.2014852>
- Seyfried, A., Rupprecht, T., Passon, O., Steffen, B., Klingsch, W., & Boltes, M. (2007). *New insights into pedestrian flow through bottlenecks* (arXiv:physics/0702004). arXiv.
<https://doi.org/10.48550/arXiv.physics/0702004>
- Shirk, A. J., Wallin, D. O., Cushman, S. A., Rice, C. G., & Warheit, K. I. (2010). Inferring landscape effects on gene flow: A new model selection framework. *Molecular Ecology*, *19*(17), 3603–3619. <https://doi.org/10.1111/j.1365-294X.2010.04745.x>

- Smirnov, E., Dunaenko, S., & Kudinov, S. (2020). Using multi-agent simulation to predict natural crossing points for pedestrians and choose locations for mid-block crosswalks. *Geo-Spatial Information Science*, 23(4), 362–374.
<https://doi.org/10.1080/10095020.2020.1847003>
- Solmazer, G., Azık, D., Fındık, G., Üzümcüoğlu, Y., Ersan, Ö., Kaçan, B., Özkan, T., Lajunen, T., Öz, B., Pashkevich, A., Pashkevich, M., Danelli-Mylona, V., Georgogianni, D., Berisha Krasniqi, E., Krasniqi, M., Makris, E., Shubenkova, K., & Xheladini, G. (2020). Cross-cultural differences in pedestrian behaviors in relation to values: A comparison of five countries. *Accident Analysis & Prevention*, 138, 105459.
<https://doi.org/10.1016/j.aap.2020.105459>
- Stea, D. (Ed.). (1973). *Image and Environment: Cognitive Mapping and Spatial Behavior*. Routledge. <https://doi.org/10.4324/9780203789155>
- Thorhauge, M., Haustein, S., & Cherchi, E. (2016). Accounting for the Theory of Planned Behaviour in departure time choice. *Transportation Research Part F: Traffic Psychology and Behaviour*, 38, 94–105. <https://doi.org/10.1016/j.trf.2016.01.009>
- Tian, W., Song, W., Ma, J., Fang, Z., Seyfried, A., & Liddle, J. (2012). Experimental study of pedestrian behaviors in a corridor based on digital image processing. *Fire Safety Journal*, 47, 8–15. <https://doi.org/10.1016/j.firesaf.2011.09.005>
- Timmermans, H., van der Hagen, X., & Borgers, A. (1992). Transportation systems, retail environments and pedestrian trip chaining behaviour: Modelling issues and applications. *Transportation Research Part B: Methodological*, 26(1), 45–59.
[https://doi.org/10.1016/0191-2615\(92\)90019-S](https://doi.org/10.1016/0191-2615(92)90019-S)

- Todorovic, N., & Pedersen, B. (2020, August 31). What You Should Know About Jaywalking. *McLeish Orlando LLP*. https://www.mcleishorlando.com/insights/what-you-should-know-about-jaywalking/#_ftn2
- Tong, Y., & Bode, N. W. F. (2022). The principles of pedestrian route choice. *Journal of The Royal Society Interface*, *19*(189), 20220061. <https://doi.org/10.1098/rsif.2022.0061>
- Traffic Safety Act, T-18.1, Pub. L. No. T-18.1 (2006).
<https://publications.saskatchewan.ca/#/products/12208>
- Treiber, M., & Kesting, A. (2013). *Traffic Flow Dynamics: Data, Models and Simulation*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-32460-4>
- Vandroux, R., Granié, M.-A., Jay, M., Sueur, C., & Pelé, M. (2022). The pedestrian behaviour scale: A systematic review of its validation around the world. *Accident Analysis & Prevention*, *165*, 106509. <https://doi.org/10.1016/j.aap.2021.106509>
- Vanumu, L. D., Ramachandra Rao, K., & Tiwari, G. (2017). Fundamental diagrams of pedestrian flow characteristics: A review. *European Transport Research Review*, *9*(4), Article 4. <https://doi.org/10.1007/s12544-017-0264-6>
- Warner, H. W., & Åberg, L. (2006). Drivers' decision to speed: A study inspired by the theory of planned behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, *9*(6), 427–433. <https://doi.org/10.1016/j.trf.2006.03.004>
- Wightman, F. L., & Kistler, D. J. (1990). *Hearing in three dimensions: Sound localization* (NAS 1.26:187706). <https://ntrs.nasa.gov/citations/19910004658>
- Wilson, R. J. (2009). *Introduction to graph theory* (4. ed., [Nachdr.]). Prentice Hall.

- Xiao, Y., Liu, Y., & Liang, Z. (2021). Study on Road-Crossing Violations among Young Pedestrians Based on the Theory of Planned Behavior. *Journal of Advanced Transportation*, 2021, 1–11. <https://doi.org/10.1155/2021/6893816>
- Xu, J., Liu, J., Qu, W., Ge, Y., Sun, X., & Zhang, K. (2018). Comparison of pedestrian behaviors between drivers and non-drivers in Chinese sample. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58, 1053–1060. <https://doi.org/10.1016/j.trf.2018.05.034>
- Yang, L., Liu, S., Li, J., Zhu, K., & Fang, T. (2009). Information-based evacuation experiment and its cellular automaton simulation. *International Journal of Modern Physics C*, 20(10), 1583–1596. <https://doi.org/10.1142/S0129183109014618>
- Yang, & Sun, J. (2013). Study on Pedestrian Red-Time Crossing Behavior: Integrated Field Observation and Questionnaire Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2393(1), 117–124. <https://doi.org/10.3141/2393-13>
- Zhang, J., Klingsch, W., Schadschneider, A., & Seyfried, A. (2011). Transitions in pedestrian fundamental diagrams of straight corridors and T-junctions. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(06), P06004. <https://doi.org/10.1088/1742-5468/2011/06/P06004>
- Zhang, J., & Seyfried, A. (2013). Empirical Characteristics of Different Types of Pedestrian Streams. *Procedia Engineering*, 62, 655–662. <https://doi.org/10.1016/j.proeng.2013.08.111>

- Zhao, Q., Zhuang, X., Zhang, T., He, Y., & Ma, G. (2023). Pedestrian gaze pattern before crossing road in a naturalistic traffic setting. *European Transport Research Review*, *15*(1), 31. <https://doi.org/10.1186/s12544-023-00605-1>
- Zhao, X., Li, X., Rakotonirainy, A., Bourgeois-Bougrine, S., & Delhomme, P. (2022). Predicting pedestrians' intention to cross the road in front of automated vehicles in risky situations. *Transportation Research Part F: Traffic Psychology and Behaviour*, *90*, 524–536. <https://doi.org/10.1016/j.trf.2022.05.022>
- Zhou, H., Romero, S. B., & Qin, X. (2016). An extension of the theory of planned behavior to predict pedestrians' violating crossing behavior using structural equation modeling. *Accident Analysis and Prevention*, *95*, 417–424. <https://doi.org/10.1016/j.aap.2015.09.009>

Appendix A – Python Code

```
# %%
#import
import numpy as np
import pandas as pd
from plotnine import *
import matplotlib.pyplot as plt
import matplotlib.animation as animation
plt.rcParams['toolbar'] = 'toolmanager'
from matplotlib.backend_tools import ToolBase, ToolToggleBase
import ipywidgets as widgets
import seaborn as sns
import pickle
from itertools import product
from itertools import compress

from numpy.random import default_rng
rng = default_rng()
import scipy

import csv
import os
import sys
from pathlib import Path
import json

from IPython.display import clear_output

import time

from math import floor

# %%
#line necessary to open figures in Matplotlib window

# %%

#####
# Function to read the surface lict csv into the code as a dictionary
# Inputs: File name of surface dict file
```

```

# returns: surface list in form of a dictionary
#####

def ImportSurfaceTypes (surfaceFile):

    file = open(os.path.join(sys.path[0], surfaceFile), "r")
    df = pd.read_csv(file, header=0, index_col=0)
    surfaceDict = df.to_dict('index')

    return surfaceDict

# %%
#####
# Function to read the map csv files into the code as lists
# Inputs: File name of map file
# returns: map in form of a 2D list
#####
def ImportMap (mapFile):
    file = open(os.path.join(sys.path[0], mapFile), "r")
    map = list(csv.reader(file, delimiter=","))
    file.close()
    return map

# %%
#####
# Function to import vehicle trajectory data into code from VISSIM
# Inputs: File name of fzp export file from VISSIM
# returns: dataframe with vehicle trajectories
#####
def ConvertFZP (vissimFile):

# -*- coding: utf-8 -*-
# Created on Mon Feb 28 10:33:07 2022
# This script export a xlsx file from the .fzp
# VISSIM's direct output
# @author: agustinguerra

    # seek for fzp file and transfor to text
    file_name = vissimFile
    p = Path(file_name)
    # save text if no text file in directory

```

```

if os.path.exists(file_name):
    p.rename(p.with_suffix('.txt'))

# open text file get right lines to read from
file = open(p.with_suffix('.txt'), "r")
lines_list = file.readlines()[19:]
file.close()

# write to txt file the data
textfile = open("VISSIM_Results.txt", "w")
for element in lines_list:
    textfile.write(element)
textfile.close()

# read data frame and save it to notebook directory
df_VehicleTrajectories = pd.read_csv("VISSIM_Results.txt", delimiter=';', )

# translate positions along links to positions in xy space. i.e. in link 1, x
position is the position along the link - 23 metres

### Conditions for Phillip St Midblock Case ###

conditions_x = [
    (df_VehicleTrajectories[r'LANE\LINK\NO']==1), #note: must use r to avoid
unicode error in python
    (df_VehicleTrajectories[r'LANE\LINK\NO']==2)

]
values_x = [8,16]

conditions_y = [
    (df_VehicleTrajectories[r'LANE\LINK\NO']==1),
    (df_VehicleTrajectories[r'LANE\LINK\NO']==2)

]
values_y = [df_VehicleTrajectories['POS']-16,142-
df_VehicleTrajectories['POS']]

### Conditions for University & Phillip Case ###

# conditions_x = [

```

```

        # (df_VehicleTrajectories[r'LANE\LINK\NO']==1), #EBT #note: must use r to
avoid unicode error in python
        # (df_VehicleTrajectories[r'LANE\LINK\NO']==2), #WBT
        # (df_VehicleTrajectories[r'LANE\LINK\NO']==3), #SBT
        # (df_VehicleTrajectories[r'LANE\LINK\NO']==4), #NBT
        # (df_VehicleTrajectories[r'LANE\LINK\NO']==5), #EBL
        # (df_VehicleTrajectories[r'LANE\LINK\NO']==6), #WBL
        # (df_VehicleTrajectories[r'LANE\LINK\NO']==7), #SBL
        # (df_VehicleTrajectories[r'LANE\LINK\NO']==8) #NBL

# ]          #link 1          #link
2          #link3,4          #link 5          #link
6          #link 7,8
#     values_x = [df_VehicleTrajectories['POS']-23,128-
df_VehicleTrajectories['POS'],36,44,df_VehicleTrajectories['POS'],76-
df_VehicleTrajectories['POS'],40,40]

#     conditions_y = [
#         (df_VehicleTrajectories[r'LANE\LINK\NO']==1), #EBT #note: must use r
to avoid unicode error in python
#         (df_VehicleTrajectories[r'LANE\LINK\NO']==2), #WBT
#         (df_VehicleTrajectories[r'LANE\LINK\NO']==3), #SBT
#         (df_VehicleTrajectories[r'LANE\LINK\NO']==4), #NBT
#         (df_VehicleTrajectories[r'LANE\LINK\NO']==5), #EBL
#         (df_VehicleTrajectories[r'LANE\LINK\NO']==6), #WBL
#         (df_VehicleTrajectories[r'LANE\LINK\NO']==7), #SBL
#         (df_VehicleTrajectories[r'LANE\LINK\NO']==8) #NBL

# ]          #link 1,2          #link 3          #link
4          #link 5,6          #link 7          #link 8
#     values_y = [71,59,df_VehicleTrajectories['POS']+5,123-
df_VehicleTrajectories['POS'],65,65,df_VehicleTrajectories['POS']+28,86-
df_VehicleTrajectories['POS']]

df_VehicleTrajectories['X'] = np.select(conditions_x,values_x)
df_VehicleTrajectories['Y'] = np.select(conditions_y,values_y)

return df_VehicleTrajectories

# %%

#####

```

```

# Function to organize vehicle trajectory dataframe by link
# Inputs: link map and vehicle trajectory dataframe
# returns: dictionary where the key is the link no. and the value is the
dataframe for that link
#####
def createVehicleLinkList(linkMap, df_VehicleTrajectories):
    linkList = []
    for linkRow in linkMap: #loop through all cells to find list of links
        for linkTile in linkRow:
            links = linkTile.split(";")
            for link in links:
                linkList.append(link)
    linkSetUnique = list(set(linkList)) #filter to unique values
    linkSetUnique = [x for x in linkSetUnique if x != ''] #remove blank values

    vehicleLinkDict = {}
    for link in linkSetUnique: #sort data by link

        df_LinkVehicles =
df_VehicleTrajectories[df_VehicleTrajectories[r'LANE\LINK\NO']==int(link)]
        vehicleLinkDict[link] = df_LinkVehicles

    return vehicleLinkDict

# %%
#####
# Function to import signal indication data into code from VISSIM
# Inputs: File name of lsa export file from VISSIM
# returns: dictionary with signal indications (dictionary key is time)
# Original author: agustinguerra, modified by Chris Lehmann Skelton
#####
def ConvertLSA (vissimFile):

    if os.path.exists('convert.txt'):
        with open('convert.txt') as f:
            data = f.read()
            signalsAtTimes = json.loads(data)
            print('data loaded')
    else:
        # seek for fzp file and transfor to text
        file_name = vissimFile
        p = Path(file_name)
        # save text if no text file in directory
        if os.path.exists(file_name):

```

```

        p.rename(p.with_suffix('.txt'))

# open text file get right lines to read from
file = open(p.with_suffix('.txt'), "r")
lines_list = file.readlines()[18:]
file.close()

# write to txt file the data
textfile = open("VISSIM_Signal.txt", "w")
for element in lines_list:
    textfile.write(element)
textfile.close()

# read data frame and assign headings to each column
df_SignalTimings = pd.read_csv("VISSIM_Signal.txt", delimiter=';',
header=None, names=['Time', 'Unknown', 'Signal Controller', 'Signal Group',
'Indication', 'Duration', 'Controller Type', 'Unknown2','Extra'] )

#convert signal changes into continuous signal status list

listofTimes = df_SignalTimings['Time'].tolist() #list of times in signal
file

times = list(np.round(np.arange(0, listofTimes[-1], .1),1)) #create list
of every 0.1s time increment from 0 to last time in signal file

signalStatusDict ={'0': 'green', '2':'green',
                    '4':'red',
                    '6':'green',
                    '8':'red'
                    }

signalsAtTimes = {}

#lsa file only includes signal changes, so we need to loop through the
data at each time to create a record of what the indication was at each time step

for t in times:

    df_signalsatcurrenttime =
df_SignalTimings[df_SignalTimings['Time']==t]

    if df_signalsatcurrenttime.empty:
        pass
    else:

```

```

        SGstoUpdate = df_signalsatcurrenttime['Signal Group'].tolist()
        IndicationUpdates =
df_signalsatcurrenttime['Indication'].tolist()
        signalStatusDict.update(zip(SGstoUpdate, IndicationUpdates))

        signalsAtTimes[str(t)] = pickle.loads(pickle.dumps(signalStatusDict,
-1))

        with open('convert.txt', 'w') as convert_file:
            convert_file.write(json.dumps(signalsAtTimes))

    return signalsAtTimes

# %%
# %%
import heapq
import random
import time

#####
# Node class which represents each individual node object in our graph
# Inputs: Node position (x and y), distance to goal aka heuristic value (h), cost
to move to this node from neighbouring nodes
# returns: dictionary with signal indications (dictionary key is time)
#####

class Node:
    def __init__(self, t, x, y, h, cost, parent=None):
        self.t = t
        self.x = x
        self.y = y
        self.neighbours = []
        self.cost = cost
        self.g = 0 #initially g cost is 0 but this will be updated during
pathfinding
        self.h = h
        self.f = self.g + self.h
        self.parent = parent #parent node (the preceding node along path,
initially set to none but updated during pathfinding)

    def __lt__(self, other):
        return self.f < other.f

#####
# Pathfinding function

```

```

# Inputs: start node, goal node(s) Note: since pedestrians can arrive at any
time, there are a list of goal nodes.
# returns: path as list of tuples
#####

def astar_with_variable_costs(start, goal_list):
    start.g = 0 #set start node g cost to zero

    open_list = [start] #add start node to open list
    closed_list = []

    while open_list:

        current = heapq.heappop(open_list) #remove current node to

        if current in goal_list: #if the current node is in the goal list, we
have found a path!, so loop through all parents to compile the entire path
            path = []
            while current.parent!=None and
current.parent!=current:
                path.append((current.t,current.x,current.y))
                current = current.parent
            path.append((0,start.x,start.y))
            path.reverse()

            return path

        for neighbour in current.neighbours: #investigate neighbours of current
node
            if ((neighbour.x - current.x)**2+(neighbour.y - current.y)**2)**0.5 >
1.0:
                g = current.g + (2**0.5)*neighbour.cost
            else:
                g = current.g + neighbour.cost
            f = g + neighbour.h
            if f < neighbour.f or neighbour.parent == None:
                neighbour.g = g
                neighbour.f = f
                neighbour.parent = current

            if ((neighbour not in open_list) and (neighbour not in closed_list)):
                heapq.heappush(open_list, neighbour)

        heapq.heappush(closed_list, current)

```

```

#####
# Function to construct the graph that will be used for pathfinding
# Inputs: pedestrian object
# returns: 3D list of node objects
#####

def build_adjacency_list(ped):
    width = len(ped.liveMap[0][0])
    height = len(ped.liveMap[0])
    duration = len(ped.liveMap)

    baseMap_array = np.array(ped.baseMap)
    liveMap_array = np.array(ped.liveMap)
    averageMap_array = np.array([np.array(ped.averageMap)]*duration)
    linkMap_array = np.array(ped.linkMap)

    baseMap_indexes = np.transpose(np.nonzero(baseMap_array))

    Distance_arrayvalues = scipy.spatial.distance.cdist(baseMap_indexes,
np.array([[ped.position_y,ped.position_x]]))

    Distance_array = Distance_arrayvalues.reshape(height,width) #array of same
dimesnions as baseMap which has the distance of each cell relative to the current
ped position

    GoalDistance_arrayvalues = scipy.spatial.distance.cdist(baseMap_indexes,
np.array([[ped.destination_y,ped.destination_x]]))

    GoalDistance_list = GoalDistance_arrayvalues.reshape(height,width).tolist()
#array of same dimesnions as baseMap which has the distance of each cell relative
to the goal position

    def liveCosts(ped, distance, links): #determines if this tile should be
calculated using live costs (1 if true)
        link = links.split(';')
        if distance <= ped.lookAheadDistance:
            return 1
        if any(check in link for check in ped.listofLinksInPerceptionRange):
            return 1
        else: return 0

    liveCostFunction = np.vectorize(liveCosts)

```

```

liveCost_array =
np.array([liveCostFunction(ped,Distance_array,linkMap_array)]*duration) #array of
cells where we are using live costs based on distance
averageCost_array = np.where(liveCost_array>0,0,1) #use average costs at
tiles we aren't using live costs

Time_array = np.zeros((duration,height,width))
Time_array[0:ped.lookAheadTimeSteps]=1
Time_liveCost_array = Time_array
Time_averageCost_array = np.where(liveCost_array>0,0,1) #use average costs at
times we aren't using live costs

perceivedMap_array =
Time_averageCost_array*averageCost_array*averageMap_array +
Time_liveCost_array*liveCost_array*liveMap_array #the final perceived map array
whoch is a mix of live and average costs

ped.perceivedMap = perceivedMap_array.tolist()

theMatrixofNodes = []

t1 = time.time()
for t in range(min(ped.lookAheadTimeSteps,duration)):
    yrow = []
    for y in range(height): #create nodes and assign cost values
        xrow = []
        for x in range(width):
            node = Node(t,x,y,
GoalDistance_list[y][x],ped.perceivedMap[t][y][x])
            xrow.append(node)
        yrow.append(xrow)
    theMatrixofNodes.append(yrow)

t2 = time.time()
print (t2-t1)

t1 = time.time()
for t in range(min(ped.lookAheadTimeSteps,duration)):
    t_tindex = min(t + 1,ped.lookAheadTimeSteps-1,duration-1)
    for y in range(height): #loop through list of nodes and create list of
neighbours for each node
        for x in range(width):
            theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[t_tindex][y][x])
            if 0 < x:

```

```

        theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[
t_tindex][y][x-1])
        if x < width-1:
            theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[
t_tindex][y][x+1])
        if 0 < y:
            theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[
t_tindex][y-1][x])
        if y < height-1:
            theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[
t_tindex][y+1][x])
        if 0 < y and 0 < x:
            theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[
t_tindex][y-1][x-1])
        if 0 < y and x < width-1:
            theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[
t_tindex][y-1][x+1])
        if y < height-1 and 0 < x:
            theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[
t_tindex][y+1][x-1])
        if y < height-1 and x < width-1:
            theMatrixofNodes[t][y][x].neighbours.append(theMatrixofNodes[
t_tindex][y+1][x+1])
    t2 = time.time()
    print (t2-t1)

    return theMatrixofNodes

# %%

#####
# Class which represents an individual pedestrian
# Inputs: maps, surface dictionary, start time when pedestrian enters simulation
#####
class Pedestrian():
    def __init__(self,baseMap, linkMap, signalPhaseMap,surfaceDict, startTime):

        self.baseMap = baseMap
        self.linkMap = linkMap
        self.signalPhaseMap = signalPhaseMap
        self.surfaceDict = surfaceDict

        self.startTime = startTime

        rand_y = rng.choice([-2,-1,0,1,2])

```

```

rand_x = rng.choice([-2,-1,0,1,2])
self.origin = (22 + rand_x,20 + rand_y)

rand_y = rng.choice([-2,-1,0,1,2])
rand_x = rng.choice([-2,-1,0,1,2])
self.destination = (2 + rand_x,77 + rand_y)

self.origin_x = self.origin[0]
self.origin_y = self.origin[1]

self.destination_x = self.destination[0]
self.destination_y = self.destination[1]

self.position_x = self.origin_x
self.position_y = self.origin_y

self.plannedPath = [] #the path the pedestrian is planning to take
self.actualPath = [] #the path the pedestrian actually takes
self.active_status = True

# perception variables
self.walkSpeed = rng.normal(loc = 1.34, scale = 0.37)
road_width = 14 #metres
self.lookAheadTimeSteps = int(round(road_width/self.walkSpeed,0))
#int(round(14/self.walkSpeed,0))
self.lookAheadDistance = 5 #d_A in model

# cost variables
self.AlphaBase = rng.triangular(0,0.5,1)
self.AlphaLegalViolation = rng.triangular(0,0.5,1)
self.desiredGap = road_width/self.walkSpeed

self.liveMap = []
self.averageMap = []
self.perceivedMap = []
self.listofLinksInPerceptionRange = {}

self.simulationResults = {}

#####
# Function to find links pedestrian can perceive, construct the graph, and
find path
# other functions are called to do this, this function essentially organizes
that process

```

```

# Inputs: starting position for pathfinding
#####
def FindPath (self, start_x, start_y):

    print("finding links in range")
    self.FindLinksInPerceptionRange() #find links in range of pedestrian
perception
    print("building adjacency list")
    theMatrixofNodes = build_adjacency_list(self) #construct graph

    start_node = theMatrixofNodes[0][start_y][start_x] #starting node object
    goal_node_list = [i[self.destination_y][self.destination_x] for i in
theMatrixofNodes] #ending node is list of node x,y and all time steps
    print("finding path")
    path = astar_with_variable_costs(start_node, goal_node_list) #find path

    self.plannedPath = path #update planned path

#####
# Function to find all the links that are within the perception range of the
pedestrian
# Inputs:
#####

def FindLinksInPerceptionRange (self):
    linksInRange = np.array(self.linkMap)

    listofLinks = linksInRange[self.position_y-
self.lookAheadDistance:self.position_y+self.lookAheadDistance, self.position_x-
self.lookAheadDistance:self.position_x+self.lookAheadDistance].flatten().tolist()
    listofLinks = [i.split(';') for i in listofLinks if i!='0']

    self.listofLinksInPerceptionRange = []
    for i in listofLinks:
        for j in i:
            self.listofLinksInPerceptionRange.append(j)

    self.listofLinksInPerceptionRange =
list(set(self.listofLinksInPerceptionRange))

    if ('1' in self.listofLinksInPerceptionRange) or ('2' in
self.listofLinksInPerceptionRange): #if the pedestrian can see one direction
(link) of a road, they can also see the other direction
        self.listofLinksInPerceptionRange.append('1')
        self.listofLinksInPerceptionRange.append('2')

```

```

        self.listofLinksInPerceptionRange.append('5')
        self.listofLinksInPerceptionRange.append('6')
        if ('3' in self.listofLinksInPerceptionRange) or ('4' in
self.listofLinksInPerceptionRange):
            self.listofLinksInPerceptionRange.append('3')
            self.listofLinksInPerceptionRange.append('4')
            self.listofLinksInPerceptionRange.append('7')
            self.listofLinksInPerceptionRange.append('8')

#####
# Function to find links pedestrian can perceive, construct the graph, and
find path
# other functions are called to do this, this function essentially organizes
that process
# Inputs: vehicle and signal information
# Returns: 3D list with live costs
#####

def CreateLiveMap (self,
vehicleLinkDict_X,vehicleLinkDict_Y,vehicleLinkDict_Speed,signalStatusDict):
#interpret the basemap information based on the pedestrians own preferences

    width = len(self.baseMap[0]) # Get the length and width of the imported
grid
    height = len(self.baseMap)

    Cpave = 1
    Crisk = 141
    vehicleLength = 7
    d_viol = 30 #violation distance beyond which

    dNearCrosswalk = 3 #distance from crosswalk that reduced risk area
extends
    pYieldCW = 1 #perceived probability of vehicle yielding at CW
    pYieldMidblock = 0 #perceived probability of vehicle yielding midblock

    baseMap_array = np.array(self.baseMap)
    linkMap_array = np.array(self.linkMap)
    signalStatus_array = np.array(self.signalPhaseMap)
    redPhases = [k for k, v in signalStatusDict.items() if 'green' not in v]

```

```

    redSignal_array =
np.where(np.isin(signalStatus_array,redPhases),True,False) #value of 1 if not
green, 0 if green
    #greenPhases = np.where(redSignal_array==1,0,1)

    redLinks = linkMap_array[redSignal_array]
    redLinksstring = ';' .join(np.unique(redLinks[redLinks != 0]))

    CW_array = np.where(baseMap_array=='10',1,0)
    PVMA_array = np.where(baseMap_array=='42',1,0)
    CWandPVMA_array = CW_array+PVMA_array #1 if PVMA or CW surface, zero
otherwise

    ##### base costs #####
    Cbase_dict = {}
    for key in self.surfaceDict:
        Cbase_dict[str(key)]=self.surfaceDict[key]['Cbase']

    Cbase_array = np.vectorize(Cbase_dict.__getitem__)(baseMap_array)
    baseCosts_array = self.AlphaBase*(Cbase_array-Cpave)+Cpave #calculate
base cst for ped
    ##### end base costs #####

    ##### violation costs #####

    CWindex_indexes = np.transpose(np.nonzero(CW_array))
    listOfCWs = CWindex_indexes.tolist()
    baseMap_indexes = np.transpose(np.nonzero(baseMap_array))

    CWdistance_arrayvalues = scipy.spatial.distance.cdist(baseMap_indexes,
CWindex_indexes).min(axis=1)
    CWdistance_array = CWdistance_arrayvalues.reshape(height,width)

    def violationCosts(d,link, d_viol, redLinksstring):
        #dviol parameters
        left = 0
        peak = d_viol/2 #peak

        Clegalviolation_calibrated = 9.5 #value calibrated in previous work
        Clegalviolation_at_midblock = 0

        redLinks = redLinksstring.split(';')

        links = link.split(';')

```

```

        if not set(links).isdisjoint(redLinks) : Clegalviolation_at_CW = 0
        else: Clegalviolation_at_CW = 0

        Cpeak = max(Clegalviolation_calibrated,
Clegalviolation_at_CW,Clegalviolation_at_midblock) #the peak value is the max of
calibrated peak, midblock and crosswalk,

        #this prevents a case where cw is less appealing than
violating, which wouldn't make sense

        #Violation determination
        if d == 0:
            Clegalviolation = Clegalviolation_at_CW
        elif d >= peak and d < d_viol:
            Clegalviolation = (Clegalviolation_at_midblock-Cpeak)/(d_viol-
peak)*(d-peak) + Cpeak
        elif d >= left and d < peak:
            Clegalviolation = (Cpeak-Clegalviolation_at_CW)/(peak-left)*d +
Clegalviolation_at_CW
        else:
            Clegalviolation = Clegalviolation_at_midblock
        return Clegalviolation

        vectViolCostFunc = np.vectorize(violationCosts)
        Clegalviolation_array =
vectViolCostFunc(CWdistance_array,linkMap_array,d_viol,
redLinksstring)*CWandPVMA_array #for all road and CW tiles, find the
legalviolation cost
        violationCosts_array = Clegalviolation_array*self.AlphaLegalViolation
#calculate LV cost for pedestrian

        ##### end violation costs #####

        #ignore these commented out lines, part of a future effort to vectorize
the code
        #gamma
        #gammaRisk_array = CWdistance_array*PVMA_array
        #gammaRisk_array = np.where(gammaRisk_array<= dNearCrosswalk,(pYieldCW-
pYieldMidblock)*(gammaRisk_array)/dNearCrosswalk+(1-pYieldCW),pYieldCW)
        ##### agent costs #####
        agentMap = []
        for y in range(height): #interpreting the base map using information from
the surface dictionary
            row = []

```

```

for x in range(width):
    active_crosswalk = False
    GammaRisk = 0
    link = self.linkMap[y][x].split(';')

    if link[0] != '0':

        if [x,y] in listofCWs:

            if 'green' in
signalStatusDict[self.signalPhaseMap[y][x]]: #if traffic signal is green,
crosswalks have the same attributes as sidewalks, otherwise same as roads
                active_crosswalk = True
            else:
                active_crosswalk = False

            if CWdistance_array[y][x] <= dNearCrosswalk and (x,y) not in
listofCWs and 'green' in signalStatusDict[self.signalPhaseMap[y][x]]:
                GammaRisk = (pYieldCW-
pYieldMidblock)*(CWdistance_array[y][x])/dNearCrosswalk+(1-pYieldCW) #if lower
than alpha threshold, amount of risk experienced scales linearly with distance
            else:
                GammaRisk = pYieldCW

            closestVehicleTTC = 10000
            xDistance = 10000
            yDistance = 10000
            selectedLink = 0

            # this section is commented out for Phillip Midblock Example
            # if '1' in link and x > 42: #related links, allow for the
conflict zone to be projected beyond the link the vehicle is currently located on
            #     link.append('4')
            #     link.append('7')
            # if '2' in link and x < 39:
            #     link.append('3')
            #     link.append('8')
            # if '3' in link and y > 62:
            #     link.append('6')
            #     link.append('1')
            # if '4' in link and y < 59:
            #     link.append('2')
            #     link.append('5')

```

```

        for i in link:
            vehicleDistanceList = []
            for a, b in zip(vehicleLinkDict_X[i],
vehicleLinkDict_Y[i]):
                xDistance = x-a
                yDistance = y-b
                vehicleDistanceList.append(((xDistance)**2+(yDistance
)**2)**0.5)

            try:
                closestVehicleDistance = min(vehicleDistanceList)
                closestVehicleIndex=
vehicleDistanceList.index(min(vehicleDistanceList))
                VehicleSpeedList = vehicleLinkDict_Speed[i]
                closestVehicleTTC_check =
closestVehicleDistance/(VehicleSpeedList[closestVehicleIndex]/3.6)

                if closestVehicleTTC_check < closestVehicleTTC:
                    closestVehicleTTC = closestVehicleTTC_check
                    selectedLink = i

            except:
                pass

        if active_crosswalk == True:
            BetaRisk = 0

        ### if statements for Phillip Midblock ###
        elif selectedLink == '1' :
            if yDistance >= -vehicleLength and yDistance < 0: #if
area within vehicle length and located behind front of vehicle
                BetaRisk = 1
            elif closestVehicleTTC < self.desiredGap and yDistance >=
0: #if in danger zone and in front of vehicle
                BetaRisk = 1
            else:
                BetaRisk = 0
        elif selectedLink == '2' :
            if yDistance <= vehicleLength and yDistance > 0: #if area
within vehicle length and located behind front of vehicle
                BetaRisk = 1
            elif closestVehicleTTC < self.desiredGap and yDistance <=
0: #if in danger zone and in front of vehicle
                BetaRisk = 1
            else:

```

```

BetaRisk = 0

### if statements for University & Phillip ###

# elif selectedLink == '3' : #SBT
#     if yDistance >= -vehicleLength and yDistance < 0: #if
area within vehicle length and located behind front of vehicle
#         BetaRisk = 1
#     elif closestVehicleTTC < self.desiredGap and yDistance
>= 0: #if in danger zone and in front of vehicle
#         BetaRisk = 1
#     else:
#         BetaRisk = 0
# elif selectedLink == '4' : #NBT
#     if yDistance <= vehicleLength and yDistance > 0: #if
area within vehicle length and located behind front of vehicle
#         BetaRisk = 1
#     elif closestVehicleTTC < self.desiredGap and yDistance
<= 0: #if in danger zone and in front of vehicle
#         BetaRisk = 1
#     else:
#         BetaRisk = 0
# elif selectedLink == '2' : #WBT
#     if xDistance <= vehicleLength and xDistance > 0: #if
area within vehicle length and located behind front of vehicle
#         BetaRisk = 1
#     elif closestVehicleTTC < self.desiredGap and xDistance
<= 0: #if in danger zone and in front of vehicle
#         BetaRisk = 1
#     else:
#         BetaRisk = 0
# elif selectedLink == '1' : #EBT
#     if yDistance >= -vehicleLength and xDistance < 0: #if
area within vehicle length and located behind front of vehicle
#         BetaRisk = 1
#     elif closestVehicleTTC < self.desiredGap and xDistance
>= 0: #if in danger zone and in front of vehicle
#         BetaRisk = 1
#     else:
#         BetaRisk = 0
# elif selectedLink == '7' : #SBL
#     if yDistance >= -vehicleLength and yDistance < 0: #if
area within vehicle length and located behind front of vehicle
#         BetaRisk = 1

```

```

        #     elif closestVehicleTTC < self.desiredGap and yDistance
>= 0: #if in danger zone and in front of vehicle
        #         BetaRisk = 1
        #     else:
        #         BetaRisk = 0
        # elif selectedLink == '8' : #NBL
        #     if yDistance <= vehicleLength and yDistance > 0: #if
area within vehicle length and located behind front of vehicle
        #         BetaRisk = 1
        #     elif closestVehicleTTC < self.desiredGap and yDistance
<= 0: #if in danger zone and in front of vehicle
        #         BetaRisk = 1
        #     else:
        #         BetaRisk = 0
        # elif selectedLink == '6' : #WBL
        #     if xDistance <= vehicleLength and xDistance > 0: #if
area within vehicle length and located behind front of vehicle
        #         BetaRisk = 1
        #     elif closestVehicleTTC < self.desiredGap and xDistance
<= 0: #if in danger zone and in front of vehicle
        #         BetaRisk = 1
        #     else:
        #         BetaRisk = 0
        # elif selectedLink == '5' : #EBL
        #     if yDistance >= -vehicleLength and xDistance < 0: #if
area within vehicle length and located behind front of vehicle
        #         BetaRisk = 1
        #     elif closestVehicleTTC < self.desiredGap and xDistance
>= 0: #if in danger zone and in front of vehicle
        #         BetaRisk = 1
        #     else:
        #         BetaRisk = 0
        else:
            BetaRisk = 0
    else:
        BetaRisk = 0

    agentCosts = GammaRisk*BetaRisk*Crisk
    row.append(agentCosts)
    agentMap.append(row)

#### end agent costs ####

```

```

        interpretedMap_array = baseCosts_array + violationCosts_array +
np.array(agentMap)
        interpretedMap = interpretedMap_array.tolist()
        return interpretedMap

#####
# Function to update the pedestrian path, remove one time step from maps to
prepare for next time step
# Inputs: starting position for pathfinding
#####

def UpdatePosition(self):
    # Update pedestrian positions and remove if they have reached the
destination

    if (self.position_x == self.destination_x) and (self.position_y ==
self.destination_y): #if path complete, deactivate pedestrian
        self.active_status = False
        self.actualPath.append((self.position_x,self.position_y))
    else:
        self.actualPath.append((self.position_x,self.position_y))

        self.plannedPath.pop(0)
        self.liveMap.pop(0)
        self.position_x = self.plannedPath[0][1]
        self.position_y = self.plannedPath[0][2]
        self.perceivedMap.pop(0)
        self.FindPath(self.position_x,self.position_y)

# %%

# animations for gif
class SimulationAnimation():
    def __init__(self,ped,simulationSteps, t_conversion,df_VehicleTrajectories,
signalsAtTimes):

        grid_kws = {'width_ratios': (0.9, 0.05), 'wspace': 0.2}
        fig, (ax, cbar_ax) = plt.subplots(1, 2, gridspec_kw = grid_kws, figsize =
(10, 8))

        self.ani = animation.FuncAnimation(fig = fig, func = self.Animate, fargs
= [ped, simulationSteps, t_conversion, df_VehicleTrajectories, signalsAtTimes,ax,
cbar_ax],init_func=self.Initial, frames = simulationSteps, repeat = False,)
        #plt.show()

```

```

    def Animate(self, t, ped, simulationSteps, t_conversion,
df_VehicleTrajectories, signalsAtTimes, ax, cbar_ax):
        simTime = t/(1/t_conversion) #the real time equivalent value of the
simulation

        index = round(floor((simTime + ped.startTime) / round(1/ped.walkSpeed,1))
* round(1/ped.walkSpeed,1),1)

        if index in ped.simulationResults:
            pedData = ped.simulationResults[index]
            map_np = np.asarray([[float(x) for x in line] for line in
pedData[1][0]]) # create map array compatible with matplotlib using np
            ax.cla()

            sns.heatmap(ax = ax, data = map_np, cmap = 'RdYlGn_r', cbar_ax =
cbar_ax, vmax=50, annot = False).set(title = 't = ' +
str(round(simTime+ped.startTime,1))+ ' South CW = ' +
signalsAtTimes[str(round(simTime+ped.startTime,1))]['2'])

            sns.scatterplot(x = [ped.origin_x], y = [ped.origin_y],
color='purple', ax = ax, marker = "$O$") #plot origin positions
            sns.scatterplot(x = [ped.destination_x],y = [ped.destination_y],
color='blue', ax = ax, marker = "$D$") #plot destination positions

            shownPath = pedData[0]
            t_shownpath, x_shownPath, y_shownPath = zip(*shownPath)
            sns.scatterplot(x = x_shownPath, y = y_shownPath, ax = ax, color =
'blue') #plot ped path of displayed pedestrian

            df_vehiclesatcurrenttime =
df_VehicleTrajectories[df_VehicleTrajectories['$VEHICLE:SIMSEC']==simTime+ped.sta
rtTime]

            sns.scatterplot(x = df_vehiclesatcurrenttime['X'],y
= df_vehiclesatcurrenttime['Y'], color='red', ax = ax, marker = "D") #plot
vehicle positions

    def Initial(self):
        pass

# %%
'''
Main

```

```

'''
mapFile = 'map.csv'
linkFile = 'links.csv'
signalPhaseFile = 'SignalPhases.csv'
surfaceFile = 'SurfaceList.csv'
vissimTrajectoryFile = 'Phillip St Miblock Location_001.fzp'
vissimSignalFile = 'University Ave & Phillip_001_Signal.lsa'

if __name__ == '__main__':

    baseMap = ImportMap(mapFile)
    linkMap = ImportMap(linkFile)
    signalPhaseMap = ImportMap(signalPhaseFile)
    df_VehicleTrajectories = ConvertFZP(vissimTrajectoryFile)
    signalsAtTimes = ConvertLSA(vissimSignalFile)
    surfaceDict = ImportSurfaceTypes (surfaceFile)

    simulationTime = 200 #seconds

    numPeds = 1 # number of peds to simulate per run

    vehicleLinkDict_X={}
    vehicleLinkDict_Y={}
    vehicleLinkDict_Speed={}

    for j in range(1): # if desired can run the simulation multiple times and
save the plot, so you don't have to sit at the computer
        ListofPeds = []
        simulationResultsList = []

        for p in range(numPeds):
            startTime = rng.choice(list(np.round(np.arange(900, 4000, .1),1)))
            # while 'red' not in signalsAtTimes[str(startTime)][2]: #this code
is useful if trying to start the pedestrian during a specific signal indication
            #     startTime = rng.choice(list(np.round(np.arange(900, 4000,
.1),1)))

            ped = Pedestrian(baseMap,linkMap, signalPhaseMap,
surfaceDict,startTime)# create new ped to add to simulation

```

```

timeSteps = list(np.round(np.arange(startTime,
startTime+simulationTime, .1),1))

for t in timeSteps: # calculate live costs at each time step

    if(round((t) / round(1/ped.walkSpeed,1),1)) % 1 == 0:
        timefiltered_VehicleTrajectories =
df_VehicleTrajectories[(df_VehicleTrajectories['$VEHICLE:SIMSEC'] == t)] #filter
and organize vehicle data

        vehicleLinkDict = createVehicleLinkList(linkMap,
timefiltered_VehicleTrajectories)
        for key,value in vehicleLinkDict.items():
            vehicleLinkDict_X[key]= value['X'].tolist()
            vehicleLinkDict_Y[key]= value['Y'].tolist()
            vehicleLinkDict_Speed[key]=value['SPEED'].tolist()

        mapAtTime =
ped.CreateLiveMap(vehicleLinkDict_X,vehicleLinkDict_Y,vehicleLinkDict_Speed,signa
lsAtTimes[str(t)]) #calculate live costs
        ped.liveMap.append(mapAtTime)

array = np.array(ped.liveMap)
ped.averageMap = np.mean(array, axis = 0) # calculate the expected
costs

ped.FindPath(ped.origin_x,ped.origin_y) # first time find path

for t in timeSteps:
    if(round((t) / round(1/ped.walkSpeed,1),1)) % 1 == 0:
        ped.simulationResults[t]=g =
pickle.loads(pickle.dumps([ped.plannedPath,ped.perceivedMap], -1)) #save
information for records
        ped.UpdatePosition() ##creating graph, recalculate all ped
paths, and update their positions
        print(t)
        print(signalsAtTimes[str(t)])
        print(ped.actualPath)
        if ped.active_status == False:
            break

ListofPeds.append(ped)

```

```

##### gif output #####
simulationSteps = int(len(ListofPeds[0].actualPath)/0.1)

currentSimulation = SimulationAnimation(ListofPeds[0],simulationSteps,
0.1, df_VehicleTrajectories, signalsAtTimes)

# currentSimulation.ani.save(os.path.join(sys.path[0], 'simulationVideo' +
str(j) + '.gif'), writer="pillow", fps=30)

##### still plot output #####
grid_kws = {'width_ratios': (0.9, 0.05), 'wspace': 0.2}
fig, (ax, cbar_ax) = plt.subplots(1, 2, gridspec_kw = grid_kws, figsize =
(10, 8))

map_np = np.asarray([[float(x) for x in line] for line in
ped.perceivedMap[0]]) # create map array compatible with matplotlib using np

sns.heatmap(ax = ax, data = map_np, cmap = 'RdYlGn_r', cbar_ax = cbar_ax,
annot = False).set()

get_colors = ["#%06x" % random.randint(0, 0xFFFFFF) for _ in
range(len(ListofPeds))]

for i in range(len(ListofPeds)):
    ped = ListofPeds[i]

    colour = get_colors[i]

    shownPath = ped.actualPath
    x_shownPath, y_shownPath = zip(*shownPath)
    sns.scatterplot(x = x_shownPath, y = y_shownPath, ax = ax,
color=colour, alpha = 0.5) #plot ped path of displayed pedestrian
    sns.scatterplot(x = [ped.origin_x], y = [ped.origin_y],
color='black', ax = ax, marker = "$O$") #plot origin positions
    sns.scatterplot(x = [ped.destination_x], y = [ped.destination_y],
color='black', ax = ax, marker = "$D$") #plot destination positions

    #plt.savefig(os.path.join(sys.path[0], 'plot'+ str(j) + '.png'))

plt.show()

```

```
''' #old code to write results to csv
    with open('out.csv','w',newline='') as f:

        # fieldnames lists the headers for the csv.
        w =
csv.DictWriter(f,fieldnames=['actualPath','origin','destination'])
        w.writeheader()

        for ped in ListofPeds:
            # Build a dictionary of the member names and values...
            w.writerow({'actualPath':getattr(ped,'actualPath'),'origin':getat
tr(ped,'origin'),'destination':getattr(ped,'destination')})
    '''
```

Appendix B Fisher's Test R Code

Example for O-D 3-1 in Test Case #1

```
TAB = rbind(c(12, 88), c(4, 26)); TAB
```

```
fisher.test(TAB)
```

Output:

```
      [,1] [,2]
[1,]   12   88
[2,]    4   26

      Fisher's Exact Test for Count Data

data:  TAB
p-value = 0.7624
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.2413264 4.0952728
sample estimates:
odds ratio
 0.8872189
```

Since $p \gg 0.05$ we reject the null hypothesis.