Microscopic Agent-Based Modeling and Simulation of Cyclists on Off-Street Paths

Hossameldin Mohammed

B.Sc., Cairo University, 2012

M.Sc., Cairo University, 2016

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Civil Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2023

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Microscopic Agent-Based Modeling and Simulation of Cyclists on Off-Street Paths

submitted by Hossameldin Mohammed in partial fulfilment of the requirements for

the degree of Master of Applied Science

in Civil Engineering

Examining Committee:

Dr. Alexander Bigazzi, Associate Professor, Department of Civil Engineering, UBC
Supervisor

Dr. Tarek Sayed, Professor, Department of Civil Engineering, UBC
Co-supervisor

Abstract

Inclusion of bicycle traffic in microsimulation tools is essential for evaluating bicycle-accessible infrastructure projects. However, the representation of bicycles in microsimulation models is still at an early stage of development. A better understanding of cyclist behaviour during various interactions is needed to enhance bicycle microsimulation models, which is a pre-requisite for accurate microscopic modeling of bicycle traffic operations. Due to the limited availability of detailed data, the inherent complexity of cyclist decision-making, and the substantial heterogeneity in cycling behaviour, modeling cyclist operation behaviour requires novel methods and techniques. This thesis aims first to characterize cyclist maneuvers in following and overtaking interactions using multivariate finite mixture model-based clustering. Second, an agent-based bicycle simulation method is proposed to model cyclists as intelligent agents making operational and tactical decisions based on their observations of the operating environment. Cyclist position data associated with time stamps are used to infer state and future decisions. The data are extracted from videos collected in Vancouver, BC, Canada using computer vision techniques. For segmenting behavioural states, observations of cyclists in following interactions are clustered into constrained and unconstrained states. Observations of overtaking cyclists are clustered into initiation, merging and post-overtaking states. Generative adversarial imitation learning (GAIL) is used to infer the uncertain intentions and preferences of cyclists from observational data. The model is validated by comparing multivariate distributions of variables such as speed, direction, and spacing of observed and simulated cyclist trajectories. The model performs well in comparison to two other cyclist simulation models from the literature. The proposed approach to miscrosimulation is a significant advancement in agent-based modeling methods, with continuous,

non-linear, and stochastic representation of states, decisions, and actions. By modeling cyclist heterogeneity, the proposed approach can enhance applications in bicycle facility planning and design, safety modeling, and energy modeling with consideration of the full diversity of cyclists. Such an advancement is necessary for developing bicycle networks for all ages and abilities of riders.

Lay Summary

Developing frameworks for understanding and modeling cyclist movements on pathways is an important area of research. Integrating cyclist-specific behavioural models into traffic microsimulation tools can enhance their utility for active transportation infrastructure projects and studies. In this thesis, machine learning is used first to segment the stages of cyclist movements to gain a better understanding of their movement patterns. A machine learning model is then developed to predict the actions of cyclists in different situations. These models are trained on data extracted from videos recorded in Vancouver, BC, Canada.

Preface

This thesis chapters are published as two papers in respectable journals. These are as follows:

1. Portions of the introduction in Chapter 1, the literature review in Chapter 2, the data collection in Chapter 3, the method as well as the results and discussion in Chapter 4 and conclusions in Chapter 6 have been published in:

Mohammed, H., Bigazzi, A. Y., & Sayed, T. (2019). Characterization of bicycle following and overtaking maneuvers on cycling paths. Transportation research part C: emerging technologies, 98, 139-151.

I conducted all the analyses and wrote most of the manuscript under the supervision of Dr. Sayed and Dr. Bigazzi.

2. Portions of the introduction in Chapter 1, the literature review in Chapter 2, the data collection in Chapter 3, the method as well as the results and discussion in Chapter 5 and conclusions in Chapter 6 have been published in:

Mohammed, H., Sayed, T., & Bigazzi, A. (2022). Microscopic modeling of cyclists on off-street paths: a stochastic imitation learning approach. Transportmetrica A: transport science, 18(3), 345-366.

I conducted all the analyses and wrote most of the manuscript under the supervision of Dr. Sayed and Dr. Bigazzi.

Table of Contents

List of Tables

List of Figures

List of Symbols

$\mathcal{A}$	Set of cyclist actions
$\mathcal{D}$	Set of expert demonstrations
$\mathcal{L}$	Neural network loss function
$\mathcal{R}$	Reward function for a given state $s \in \mathcal{S}$
$\mathcal{S}$	Set of cyclist states
$\mathcal{T}$	Set of state transition probabilities
γ	Discount factor for future rewards $\in [0,1]$
θ	Set of parameters
$\pi_\theta(a\|s)$	Policy function for states and actions
ζ	Trajectory of a cyclist

List of Abbreviations

GAIL	Generative Adversarial Imitation Learning
GFMM	Finite Mixed Mixture Model
IID	Independent and Identically Distributed
KL	Kullback-Leibler Divergence
LSTM	Long-Short Term Memory
MDP	Markov Decision Process
VAE	Variational Autoencoder

Acknowledgements

I offer my enduring gratitude to Dr. Tarek Sayed and Dr. Alex Bigazzi for their continuous support, inspiration and patience throughout my research journey. I owe much of my scholarship development to their continuous guidance.

I would also like to thank colleagues at the transportation group for their friendship and for providing a welcoming environment.

Dedication

I dedicate this thesis to my wife Marwa, and my daughter Yasmina!

Chapter 1: Introduction

1.1 Motivation

Active modes of transportation such as cycling are promoted as a way to provide health benefits, mitigate traffic congestion and reduce air pollution (Götschi, Garrard, & Giles-Corti, 2016). As such, there is significant interest in research to develop cycling behaviour models to enhance the design and evaluation of investments in bicycle infrastructure (Heinen, Bert, & Kees, 2010). Better understanding of how cyclists move and interact is essential for developing bicycle microsimulation models, which can be used in traffic modeling, safety evaluation, energy and health modeling, and more.

Traffic microsimulation model development to date has focused on motor vehicle traffic. These existing tools can help guide the development of bicycle microsimulation, though the operation of bicycles differs from that of motor vehicles in fundamental ways. Advances in motor vehicle traffic simulation models can guide the development of bicycle traffic simulation models. Since the 1960s, many car-following and lane-change models have been developed for motor vehicle traffic simulation. Car-following models can be generally classified into three main types (Aghabayk, Sarvi, & William , 2015): stimulus response models such as the General Motors (GM) model (Chandler, Herman, & Montroll, 1958) & (Gazis, Herman, & Rothery, 1961), safe distance models such as the Gipps model (Gipps, 1981), and action point models such as the Wiedemann model (Leutzbach & Wiedemann, 1986). Most of these models follow a general framework of first identifying the state or the regime a vehicle is in by acquiring perceptual variables such as speeds and positions of that vehicle and its adjacent interacting vehicles (Aghabayk, Sarvi, & William , 2015), (Ahmed & Ben-Akiva, 1996). For example, VISSIM (PTV, 2011) utilizes an action point model by identifying different regimes for each vehicle in each time step, i.e. free driving,

approaching, following, and braking regime (Fellendorf & Vortisch, 2001). For each regime, VISSIM predicts a vehicle's acceleration or deceleration in the next time step based on speed, speed difference, distance and individual characteristics of driver and vehicle. A vehicle can transition from one regime to another if it exceeds a certain threshold described as a combination of speed difference and distance. Another example is AIMSUN (AIMSUN User's Manual, 2005), which utilizes a safe distance model. A vehicle in AIMSUN can be in a free-flow or constrained regime. Different rules are applied to vehicles in each state to predict acceleration in the next time step (Rakha & Wang, 2009). Lane change models follow a similar approach to car-following models by modeling the lane change process as a series of decisions, such as the decision to initiate a lane change and the target lane (Ahmed, 1999).

Bicycle traffic is more maneuverable, and more affected by operator attributes including physical capability, gender, age, and fatigue (Twaddle, Schendzielorz, & Fakler, 2014). The two basic models used for motor vehicle microsimulation are car-following and lane change. These two models describe the decisions of drivers to take different guidance actions (e.g., accelerate, decelerate, change lane) based on stimuli from the environment and interacting motor vehicles (Aghabayk, Sarvi, & William, 2015). The adaptation of these methods to bicycle traffic must account for differences such as the absence of lanes, which allows for flexible utilization of lateral road space by cyclists and changing lateral positions to allow for more unconstrained riding or overtaking. Most car-following models assume a stimuli response mechanism, where motor vehicle actions are highly determined by leading vehicle actions. We cannot assume these same stimuli-response models apply to bicycle traffic given the differences in independence of movement. Also, stochasticity in behaviour is more pronounced in bicycle traffic, in the form of inter-cyclist (differences between cyclists due to different characteristics) and intra-cyclist

(variability in behaviour of an individual cyclist when encountering similar situation in different time periods) heterogeneity (Taylor & Mahmassani, 1998).

Cycling behaviour and bicycle maneuvers in traffic are different from those of motor vehicles, due to distinct physical and dynamic characteristics (Twaddle, Schendzielorz, & Fakler, 2014). Bicycle traffic flow is often non-lane based and cyclists more freely utilize lateral road space (Taylor & Mahmassani, 1998). Bicycle traffic is also characterized by high heterogeneity in behaviour and performance among cyclists (Hoogendoorn & Daamen, 2016). Some commercial traffic microsimulation software such as VISSIM, PARAMICS and AIMSUN provide the option to simulate bicycle traffic using modified variations of motor vehicle microsimulation models. The performance of these models has not been sufficiently validated (Carrignon, 2009), (COWI, 2012).

There is increasing interest in bicycle microsimulation models that can be used for the design and evaluation of bicycle infrastructure investments (Heinen, Bert, & Kees, 2010). The developments in bicycle microsimulation models is relatively new, compared to motor vehicle microsimulation. The concepts from motor vehicle microsimulation models can help guide the development of bicycle microsimulation models, although the operation of bicycles differs from that of motor vehicles in several fundamental ways (Twaddle, Schendzielorz, & Fakler, 2014). The two basic models used for modeling motor vehicle microsimulation are car-following and lane change models. These two models describe the decisions of motor vehicle drivers to take different actions (e.g., accelerate, decelerate, change lane to a target lane), given previous stimuli from the environment and other interacting motor vehicles (Aghabayk, Sarvi, & William , 2015). Adapting concepts from motor vehicle microsimulation models for bicycle traffic must take into account differences such as being a non-lane-based traffic that can freely utilize lateral road space (Taylor & Mahmassani, 1998). Bicycle traffic is also characterized by high heterogeneity in behaviour and

performance among cyclists (Hoogendoorn & Daamen, 2016). Some commercial traffic microsimulation software such as VISSIM, PARAMICS and AIMSUN provide the option to simulate bicycle traffic using modified variations of motor vehicle microsimulation models. The performance of these models has not been sufficiently validated (Carrignon, 2009; COWI, 2012).

Agent-Based Modeling (ABM) is a potentially powerful approach for realistic modeling of human-directed movements and interactions in roadway environments. ABM offers flexibility in developing models that are robust and scalable, which can capture the complexity of real-world behaviour that emerges from the microscale into the macroscale (Jennings, 2000). One of the tasks in developing ABM is modeling of the agent's goals and strategies. Agent strategies can be modeled in a rule-based fashion by heuristically extracting rules that govern their behaviour in different situations (Baster, Duda, Maciol, & Rebiasz, 2013; Hussein & Sayed, 2017). Another approach in ABM is to design intelligent adaptive agents that can learn from experiences and interactions with other agents, imitate experts' demonstrations, and evolve by changing their goals and strategies over time (Plekhanova, 2003).

Building intelligent bicycle microsimulation agents is a promising approach to model bicycle traffic. Developing accurate bicycle microsimulation models can have many implications on modeling bicycle infrastructure, cyclist power and energy modeling, and in-depth analysis of bicycle behavioural characteristics for safety (Heinen, Bert, & Kees, 2010).

1.2 Problem Statement and Research Objective

The overall goal of this research is to develop a functioning agent-based bicycle microsimulation model. An agent-based microsimulation model requires a model that can describe all aspects in microscopic cyclist motion. In order to be able to generalize the model, evaluation of uncertainty of behaviour in various interactions needs to be done. To achieve the overall goal

of our research, it needs to be divided into certain interconnecting steps that form our final objective together.

The first step of our research is the exploratory analysis of the data. That is done by characterizing cyclist interactions and the identification of the distributions of variables in each state is an important step toward the development of agent-based microscopic simulation models. Exploratory analysis of cyclist trajectories can reveal interesting insights about the multivariate relationships between different motion variables during following and overtaking interactions. The first research objective can be stated as:

1) "Characterize cyclist states in following and overtaking interactions".

The scope of investigation includes unidirectional cyclists on dedicated bicycle paths. A Gaussian Finite Mixture Model (GFMM) is used to cluster the perceptual variables of longitudinal distances, lateral distances and speed differences between interacting cyclists.

Agent-based microsimulation models require rules for agents in distinct traffic situations to predict decisions such as acceleration, change in longitudinal and lateral positions, etc. For example, the rules applied for cyclist agents in free flow regime will be different from rules applied in constrained or overtaking regimes. The determination of thresholds of perceptual variables between regimes is an essential step in bicycle microsimulation. Determination of cyclist regimes is also important for prediction of cyclist behaviour in different situations, which is essential in safety evaluation, particularly for shared-space infrastructure and connected vehicle technology.

By modeling cyclist motion as Markov Decision Process (MDP), a cyclist is an agent choosing actions based on state sensory variables. The agent chooses the actions which maximize the discounted long-term reward. The objective is to estimate the parameters of a generative neural network model that uses generative adversarial imitation learning to infer the movement patterns

of cyclists at different situations and predict their actions in a simulation framework. The second research objective can be stated as:

2) "Build a generative predictive neural imitation learning model for cyclist movement prediction and simulation"

Although unconstrained cyclists may seem to move arbitrarily, it is hypothesized that their movements can be modeled given variables such as acceleration, change in direction and deviation from path centerline. Unconstrained cyclist trajectories can be thought of as correlated random walk movements, in which the state of the present time step is dependent on the previous time step. Variational autoencoders (VAE) can be used to infer the hidden states that cyclist encounter at each time step given observations about their movement. The third research objective can be stated as:

3) "Formalize a generative model that can describe unconstrained cyclist movement and be used to generate cyclist trajectories that are similar to observed ones. This model can then be combined with other constrained bicycle models to generate realistic interactions for bicycle microsimulation".

As a summary, the overall goal to build a functioning bicycle microsimulation agent-based model is achieved by following the listed sub-goals and integrating them together. Due to the high variability in cyclists' behaviour, the uncertainty of the developed models needs to be assessed and taken into account in the simulation platform. Modeling the unconstrained cyclists will give more flexibility to the model to be used in different situations by forming a concrete model of environment dynamics.

Chapter 2: Literature Review

In recent years, there is an increase in cycling frequency for transportation not only for recreational purposes, but also for commuting, shopping and other trip purposes (Heinen, Bert, & Kees, 2010). That attracted some attention among transportation researchers to gain more understanding of cyclists' microscopic behaviour. Insights about different cyclist motion regimes, such as cruising, acceleration and deceleration were gained from the analysis of bicycle GPS data (Ma & Luo, 2016). Distributions of interacting bicycles' headways were developed and segmented to constrained and unconstrained headways (Hoogendoorn & Daamen, 2016). Also, bicycle overtaking interactions were analyzed using observations of bicycle maneuvers on separated bicycle facilities by comparing speeds of overtaking bicycles at different overtaking states. In addition, attempts of experimental research involving bicycles on a circular track was conducted to develop a model for bicycle following and a unified following model for motor vehicle, bicycle and pedestrian flow. To better understand cyclists' acceleration process, a methodology was developed to predict acceleration profiles given information only about the speed. More understanding of how bicycles move and interact in following and overtaking maneuvers is still needed for the development of accurate agent-based microsimulation models.

A number of bicycle specific microsimulation models were developed using cellular automata (CA) modeling approach. A stochastic cellular automata model that discretizes the environments into cells containing multiple bicycles was developed by (Jiang, Jia, & Wu, 2004), but no validation with field data was conducted. A CA model for bicycle traffic that assumes that each cell represents one bicycle and bicycle speed depends on the leading bicycle speed in preceding time steps in a discrete environment (Xue, Jia, Jiang, Li, & Shan, 2017). Other CA modeling efforts have been made towards mixed bicycle and electric bicycle traffic (Zhao D. , et

al., 2013). The developed model was used to predict the number of the overtaking events at different bicycle traffic densities. There are numerous advantages to the CA modeling approach to the generic analytical models. These advantages include taking into account the correlation between different input variables and the flexibility in making changes to the modeling environment (Wolfram, 1984). However, CA models are criticized by using discrete representation, which in many cases does not provide accurate representation of the problem (Chopard, 2012). Other models such as social force (SF) model was used by Liang et al. (2012), in which they developed a SF model that describes bicycle traffic as either free-flow or congested flow regimes. Further research is needed to characterize cyclist states in various types of interactions and develop representative bicycle traffic microsimulation models in a variety of contexts.

Agent-Based Modeling (ABM) is a potentially powerful approach for realistic modeling of human-directed movements and interactions in roadway environments. ABM offers flexibility in developing models that are robust and scalable, which can capture the complexity of real-world behaviour that emerges from the microscale into the macroscale (Jennings, 2000). One of the tasks in developing ABM is modeling of the agent's goals and strategies. Agent strategies can be modeled in a rule-based fashion by heuristically extracting rules that govern their behaviour in different situations (Baster, Duda, Maciol, & Rebiasz, 2013; Hussein & Sayed, 2017). Another approach in ABM is to design intelligent adaptive agents that can learn from experiences and interactions with other agents, imitate experts' demonstrations, and evolve by changing their goals and strategies over time (Plekhanova, 2003).

Microscopic behaviour of a cyclist moving along a path can be framed as a finite-state Markov Decision Process (MDP). This framework defines a sequential decision process in which

a decision and consequent action at each time step is dependent on the current state of the cyclist and a decision policy that seeks to maximize a utility or reward function. To solve an MDP and calculate the optimal policy (representing the predicted decision sequence and thus behaviour), the reward function must be defined, and the environment dynamics expressed by transition probabilities between states conditional on each possible action in the decision space. The environment dynamics (transition probabilities) can be estimated from an inherent understanding of the environment, or from a sufficiently large sample of action/state sequences. Specifying and estimating the reward function can be more difficult. Inverse reinforcement learning (IRL), developed by Ng. and Russel (Ng & Russel, 2000), aims to solve this problem by estimating a reward function given demonstrations from an expert. In the case of traffic modeling, the expert demonstration data can be observations of real-world decisions/behaviour.

The problem of estimating agent policy without knowing the underlying reward function with the help of observed demonstrations from an expert is often called imitation learning or apprenticeship learning. One solution to the imitation learning problem is using behaviour cloning (BC) (Bain & Sammut, 1995). BC aims to directly learn a policy from expert demonstrations using any supervised learning approach. One major drawback of behaviour cloning is the lack of generality in the learned policy (Ratliff, Silver, & Bagnell, 2009). Learning the reward function is usually called inverse reinforcement learning (IRL) or inverse optimal control (IOC). The major advantage of inductive (IRL) over deductive (BC) learning is the capability to produce generalizable theories not limited to the training information (Plekhanova, 2003). A comparison between IRL and BC is depicted in Table 2-1.

Table 2-1: Comparison between inverse reinforcement learning and behaviour cloning.

Inverse Reinforcement learning	**Behaviour Cloning**
Learning a cost (reward) function	Learning direct mapping between input and output
The main objective is to find reward function that explains the demonstrations	The main objective is to generate similar demonstrations
Generalizable	Generalization is not guaranteed

The IRL problem was first introduced by Ng and Russel (2000), suggesting a linear programming solution for optimizing the difference between observed and demonstrated behaviour. Ng and Russel (2000) described the IRL as an ill-posed problem, which can have many degenerate solutions to the reward function (e.g., all zeros). As a solution, some previous studies such as (Abbeel & Ng, 2004; Ratliff, Bagnell, & Zinkevich, 2006) proposed a methodology for estimating a reward function that has the maximum margin to any alternative solutions. Ziebart et al. (2008) introduced using of the principle of maximum entropy (Jaynes, 1957) for better inference of reward function that does not stick to the assumption of agent optimality, assuming exponential increase in trajectory probability with reward increase. Gaussian Processes (GP) was used by (Levine, Popovic, & Koltun, 2010) to estimate the nonlinear structure of the reward function and can return the distribution of the reward function not only its mean values.

A major drawback with the previously discussed methods is that they require discrete formulation of state and action features. The behaviour of bicycles in interactions is a continuous control task. Trying to discretize all the state sensory and action motor features would be a challenging task that would significantly increase the dimensionality of the problem causing what

is referred to as the “curse of dimensionality” (Sutton & Barto, 1998). The computational complexity is even more pronounced in IRL than the forward RL, as IRL algorithms need to run RL at every iteration in an internal loop for policy evaluation. Levine & Koltun (2012) introduced a methodology to estimate the reward function in continuous state and action environments by considering the shape of the learned reward function in the neighborhood of the expert demonstration instances. The method requires the calculation of the gradient and hessian matrices of the reward function with respect to states and actions instead of the discrete tabular form, in which the algorithm involves iterating over a three-dimensional array.

Trajectory planning models specialize in finding paths that are collision free and optimal given the agent’s reward function (Mahajan & Marbate , 2013). Trajectory planning algorithms usually rely on a search algorithm that finds the optimal path given the environment sensory inputs (Biagiotti & Melchiorri, 2008). The environment in trajectory planning algorithms is either static (Garrido, Moreno, & Abderrahim, 2006), or dynamic (Sud, Anderson, Curtis, Lin, & Manocha, 2007). Path planning has recently been extensively used in transportation research for autonomous vehicles trajectory planning (Wei, et al., 2019; Li & Li, 2019).

As a summary, Previous research attempts have been made to accurately model microscopic bicycle traffic for traffic simulation applications, but most of the research lacked calibrating and validating these models using field data. Also, most of the developed models in the literatures lack accurate representation of behaviour by discretizing the simulation environment and/or not including all the significant factors that may affect cyclists’ microscopic traffic characteristics. This research is trying to fill-in these gaps in the literature by building data-based models using field video data and using state-of-the-art techniques to capture most of the factors that affect cyclists’ microscopic behaviour.

The most commonly used method for modeling microscopic cyclist behaviour is cellular automata (CA) (Wolfram, 1983). CA uses a discrete space formulation wherein cells occupied by cyclists change according to the movement of cyclists following a set of rules controlling their motion and interactions. Jiang et al. (2004) developed a stochastic CA model that discretizes the environment into cells containing multiple bicycles. Jia et al. (2007) developed a CA model that differentiates fast-moving and slow-moving cyclists, and compared deterministic and stochastic rules. Gould et al. (2009) extended the CA modeling approach for multiple lanes and multiple types of cyclists. Xue et al. (2017) developed a CA model in which each cell can represent one bicycle, and each bicycle's speed is affected by a leading bicycle's speed in preceding time steps. Tang et al. (2018) formulated a CA model that can account for group behaviour. Although CA has been extensively used in the bicycle microsimulation literature, it has limitations including the lack of a continuous state-space representation. Also, CA models only allow a preset number of agent groups, each with its own set of rules, which limits the ability to model multiple levels of heterogeneity and environment dynamics (Wolfram, 1985). These limitations of CA models restrict the flexibility required to model bicycle traffic heterogeneity.

Other methods of modeling microscopic cyclist operations are derived from homogeneous rules of behaviour. Liang et al. (2012) developed a psychological-physical force model that assumes cyclists make guidance decisions (acceleration and direction) based on forces of collision avoidance and friction between interacting cyclists. Zhao and Zhang (2017) developed a following model that can be applied to motor vehicles, bicycles, and pedestrians by assuming the same stimuli-response behavioural mechanism, but with different parameter values. The parameters were estimated from experimental data of motor vehicles, bicycles and pedestrians moving in a

circle. These methods model the average behaviour of cyclists, but cannot represent inter-cyclist and intra-cyclist behaviour heterogeneity.

Agent-Based Modeling (ABM) is a potentially powerful approach for realistic modeling of human-directed movements and interactions in roadway environments. ABM offers flexibility in developing models that are robust and scalable, which can capture the complexity of real-world behaviour that emerges from the microscale into the macroscale (Jennings, 2000). Using ABM approach allows models to capture the variability in decision-making rules in varying situations and between heterogeneous agents. For example, interacting cyclists could encounter different behavioural schemes depending on their relative values of motion variables (Mohammed, Bigazzi, & Sayed, 2019).A key task in developing ABM is identification of agent goals and strategies. One approach to this task is using a structured analytical model with parameters that govern agent behaviour in different situations, selected using heuristic techniques (Baster, Duda, Maciol, & Rebiasz, 2013; Hussein & Sayed, 2017). This approach to ABM has been used in various transportation applications, including modeling the effect of 'Mobility as a Service' trends (Djavadian & Joseph, 2017), solving transit network design problems (Liu & Zhou, 2016), and modeling the interactions between autonomous vehicles using a multi-agent framework (de Oliveira, 2017). The drawback of this approach for bicycle microsimulation is the use of pre-set rules; agents do not learn from the environment or other agents, nor do they evolve through the learning process (Abdou, Hamill, & Gilbert, 2012).

Another approach to ABM development is designing intelligent adaptive agents that can learn from experience by imitating expert demonstrations and evolving their goals and strategies over time (Plekhanova, 2003). To apply this approach to bicycle microsimulation, the navigation and guidance behaviour of a cyclist moving along a path can be framed as a finite-state Markov

Decision Process (MDP). In this framework, a cyclist is an agent interacting with the environment (roadway features, other road users, traffic controls, etc.) and executing a sequential decision process. At each time step, cyclists base their decisions on a function which is called a policy throughout this paper. Policy is defined as a mapping function from the agent's current state to its consequent action. The main component of a policy is its reward function that represents the attractiveness of potential future states to the agent, following the formal approach of (forward) reinforcement learning. Solving a reinforcement learning problem requires a defined reward function, which is usually defined by the analyst for *ad hoc* control tasks. For the problem of modeling cyclist guidance behaviour, *a priori* definition of a reward function is difficult due to the unknown multivariate relationships among state and action variables that describes the exact trade-offs between different desiderata, and the unknown structure of the reward function itself. This issue is similar to the difficulty in specifying a reward function for robot driving tasks described in (Abbeel & Ng, 2004), or in the discrete model developed for inferring a discrete reward function for cyclists' on cycling paths (Mohammed, Sayed, & Bigazzi, 2019).

Imitation learning can be used to avoid *a priori* definition of a reward function; imitation learning is the estimation of an agent policy (without a pre-defined reward function) based on observed demonstrations of "expert" agents. Behaviour cloning aims to directly learn a policy from expert demonstrations using a supervised learning approach (Bain & Sammut, 1995). A major drawback of behaviour cloning is the lack of generality in the learned policy (Ratliff, Silver, & Bagnell, 2009). Learning the reward function is another approach, usually called Inverse Reinforcement Learning (IRL) or Inverse Optimal Control (IOC). The major advantage of inductive (IRL) over deductive (behaviour cloning) learning approaches is the ability to produce generalizable theories not limited to the training information (Plekhanova, 2003).

Reward function learning has several distinct advantages over supervised learning in the context of sequential decision making. The reward in reinforcement learning is defined for all states, allowing an agent to receive a learning signal even from states that are not observed in the training dataset. In contrast, in supervised learning, an agent only receives a score for states included in the labeled observations. In addition, reinforcement learning maximizes the global expected return on a trajectory, rather than the local response to a specific observation. Agents in reinforcement learning have the ability to plan their trajectories considering all future scenarios and may take unusual short-term actions to avoid highly impactful later states.

In summary, a major gap in bicycle microsimulation literature is the limited representation of heterogeneity and uncertainty in cyclist behaviour. The current state-of-the-art bicycle microsimulation models use either cellular automata or homogeneous analytic approaches, which do not have the ability to represent the full heterogeneity of cyclist preferences, intentions and interactions. This paper proposes a new methodology for modeling microscopic behaviour of cyclists, imitation learning, that allows inference of hidden cyclist preferences which can then be used to predict stochastic cycling behaviour in changing environments.

Few studies have investigated states or regimes in bicycle traffic. Liang et al. developed a social force model that describes bicycle traffic as either free-flow or congested flow regimes (Liang, Baohua, & Qi, 2012). Ma and Luo used bicycle GPS data to identify cruising, acceleration and deceleration regimes; interactions were not examined (Ma & Luo, 2016). Zhao and Zhang (Zhao & Zhang, 2017) used an experiment of bicycles on a circular track to develop a model for bicycle following and a unified following model for motor vehicle, bicycle and pedestrian flow, but the model was not validated with on-road data. Hoogendoorn and Daamen (Hoogendoorn & Daamen, 2016) developed distributions of following bicycle headways and segmented into

constrained and unconstrained headways. Khan and Raksuntorn (Khan & Raksuntorn, 2001) studied bicycle overtaking maneuvers on separated bicycle facilities by comparing speeds of overtaking bicycles at different overtaking states. Zhao et. al. (Zhao D. , et al., 2013) developed and calibrated a cellular automata model for modeling overtaking decisions of mixed bicycle traffic consisting of conventional and electric bicycles. The developed model was used to predict the number of the overtaking events at different bicycle traffic densities. Further research is needed to characterize cyclist states in various types of interactions and develop representative bicycle traffic microsimulation models in a variety of contexts.

Chapter 3: Data Collection

Video data were collected at three locations. The first two locations were on the Brooklyn Bridge, New York City, USA (Figure 3-1). Video data with a frame frequency of 15 Hz were obtained from two cameras mounted on the bridge, recorded on September 20, 2016, from 8:00 to 9:00 and from 17:00 to 18:00 (4 hours total). The video images included a pedestrian-bicycle path four meters in width, divided equally between pedestrians and bicycles by a solid lane marking. The bicycle traffic is bidirectional with a 1% grade in the direction away from the camera. The third location is the Burrard Bridge, Vancouver, Canada (Figure 3-2). Video data with a frame frequency 30 Hz were obtained over three days, April 12 through 14, 2016 from 7:00 to 19:00 (29 hours total). The video image included bicycle traffic in a dedicated unidirectional path (without pedestrians) with a grade of 1% in the direction of travel.

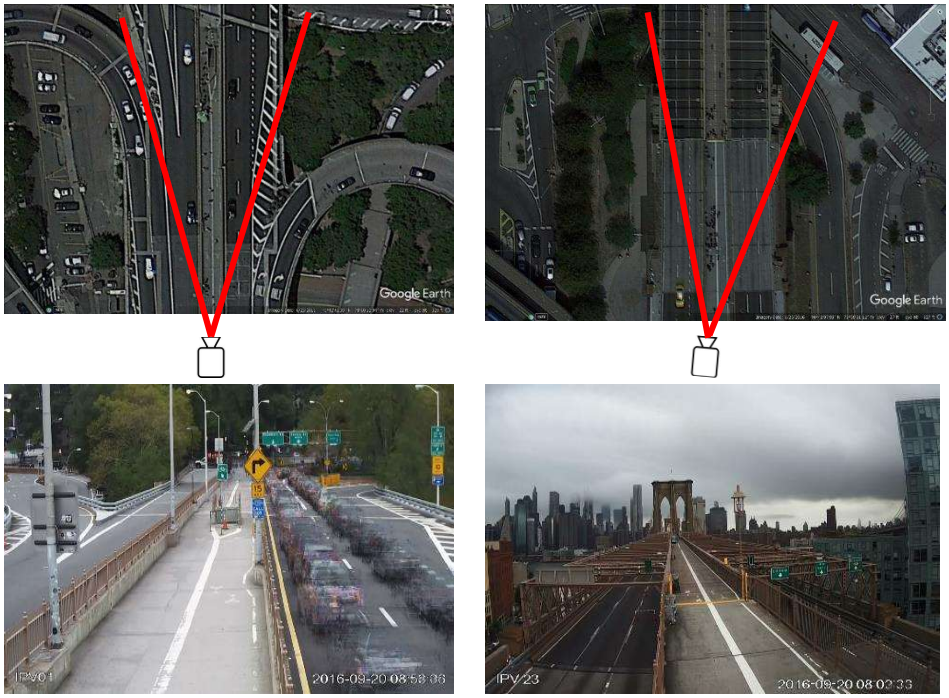

Figure 3-1: Camera positions and field of view of the two video scenes

Figure 3-2: Camera positions and field of view the video scene on Burrard Bridge

Bicycle trajectories were extracted from the videos using automated computer vision techniques developed at the University of British Columbia. A detailed description of the video analysis procedure can be found in (Saunier & Sayed, 2006). The system was used in the collection of microscopic cyclist data (Zaki & Sayed, 2016), (Zaki, Sayed, & Cheung, 2013). The first step

of the video analysis is camera calibration, which relates the two-dimensional video data with real-world three-dimensional coordinates (Ismail, Sayed, & Saunier, 2008). Camera calibration involves determining the location and orientation of the camera, and then estimating a homography matrix which represents the projection parameters between the video data and real-world space.

Moving features in the video data were differentiated from background fixed features and tracked using the Kanade-Lucas-Tomasi Feature Tracker (Lucas & Kanade, 1981), (Tomasi & Kanade, 1994). The tracked features were then grouped based on parameters of vision cues such as spatial proximity and typical size of tracked objects (Figure 3-3). The computer vision detection accuracy decreases with the increase in depth of the trajectory. All the trajectories were truncated at a length of 65 m to account for the decrease in location detection accuracy with the increase in depth. That truncation length was determined by observing higher projection plane distortion after this distance. Counts of the detected bicycles and pedestrians in the four hours of video data are shown in Table 3-1: Bicycle and pedestrian volume counts. Pedestrian trajectories were excluded from further analysis as few interactions with cyclists were observed due to the separation of bicycle and pedestrian lanes on the paths.

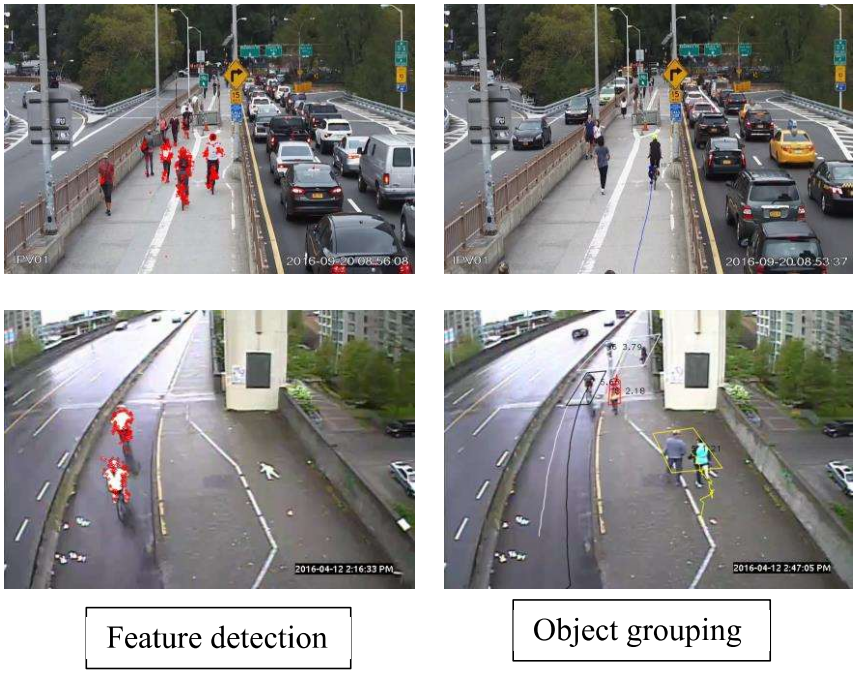

Figure 3-3: Feature detection and object grouping visualization.

Table 3-1: Bicycle and pedestrian volume counts

	Scene 1 Sept. 20, 2016 8:00-9:00	**Scene 1 Sept. 20, 2016 17:00-18:00**	**Scene 2 Sept. 20, 2016 8:00-9:00**	**Scene 2 Sept. 20, 2016 17:00-18:00**	**Total**
Bicycle Volume	421	223	349	131	1124
Pedestrian Volume	365	976	336	1333	3010

The extracted trajectories are composed of a local longitude and latitude for each cyclist at each video frame, relative to a predetermined origin point on the world map of the camera scene view. The spatiotemporal trajectories were used to calculate speed for each cyclist at each video frame by first calculating an instantaneous speed (distance divided by time) and then applying a Savitzky–Golay filter (Savitzky & Golay, 1964) with a seven time-step (7/15 seconds) window size. A validation of the trajectory extraction process was conducted by taking a sample of 30 trajectories from each hour of video (sample of 120 trajectories) and manually measuring the speed

based on the time required to traverse a known distance in the video scene. The mean absolute error between automatically extracted speed and manually measured speed was 0.86 m/s, which was considered negligible to not affect the findings. Almost 95% of the cyclists' trajectories were detected using the computer vision technique with an average trajectory duration of 5 seconds.

Chapter 4: Characterization of Bicycle Following and Overtaking Maneuvers

4.1 Interactions

Bicycle trajectories in a single direction that co-existed in any frame of data were considered to be in potential following or overtaking interactions. The minimum observed time headway between non-interacting cyclists was 10 seconds, which is considered sufficient to treat them as not being involved in an interaction. If three or more trajectories co-existed in a single frame of data, separate interactions were identified between sequential pairs of cyclists. The average length of interacting trajectories was 3 seconds with 1.5 seconds standard deviation. Overtaking interactions were differentiated from following interactions by plotting space-time diagrams to identify overtaking events where the trajectories crossed (Figure 4-1), and then visually confirming by watching the relevant video data. The data were captured over four highly directional hours, in which the two morning hours (8:00 to 9:00) had approximately 90% of the traffic in the direction going to Manhattan, and the other two hours (17:00 to 18:00) had approximately 70% of the traffic in the direction going to Brooklyn. The extracted following and overtaking interactions were rarely concurrent with the presence of counter-flowing traffic, which was consistent with the study scope of unidirectional interactions. Any trajectories with interactions with counter-flowing traffic were excluded from the analysis data. Out of 1124 identified bicycle trajectories, 590 were involved in 316 following interactions (some cyclists were involved in two following interactions simultaneously as leader in one interaction and a follower in the other), and 68 were involved in 34 overtaking interactions (half overtaking and half being overtaken) – see Figure 4-2.

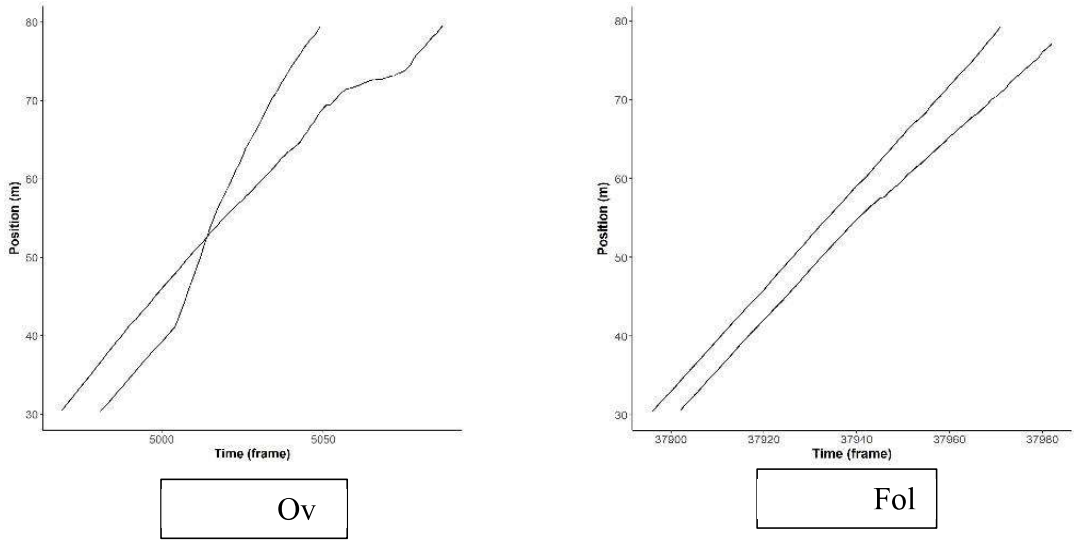

Figure 4-1: Example trajectories in overtaking and following interactions

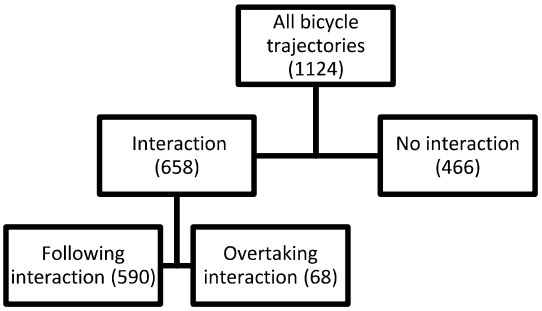

Figure 4-2: Classification of bicycle trajectories by interactions (number of trajectories)

Three perceptual variables were calculated from the trajectories in each interaction, based on previous cycling behaviour research (Liang, Baohua, & Qi, 2012; Hoogendoorn & Daamen, 2016; Zhao D. , et al., 2013): longitudinal distance, lateral distance, and speed difference. The longitudinal distance is the distance along the path between the tracking points of two interacting cyclists, negative when the initially leading cyclist is ahead (as illustrated in Figure 5). The lateral distance is the absolute distance perpendicular to the path between the tracking points of two interacting cyclists (Figure 4-3: Longitudinal and lateral distance between two interacting cyclists).

The speed difference is that of the initially following cyclist minus that of the initially leading cyclist (i.e., positive when the initially following cyclist is faster).

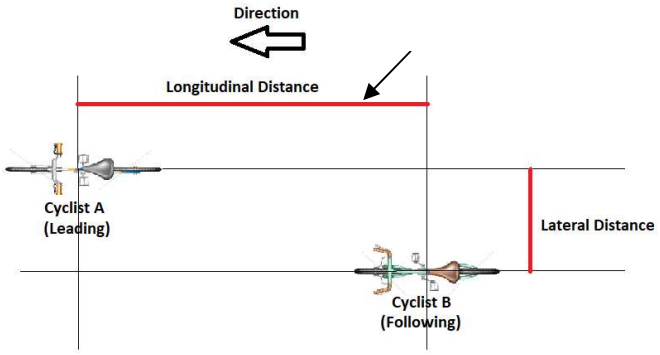

Figure 4-3: Longitudinal and lateral distance between two interacting cyclists

4.2 State identification

The characterization of states of bicycle maneuvers during interactions was conducted by classifying the disaggregate measured perceptual variables (longitudinal distance, lateral distance, and speed difference) into homogeneous subgroups within each interaction. Gaussian Finite Mixture Model (GFMM) was selected as the clustering method because it does not require an assumption of spherical variable distributions within the clusters (Everitt & Hand , 1981). GFMM attempts to recover the model of data generation assuming the data population consists of multiple subpopulations in which the variables have different multivariate joint Gaussian probability density functions. The resulting distributions can vary in features such as size, orientation, and shape, which allows for flexibility in fitting the models (Banfield & Raftery, 1993).

The twelve shape criteria based on re-parameterization of the cluster covariance matrices given in (Banfield & Raftery, 1993) were evaluated using the Bayesian Information Criteria (BIC). The shape model “VVV” (Varying size, Varying shape, and Varying orientation with ellipsoidal covariance) was selected based on lowest BIC for both interaction types. The cluster model was

estimated using the *MCLUST* package in the statistical software R (Scrucca , Fop , Murphy , & Raftery , 2016).

The GFMM clustering method assumes the input observations of the model are independent and identically distributed (IID). The pooled trajectory data used in the GFMM clustering contains repeated measurements of cyclist variables, which leads to potential serial dependence. One approach to model this dependency is to add a random effect coefficient to the finite mixtures of linear mixed models (Celeux, Martin, & Lavergne, 2005). However, previous research found no significant difference between the clustering results obtained with and without random effect coefficients (Pelosi, Alfò, Martella, Pappalardo, & Musarò, 2015). The autocorrelation coefficient was calculated for every variable in the data and the coefficient value had an average value of 0.32 which is considered low. Also, the Durbin-Watson statistic was calculated for the first order autoregressive process, and the value of the statistic was 1.6, which indicates no autocorrelation (Gujarati, 2003). The low autocorrelation coefficient could be due to the large number of individual cyclists in the data which decreases the effect of serial dependence (Guo, Logan, Glueck, & Muller, 2013). As no strong serial dependence was concluded from the data, it was decided not to model the serial dependence in the cluster analysis.

Following interactions represented by the perceptual variables of 316 pairs of cyclist trajectories were clustered into two states of motion, hypothesized to represent constrained and unconstrained flow states based on past research (Hoogendoorn & Daamen, 2016; Buckley, 1968). The overtaking interactions represented by the perceptual variables 34 pairs of cyclist trajectories were clustered into three states of motion, hypothesized to represent initiation, merging, and post-overtaking phases based on past research that examined the relationships between speeds, longitudinal distance and lateral distance in overtaking maneuvers (Khan & Raksuntorn,

Characteristics of passing and meeting maneuvers on exclusive bicycle paths, 2001) (although these three states were not defined in that study). Figure 4-4 illustrates the three hypothesized phases of an overtaking interaction. Absolute speed difference was used in clustering following interactions and speed difference was used in clustering overtaking interactions.

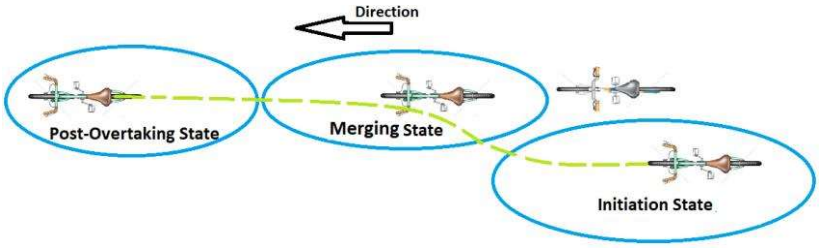

Figure 4-4: The hypothesized initiation, merging and post-overtaking phases of an overtaking interaction – blue ovals indicate the overtaking cyclist.

Intuitively, there is likely a preceding fourth state of the overtaking cyclist just before the initiation state, but an exploratory analysis of the dataset indicated no such fourth cluster in the data. For example, the longitudinal distance has a low frequency of large negative values (only 2% of the values more than -20 m), indicating that most of the captured overtaking maneuvers began in the initiation state. One reason why the data do not include a "pre-overtaking" state is the length of the captured trajectories, which limit the observation of both a pre-overtaking and post-overtaking state in the same trajectory pair. That reasoning is supported when comparing the BIC values of different numbers of clusters, as the BIC value is lower for three clusters than four clusters.

Bootstrapping was used to evaluate the uncertainty in the mixing densities and cluster distributions. The GFMM clustering method was applied individually to 1000 resamples of 90% of the original dataset, to obtain distributions of model parameters. In addition, to check the

consistency of the results, the oscillation between different states or clusters within an interaction was evaluated.

4.3 State thresholds identification

The states obtained from clustering analysis were used as input to a supervised classification problem to identify the structure and decision boundaries of perceptual variables between clusters. Decision Trees model (Breiman, Friedman, Olshen, & Stone, 1983) was selected for the classification. The tree is built by first finding the variable that best splits the data into two groups. The splitting process is applied recursively to each sub-group, until the size of the tree reaches a maximum desired value and/or the number of points in subgroups reaches a minimum value. The recursive splitting was done according to information criteria, which splits every parent node by selecting the threshold that divides the data into the child nodes that have the purest classes. The decision trees model was estimated using the *RPART* package in the statistical software R (Therneau & Atkinson, 2018). The classification algorithm first divides the data into 10 randomly selected samples. For each sample a decision tree is grown and a risk parameter for misclassification is calculated by comparing the model predicted classes and the true classes. A complexity parameter which gives the lowest risk is calculated. All the data are finally used to fit the model using the calculated complexity parameter.

4.4 Results

The clustering results for following interactions are presented in Figure 4-5 and Table 4-1. Clusters were assigned for 10422 single-frame observations within the 316 following interactions, roughly evenly split between the two clusters. Cluster 1 has longitudinal distances of -25 to -1.5 m, lateral distances of 0 to 1.2 m, and absolute speed differences of 0 to 2.2 m/s. Cluster 2 has wider ranges of all three variables: longitudinal distances with a minimum of -48 m, lateral

distances up to 3.8 m, and speed differences up to 9.1 m/s. The values of very high speed difference (9 m/s) were investigated by inspection of the video data, revealing interactions in which the following cyclist in an unconstrained state slows to observe the scenery or due to some other distraction. Also, the downgrade of 1% in one direction is another cause of observing very high speeds. Cluster 2 also has higher absolute means and larger standard deviations than cluster 1 for all three variables.

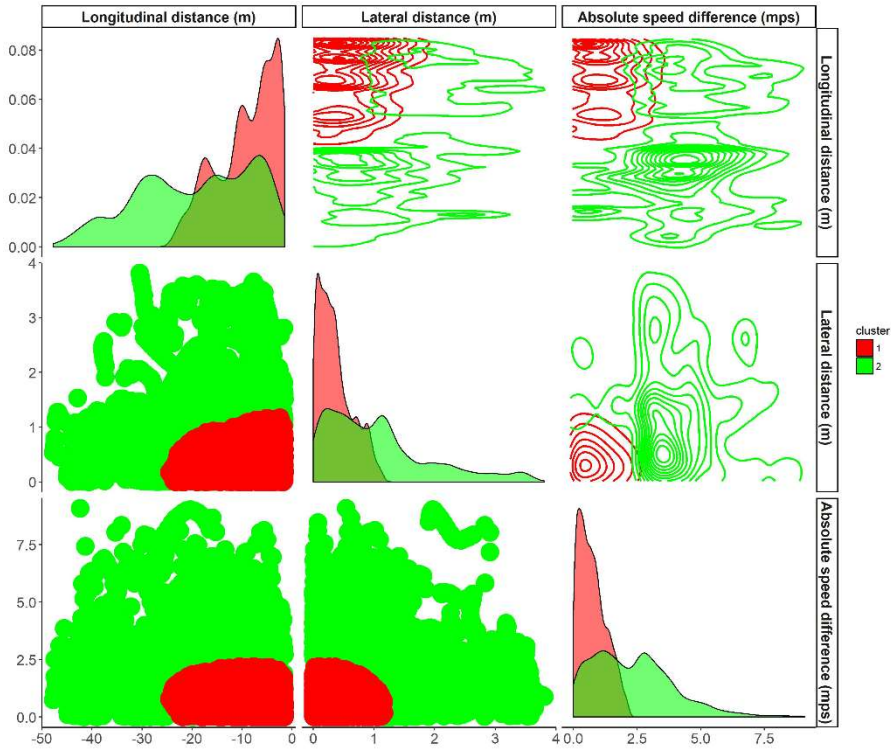

Figure 4-5: Density distributions (diagonal), cluster scatterplots (below the diagonal) and density plots (above the diagonal) for following interaction clustering results

The narrower distributions of perceptual variables in cluster 1 suggest that it represents a constrained state in a following interaction. In this state, the following cyclists have closely matched the lateral position and speed of the leading cyclist, and are in relatively close longitudinal

proximity. The univariate distributions of the clustering variables in cluster 2 are wide and include low values of all three variables, but not simultaneously. A following cyclist is unconstrained (i.e., in cluster 2) with a high value of longitudinal distance, lateral distance, or speed difference, even if the other two values are low.

Table 4-1: Summary of perceptual variables by cluster in following interactions

	Cluster 1				**Cluster 2**			
Number of observations	5117				4584			
	Mean	**Standard deviation**	**Range**		**Mean**	**Standard deviation**	**Range**	
			Min.	**Max.**			**Min.**	**Max.**
Longitudinal distance (m)	-9.27	5.97	-24.56	-1.51	-18.71	10.89	-47.75	-1.51
Lateral distance (m)	0.37	0.28	0	1.17	0.98	0.82	0	3.81
Absolute speed difference (m/s)	0.77	0.57	0	2.19	2.22	1.55	0	9.08

An example is given in Figure 4-6 of a pair of trajectories in a following interaction with clustering results. The following cyclist transitions from an unconstrained to a constrained state as the speed difference falls, and then returns to an unconstrained state as the speed difference increases again and the longitudinal distance increases as a result. The lateral distance is small throughout the interaction.

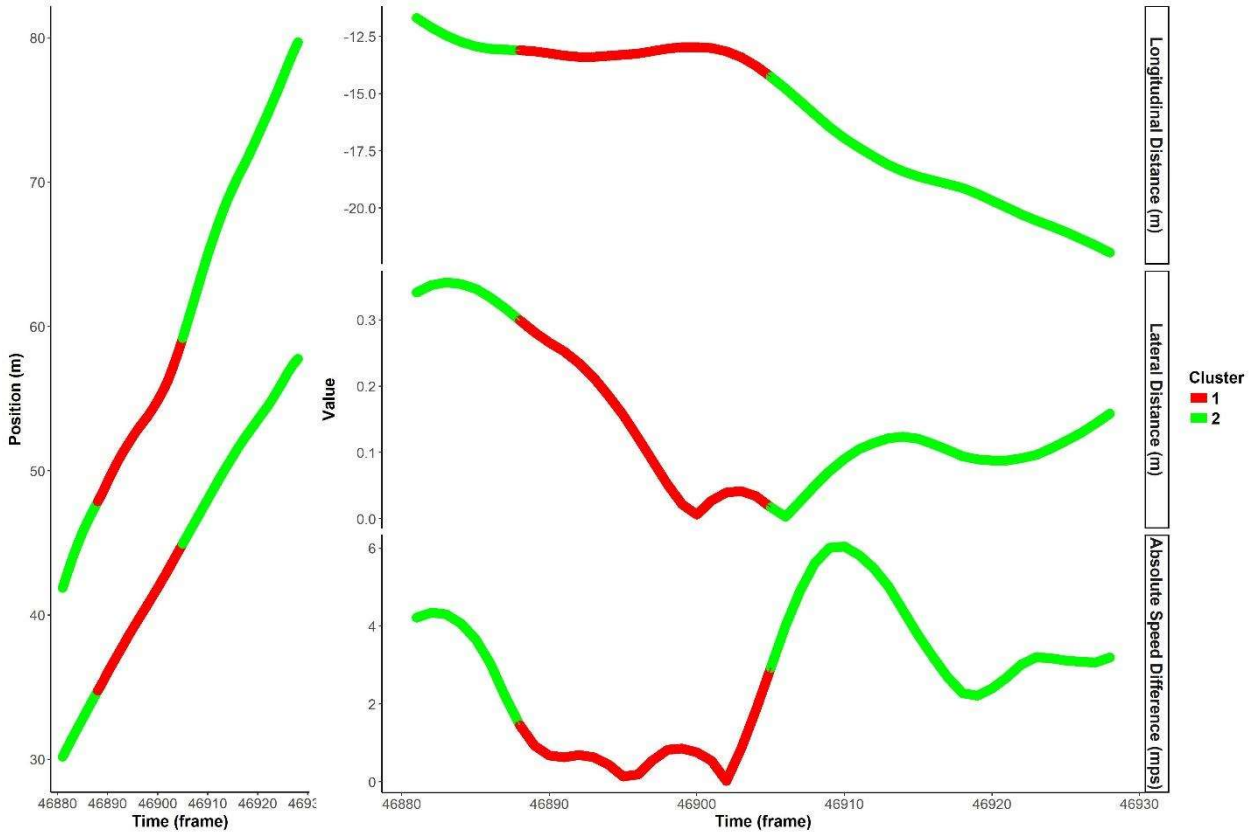

Figure 4-6: Example trajectories in a following interaction with clustering results

The clustering results for overtaking interactions are presented in Figure 4-7. Clusters were assigned for 2300 single-frame observations within the 34 overtaking interactions. Clusters 1, 2, and 3 have increasing longitudinal distances, with mean values of -5, 8, and 28 m respectively. Lateral distances are highest for cluster 1 (mean of 1.5 m), and lower for the other two clusters (means of 0.3 and 0.4 for clusters 2 and 3, respectively). The speed difference is lowest in clusters 1 and 2 (means of 1 m/s), and higher cluster 3 (mean of 5 m/s).

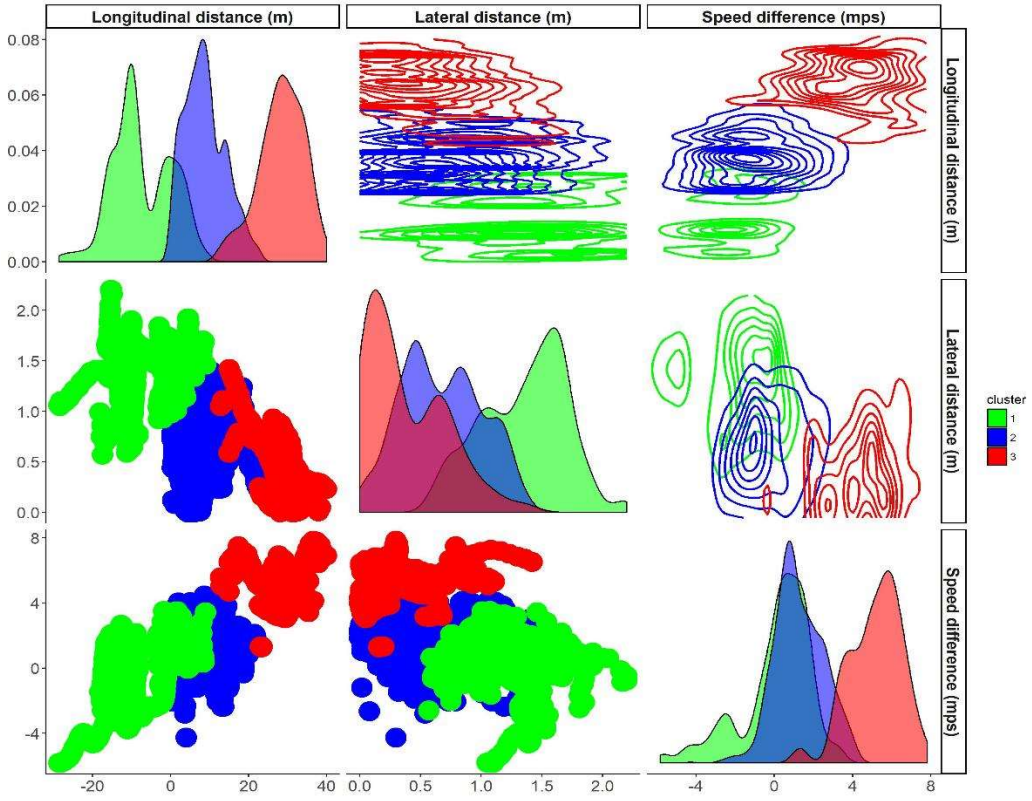

Figure 4-7: Density distributions (diagonal), cluster scatterplots (below the diagonal) and density plots (above the diagonal) for overtaking interaction clustering results

Table 4-2: Summary of perceptual variables by cluster in overtaking interactions

	Cluster 1				**Cluster 2**				**Cluster 3**			
Number of observations	630				1457				213			
	Mean	**SD**	**Range**		**Mean**	**SD**	**Range**		**Mean**	**SD**	**Range**	
			Min.	**Max.**			**Min.**	**Max.**			**Min.**	**Max.**
Longitudinal distance (m)	-5.12	7.84	-28.61	8.96	8.07	5.45	0.4	22.95	28.09	5.69	12.81	39.97
Lateral distance (m)	1.51	0.82	0.57	2.20	0.70	0.33	0	1.39	0.37	0.30	0	1.42
Speed difference (m)	0.92	1.93	-5.83	3.51	1.18	1.31	-4.27	4.45	4.82	1.64	1.26	7.82

The variable values in the three clusters are consistent with the three hypothesized phases of an overtaking interaction illustrated in Figure 4-4. Cluster 1 can be interpreted as the initiation state, in which the overtaking cyclist initiates the maneuver from an initially negative longitudinal

distance, with a positive speed difference and high lateral distance. Cluster 2 is characterized by positive but low values of longitudinal distance, lower values of lateral distances and similar speed difference, which is indicative of the merging state in which the overtaking cyclist has passed and is attempting to return to a non-interacting path. Cluster 3 can be interpreted as the post-overtaking state, in which the overtaking cyclist is far ahead of the other cyclist and has achieved their desired, higher speed, which is reflected in large positive values of longitudinal distance and speed difference, and low lateral distance.

The parallel coordinates plot given Figure 4-8 connects the three relative variable values (scaled to the variable ranges) for each observation, colored by cluster. Observations in cluster 1 have high lateral distance but low longitudinal distance and speed difference, whereas cluster 3 is the opposite and cluster 2 is a transition between them.

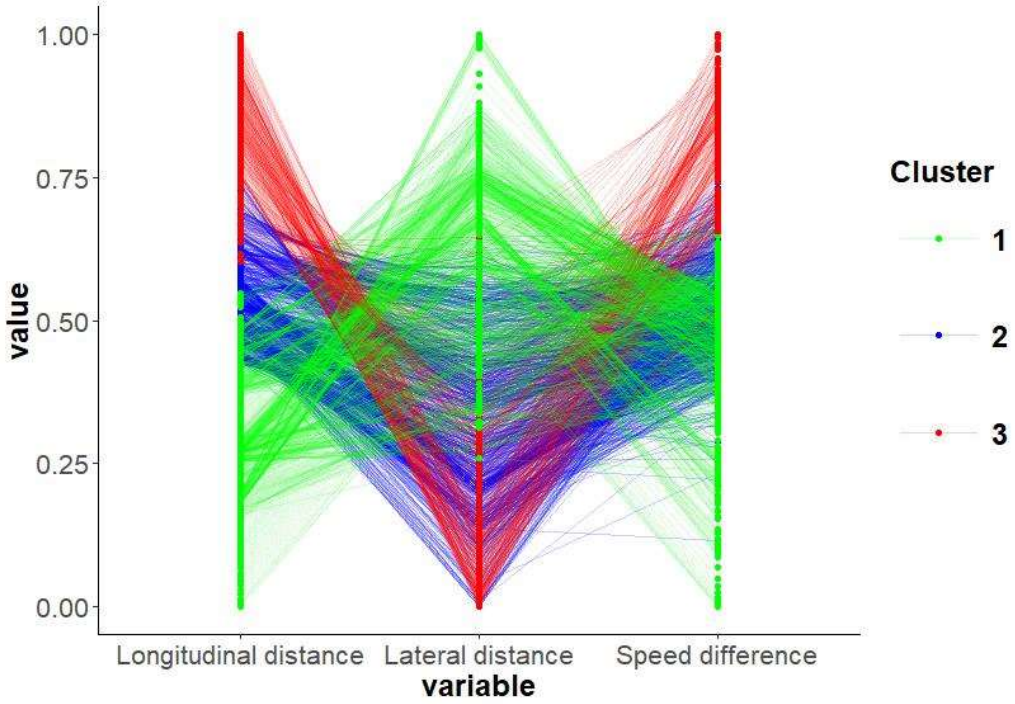

Figure 4-8: Multivariate relationships among clustering variables in overtaking interactions

Figure 4-9 gives an example of a pair of trajectories in an overtaking interaction with clustering results. The overtaking cyclist begins in an initiation state, then transitions to a merging state and finally to a post-overtaking state, consistent with the proposed labels for clusters 1, 2, and 3. The longitudinal distance increases throughout, and the lateral distance falls after completing the maneuver. The speed difference increases during the initiation state, then fluctuates, but is consistently high throughout.

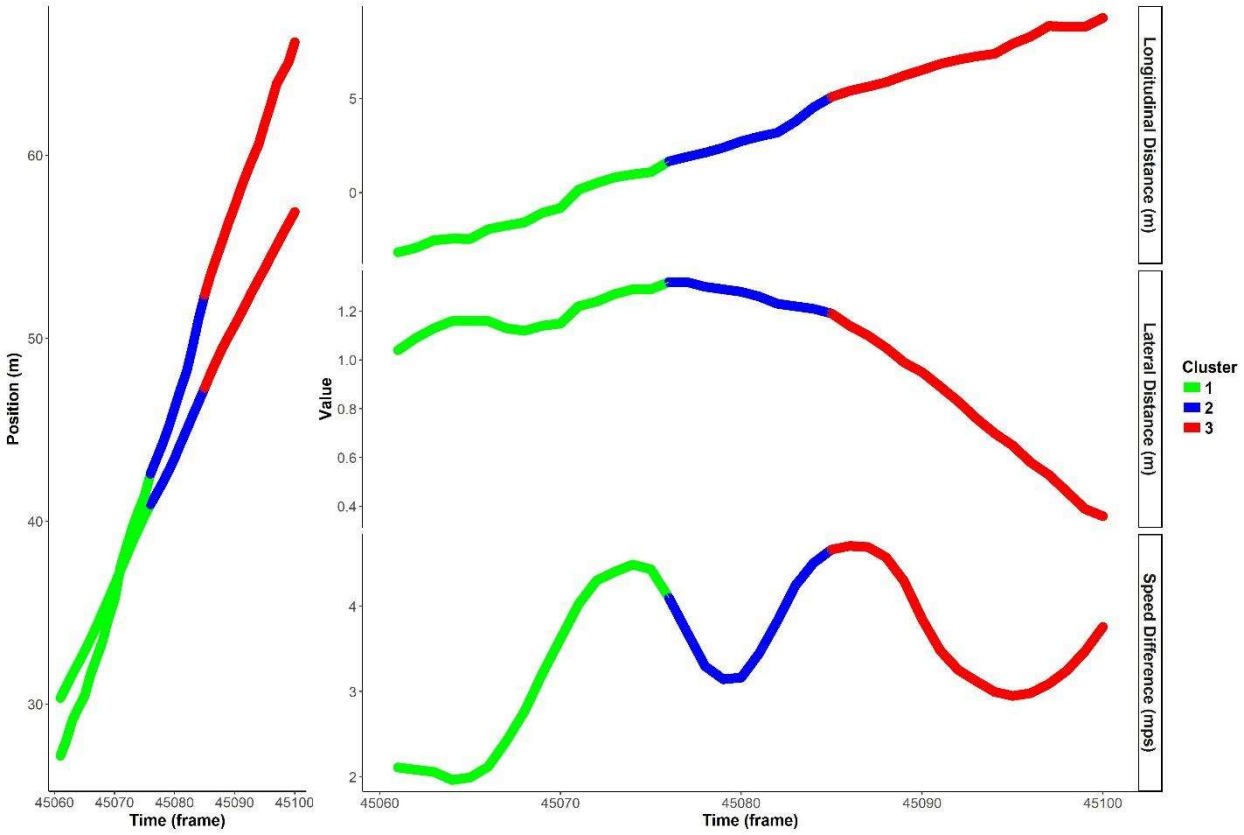

Figure 4-9: Example trajectories in an overtaking interaction with clustering results

Table 4-3 gives the results of the bootstrapping analysis for following and overtaking interactions, respectively. The distribution of variable means within clusters is low, indicating that the clustering algorithm is robust. The maximum coefficient of variation is 0.09.

Table 4-3: Bootstrapping results of following and overtaking clustering

			Longitudinal distance (m)	Lateral distance (m)	Speed difference (m/s)
Following	**Cluster 1**	**Mean**	9.13	0.37	0.76
		SD	0.15	0.01	0.01
		Range of means	(8.73; 9.72)	(0.34; 0.38)	(0.71; 0.81)
	Cluster 2	**Mean**	18.63	0.97	2.2
		SD	0.15	0.01	0.02
		Range	(18.1; 19.1)	(0.94; 1)	(2.13; 2.27)
Overtaking	**Cluster 1**	**Mean**	-5.06	1.34	0.86
		SD	0.41	0.02	0.08
		Range of means	(-5.47; -4.61)	(1.28; 1.39)	(-0.08; 1.1)
	Cluster 2	**Mean**	7.17	0.55	0.95
		SD	0.22	0.01	0.04
		Range	(6.88; 8.73)	(0.52; 0.72)	(0.83; 1.21)
	Cluster 3	**Mean**	27.94	0.4	4.98
		SD	0.49	0.02	0.23
		Range of means	(24.89; 29.39)	(0.32; 0.49)	(3.61; 5.48)

Most following interactions (57%) had zero transitions between clusters, partially due to the short duration of some of the observed interactions, while 14% had 1 transition, 14% had 2 transitions, and 15% had 3 or more transitions (maximum of 6). Most overtaking interactions (67%) had exactly 2 transitions between clusters, while 8% had 3 transitions, and 25% had 4 or more transitions (again a maximum of 6). No overtaking interactions had fewer than 2 transitions, and all overtaking interactions included all three clusters. The sequential stability of the clusters (i.e., little oscillation between clusters within an interaction), supports the robustness of the clustering approach for characterizing bicycling states of motion.

An analysis of overtaking clusters was conducted assuming four clusters to compare to the three clusters results (Figure 4-10). The results do not have an intuitive explanation, as the first cluster has a similar distribution as the clustering analysis using three clusters (Figure 4-7), but the

second and third clusters are not significantly different in longitudinal distance values, which are centered around zero. The only significant difference between the second and third clusters is the lateral distance, which does not indicate the addition of a pre-initiation state with the fourth cluster, but rather a second merging state with two different passing distances.

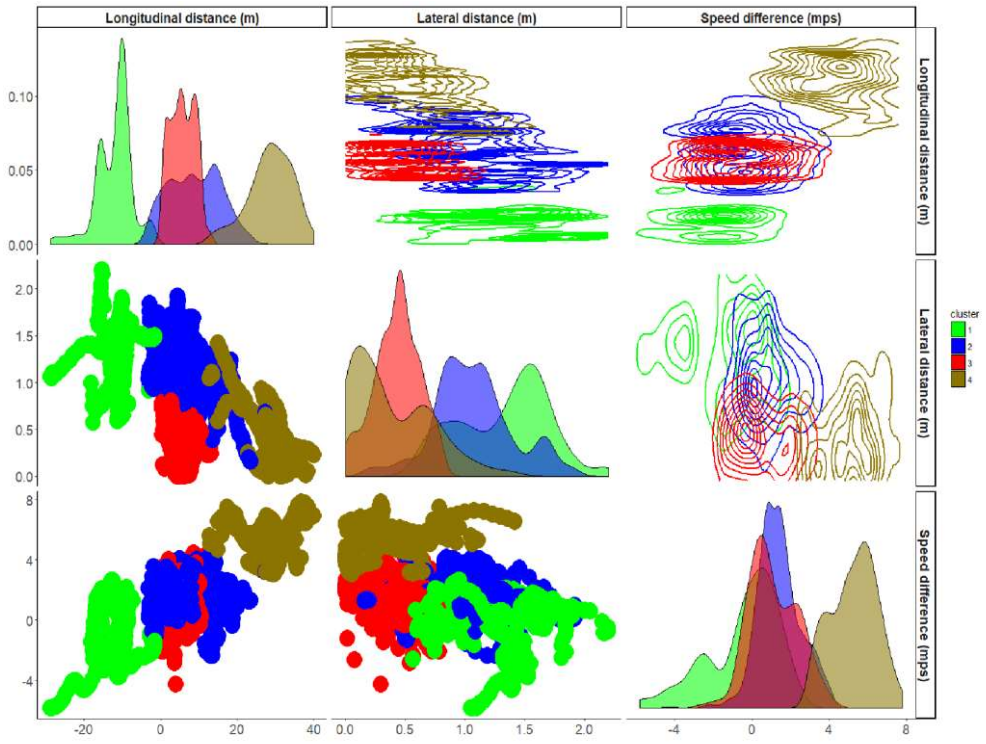

Figure 4-10: Density distributions (diagonal) and cluster scatterplots (off-diagonal) for overtaking interaction clustering results

The structure and thresholds of variables were identified by applying the decision tree model to the results of clustering analysis of the following and overtaking interactions. The results of the following interactions decision tree model are shown in Figure 4-11. The following data are initially split by the absolute speed difference at a threshold of 1.9 m/s. Data points that exceed 1.9 m/s absolute speed difference are classified into the unconstrained state (cluster 2). The remaining data are split again at a lateral distance threshold of 1 m, and then a longitudinal distance threshold

of 25 m. Thus, data points are classified into the constrained state (cluster 1) if they have an absolute speed difference below 1.9 m/s, lateral distance below 1m, and longitudinal distance less than 25 m; otherwise, they are classified as unconstrained (cluster 2).

The classification tree of overtaking states is shown in Figure 4-12. First the data points are split at a lateral distance threshold of 1.3 m. Lateral distances of 1.3 m, or more are classified as the initiation state (cluster 1). The other data points are split again at a longitudinal distance threshold of 23 m, for which longer distances are classified as the post-overtaking state (cluster 3). The next split is at a longitudinal distance threshold of -1.5 m (nearly parallel cyclists), after which observations below that threshold are split into initiation and merging states at a lateral distance of 0.58. Observations above the -1.5 m longitudinal distance threshold are split into merging and post-overtaking states at a speed difference 4.6 m/s.

The confusion matrices for the following and overtaking data classifications are given in Table 4-4. The confusion matrix is a representation of the relationship between the true and predicted classes by the model. The table shows a low misclassification error rate of 2% and 3% in following and overtaking trees, respectively.

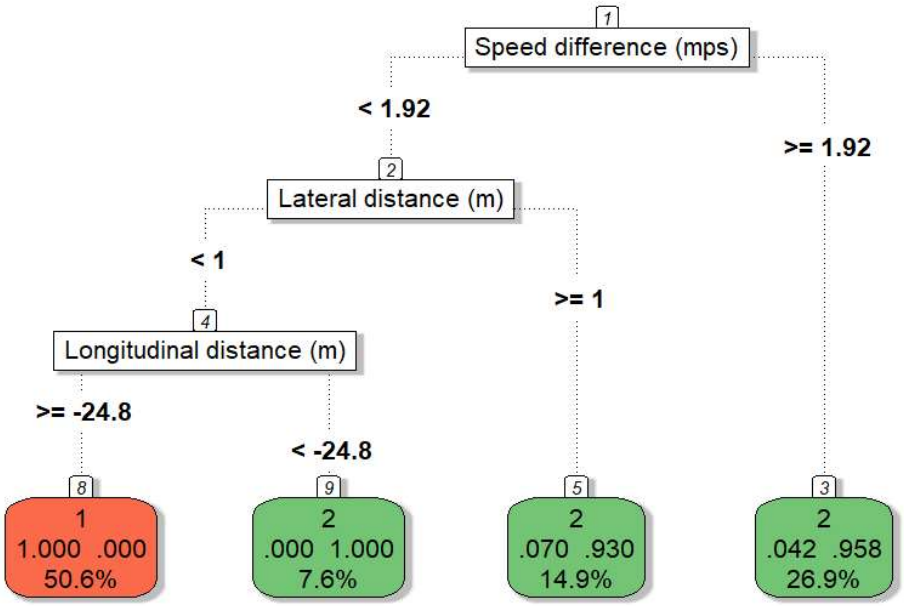

Figure 4-11: Structure and thresholds of variables in following interaction states (each final class box shows the cluster number, fraction of points correctly classified, fraction of points mis-classified, and % of total observations)

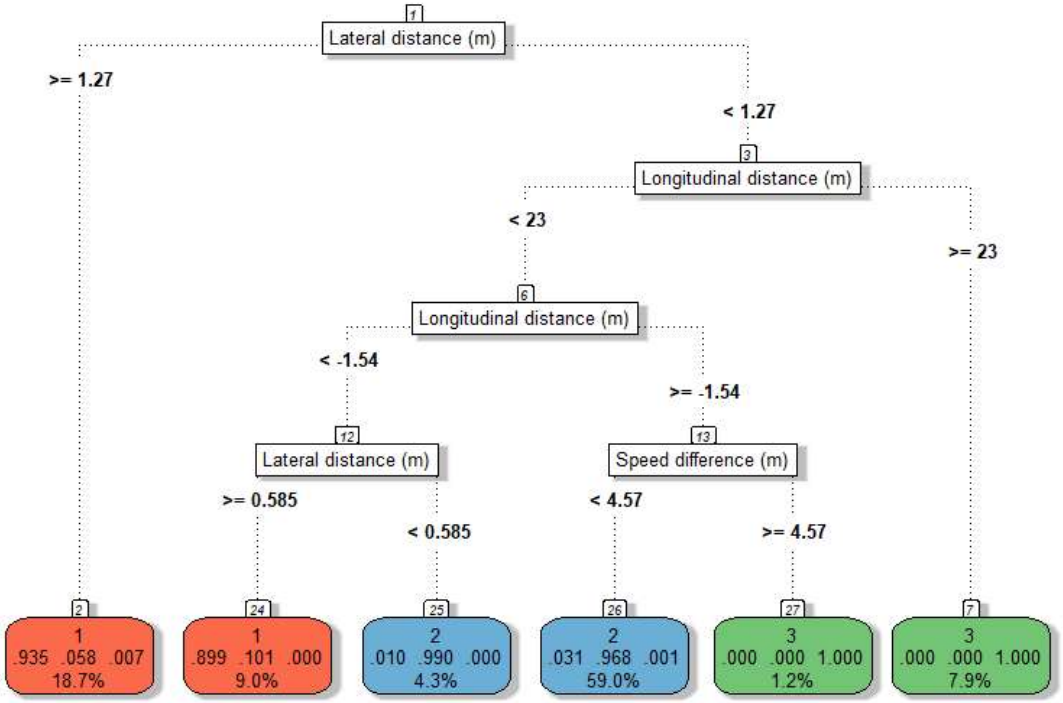

Figure 4-12: Structure and thresholds of variables in overtaking interaction states (each final class box shows the cluster number, fraction of points correctly classified, fraction of points mis-classified, and % of total observations)

Table 4-4: Confusion matrix of following and overtaking classification results

Following Confusion Matrix			
		Predicted Cluster	
		"1"	**"2"**
True Class	**"1"**	51%	0%
	"2"	2%	47%

Overtaking Confusion Matrix				
		Predicted Cluster		
		"1"	**"2"**	**"3"**
True Class	**"1"**	25%	1%	0%
	"2"	2%	62%	0%
	"3"	0%	0%	9%

Chapter 5: Microscopic modeling of cyclists on off-street paths

5.1 Conceptual and Modeling Frameworks

The decisions taken by any road user (a cyclist in this case) can be classified into strategic, tactical and operational levels (Michon, 1985) – see Figure 5-1. The strategic level describes high-level decisions about trip planning, such as mode, route, and departure time choices. The operational level describes the microscopic control decisions to operate a bicycle, predominantly pedaling and steering, in addition to finer control aspects such as balance and roll. The middle level (tactical) describes the short-term decisions required to operate the bicycle within the roadway, such as adhering to traffic controls, interacting with other road users, and avoiding obstacles. The three levels of decision-making are connected, meaning that decisions made on the strategic level affect decisions on the tactical level, which in turn affect operational-level decisions. For example, the choice of a specific route that has a right turn at some location will require a tactical decision about when to make the turning maneuver, and consequent steering and pedaling control decisions to execute the maneuver.

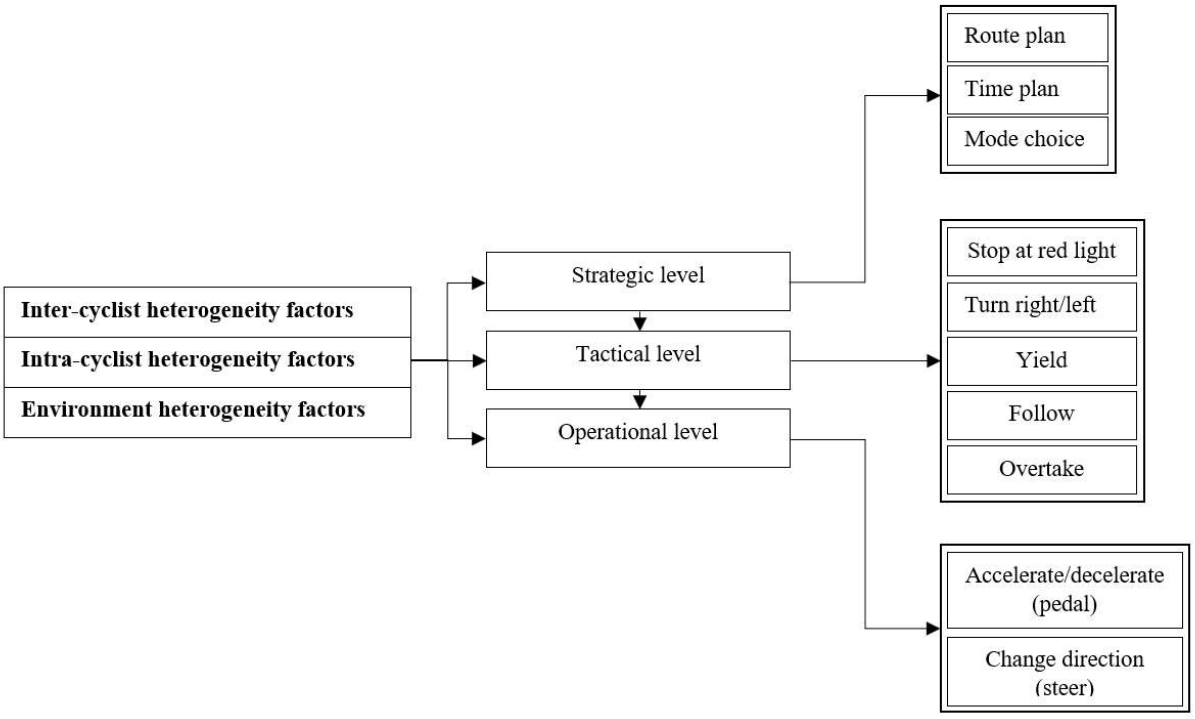

Figure 5-1: Multiple levels of cyclist on-road decision-making

Factors that vary between cyclists, such as those related to physical capability and riding equipment, lead to inter-cyclist heterogeneity in decisions. Some factors can also vary for an individual, either over the duration of a single trip or between different trips, creating intra-cyclist decision heterogeneity. Other sources of heterogeneity in decisions are due to environmental variability, such as changing traffic conditions or weather.

Focusing on operational-level decisions for microscopic simulation, we assume cyclists observe attributes of the environment (in particular the movements of other road users) and then make operational control decisions with the intention of executing a tactical-level plan. In building our conceptual framework, we need not only consider sources of heterogeneity that directly affect the operational level decision-making, but we need to consider sources of heterogeneity affecting the higher levels of decision making as well (tactical and strategic). The operational decisions are

based on future conditions anticipated by the cyclist, which may not be fully realized due to uncertain environment dynamics. The outcomes of cyclists' operational decisions that are affected by uncertain environment dynamics could change future decisions on the tactical level (deciding to overtake, keep following, keep a certain distance to the center line, etc.) or other tactical decisions that are not explicitly modeled here, such as cyclists waiting an extra block to turn right. Also, the changes in decisions on the tactical or the strategic levels affect the operational level of decisions. Although we are not explicitly modeling the effect of strategic level decisions on the operational level or vice versa, the conceptual framework is needed to scope the modeled behaviour and sources of heterogeneity.

In this paper, an agent-based stochastic modelling methodology is used to represent all three levels of heterogeneity. Cyclists are represented by agents, whose states are context-dependent; for example, states of cyclists on bicycle paths have different structures from states of cyclists on roads with mixed traffic. States of cyclists on continuous cycling paths in this paper are defined by state features of speed, position in the path, longitudinal and lateral distances from other cyclists, and speed differences from other cyclists.

The conceptual framework is reflected in the modeling framework formulation in the choice of the modeling approach itself, which allows for capturing the three levels of heterogeneity. In modeling our problem as a stochastic discrete-time continuous state and action Markov decision process. The continuous state and action spaces allow for more realistic modeling of the search space of actions taken by a cyclist facing a certain state. Also, continuous state and action spaces prevent having gaps in possible action and state variables combinations. Cyclist agents in the model map between their current state and their chosen action using a function called the policy function. This policy function uses an underlying function called the reward function

that the agent is trying to maximize. By searching the space for the best action that causes transition to a new state having the highest possible reward given the environment dynamics, the objective of the cyclist agent is being defined and the policy function is formed. This formulation is aiming to mimic the operational decision-making process of cyclists in real life by choosing actions from a continuous search space leading to change in the future state done on fine discrete intervals ($1/30^{th}$) of a second.

Two types of agents are defined in the model: unconstrained agents, whose operational decisions are not affected by the presence or movements of other road users, and constrained agents, whose decisions are affected by other road users (for example, a cyclist following another cyclist). The decisions of constrained and unconstrained agents are modeled differently. Constrained agents make decisions about their control actions by applying a learned stochastic policy to information about the environment and other agents. The policy seeks to maximize the discounted rewards of uncertain future states for the agent. Unconstrained agents apply a learned stochastic policy responding only to the environment.

5.2 Model setup

Control decisions by constrained cyclist agents are modeled as a continuous state-action discrete-time Markov Decision Process (MDP). A model that uses continuous state and action variables is crucial to avoid the exponential increase in parameter space with additional variables or finer discretization. Also, discretization of state and action spaces ignores an infinite number of overlapping state variables and action variables scenarios, which can impede accurate representation of highly detailed cycling behaviour. However, the decision was made to use discrete time instead of a continuous time scale because modeling the problem with continuous time would add an unnecessary layer of complexity. Contrary to the discretization of state and

action variables, using discrete time scale does not increase the problem dimensionality as the MDP is evaluated at each time step. Also, a high frequency of agent decisions at 1/30th of a second can be assumed as a suitable approximation of a real-world decision-making process.

An MDP consists of a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \gamma, \mathcal{R})$, where $\mathcal{S}$ is a set of states; $\mathcal{A}$ is a set of actions; $\mathcal{T} = \{P_{sa}\}$ is a set of state transition probabilities; $\gamma \in [0,1]$ is a discount factor for future rewards; and $\mathcal{R}$ is the reward function for a given state $s \in \mathcal{S}$. The cyclist agent MDP has continuous states $(s_1, \dots, s_t, \dots, s_T)^\intercal$ and continuous actions $(a_1, \dots, a_t, \dots, a_T)^\intercal$ at discrete time t in a sequence of T time steps.

Each cyclist has four state features at each time step: x coordinate, y coordinate, speed v, and direction angle Φ. The total number of state features in the model at each time step depends on the number of agents. For n cyclists existing in a certain time step (one of which will be leading or unconstrained and the others following or constrained), the complete state vector contains $n \times 4$ features: $\{(x_l, y_l, v_l, \Phi_l), (x_{f_1}, y_{f_1}, v_{f_1}, \Phi_{f_1}), (x_{f_2}, y_{f_2}, v_{f_2}, \Phi_{f_2}), \dots, (x_{f_{n-1}}, y_{f_{n-1}}, v_{f_{n-1}}, \Phi_{f_{n-1}})\}$. Each cyclist has two action features at each time step: acceleration α and yaw rate Θ (rate of change in direction angle). For n cyclists, the complete action vector is thus $\{(\alpha_l, \Theta_l), (\alpha_{f_1}, \Theta_{f_1}), (\alpha_{f_2}, \Theta_{f_2}), \dots, (\alpha_{f_{n-1}}, \Theta_{f_{n-1}})\}$. The actions of the constrained cyclists are determined by a policy derived from imitation learning, described next; the action of the unconstrained cyclist is determined by a policy derived from a variational autoencoder (described below in a subsequent subsection).

It was decided to use two separate models to predict the behaviour of constrained and unconstrained cyclists at each time-step. Constrained cyclist behaviour tends to be partially affected by the movement and actions of a leading (unconstrained) cyclist, in addition to other

variables related to the environment and individual cyclist characteristics. In contrast, the unconstrained cyclist behaviour tends to be more random in nature as cyclists have a wide range of motion variables choices (which are inherently more restricted for constrained cyclists). As a suitable approach for modeling unconstrained cyclists, GAIL was chosen as it learns the hidden stochastic policy function that is a function of relative motion variables and other environment/personal characteristics. Variational Autoencoders was found to be more suitable to modeling unconstrained cyclists as it has the ability to learn a latent distribution of the range of behavioural "styles" that differ among cyclists.

5.3 Constrained Agent Behaviour

An agent's policy is a mapping function between a state $s \in \mathcal{S}$ and the probability of taking a certain action $a \in \mathcal{A}$. The policy function can be written as $\pi_\theta(a|s)$, where θ is a set of parameters for the policy function. In imitation learning, we have a set of M trajectories or expert demonstrations $\mathcal{D} = \{\zeta_1, \dots, \zeta_m, \dots, \zeta_M\}$ that are assumed to represent optimal or sub-optimal behaviour and are used to estimate the policy and reward function. The goal of policy inference is to find the set of parameters θ that maximizes the likelihood of the observed data (i.e., the cumulative probability the model assigns to the observed trajectories). The expert demonstrations $\mathcal{D}$ are represented by an array of state and action pairs for each observed trajectory: $\zeta = (s_1, a_1, \dots, s_T, a_T)$. The likelihood of a certain trajectory (ζ) resulting from a policy function π_θ and initial state (s_1) with probability $P(s_1)$ is:

$$L_{\pi_\theta}(\zeta) = L_{\pi_\theta}(s_1, a_1, \dots, s_T, a_T) = P(s_1)\pi_\theta(a_1|s_1)\prod_{t=2}^{T}\{P(s_t|s_{t-1}, a_{t-1})\pi_\theta(a_t|s_t)\}$$

where $P(s_t|s_{t-1}, a_{t-1})$ is the state transition probability mapping from the state-action pair at one time step (s_{t-1}, a_{t-1}) to the next state s_t (estimated from the observed trajectories).

Reinforcement learning assumes that observed cyclists are following an expert policy, π_E, and that their actions seek to maximize the individual total return, which is the sum of expected discounted rewards along the trajectory. The discounted reward of a future state is represented by $\gamma^t r(s_t)$, where $r(s_t)$ is the single reward associated with visiting state s at time t. The cumulative discounted reward for a trajectory ζ is then $R_\zeta = \sum_{t \in T} \gamma^t r(s_t)$. To infer the expert policy from the observed trajectory data, an optimal policy estimation algorithm would seek to maximize the sum of discounted rewards over the observed trajectories: $\theta^* = \arg\max_\theta E_{\zeta \sim L_{\pi_E}(\zeta)}[R_\zeta]$.

The maximum entropy (ME) principle is used to represent uncertainty in the learning process (Ziebart, Maas, Bagnell, & Dey, 2008). In ME the observed trajectories are assumed to represent near-optimal behaviour; the agents are assumed to seek an optimal policy, but sub-optimal behaviour is possible. A similar representation can be used for the process of humans driving as well (Hamdar, Treiber, & Mahmassani, 2008). Relaxing the assumption that the observed trajectory strictly maximizes R_ζ, we instead assume that the probability of a trajectory increases with R_ζ. Using ME, the probability of occurrence for a specific trajectory is a function of the exponential of the aggregate reward: $P(\zeta) = \frac{\exp(R_\zeta)}{Z}$, where Z is the partition function that represents the aggregate rewards for all possible trajectories. The partition function Z is intractable to calculate, and so it is approximated using importance sampling.

Although $r(s_t)$ is unknown and unobserved, a surrogate reward $\tilde{r}(s_t)$ can be learned from the data, without imposing a structure on the reward function, using a generative adversarial imitation learning (GAIL) procedure. GAIL trains a generator G_θ to perform expert-like behaviour by rewarding it for "deceiving" a classifier or discriminator D_ψ that is trained to discriminate

between the generated and observed state-action pairs. Consider a set of simulated trajectory data ζ_θ sampled from G_θ and a set of expert trajectories ζ_E sampled from the expert dataset, $\mathcal{D}$.

For a neural network D_ψ parameterized by ψ, the GAIL discriminator loss function is given by $\mathcal{L}(D_\psi) = \underset{\zeta_E \sim \mathcal{D}}{\mathbb{E}}\left[-\log D_\psi(\zeta_E)\right] + \underset{\zeta_\theta \sim G_\theta}{\mathbb{E}}\left[\log\left(1 - D_\psi(\zeta_\theta)\right)\right]$. The first term of the discriminator's loss function is the effect of correctly classifying an expert trajectory, $D_\psi(\zeta_E)$; the second term is the effect of wrongfully classifying a generated trajectory, $D_\psi(\zeta_\theta)$. The minimum value of the loss function occurs when the discriminator becomes indifferent to trajectories sampled from expert data versus from the generator. The GAIL generator loss function is similar to that of the discriminator, but is minimized when the discriminator is confused in classifying the generated trajectories as expert ones: $\mathcal{L}(G_\theta) = \underset{\zeta_\theta \sim G_\theta}{\mathbb{E}}\left[-\log D_\psi(\zeta_\theta)\right]$.

The discriminator's classifier is trained using binary logistic regression, where $P_{\pi_E}(\zeta)$ denotes the probability of occurrence of trajectory ζ under the expert policy π_E, and $P_{\pi_\theta}(\zeta)$ denotes the probability of occurrence under the generator policy π_θ: $D_\psi(\zeta) = \frac{P_{\pi_E}(\zeta)}{P_{\pi_E}(\zeta) + P_{\pi_\theta}(\zeta)}$. Substituting the ME definition of the probability of occurrence for expert trajectories, the discriminator becomes: $D_\psi(\zeta) = \frac{\exp[R(\zeta)]}{\exp[R(\zeta)] + Z \cdot P_{\pi_\theta}(\zeta)}$. The generator's policy is modeled as a neural network with rectified linear units (ReLU) that captures non-linearity in the policy function. The policy is trained by backpropagation using minibatch gradient descent; positive examples are sampled from ζ_E and negative examples are sampled from rollouts generated by interactions of π_θ with the simulation environment.

To estimate π_θ, a surrogate reward function is formulated as: $\tilde{r}(s_t;\psi) = -\log\Big(1 - D_\psi(s_t, a_t)\Big)$. After performing the rollout with a given set of policy parameters θ, surrogate rewards $\tilde{r}(s_t;\psi)$ are calculated and trust region policy optimization (Schulman, 2015) is used to perform the policy update. Although $\tilde{r}(s_t;\psi)$ may be different from the true reward function optimized by experts, it can be used to drive π_θ into regions of the state-action space similar to those explored by π_E.

5.4 Unconstrained Agent Behaviour

A variational autoencoder (VAE) is used to model the unconstrained agent behaviour. VAE trains two neural networks, the "encoder" and the "decoder". The encoder takes as input the observed trajectories (sequence of states, s_t) of unconstrained cyclists and its output is a hidden latent representation space $\boldsymbol{z}$. The latent dimension space $\boldsymbol{z}$ is a stochastic Gaussian distribution from which samples can be drawn. The encoder process is denoted as $q_\vartheta(z|s)$, where ϑ is the encoder network parameters. The decoder process acts in reverse to the encoder, in which the input is the latent variable representation $\boldsymbol{z}$ and the output is a reconstruction of unconstrained cyclist trajectories. The decoder network is denoted as $p_\phi(s|z)$, where ϕ is the decoder network parameters. The structure of the VAE is illustrated in Figure 5-2.

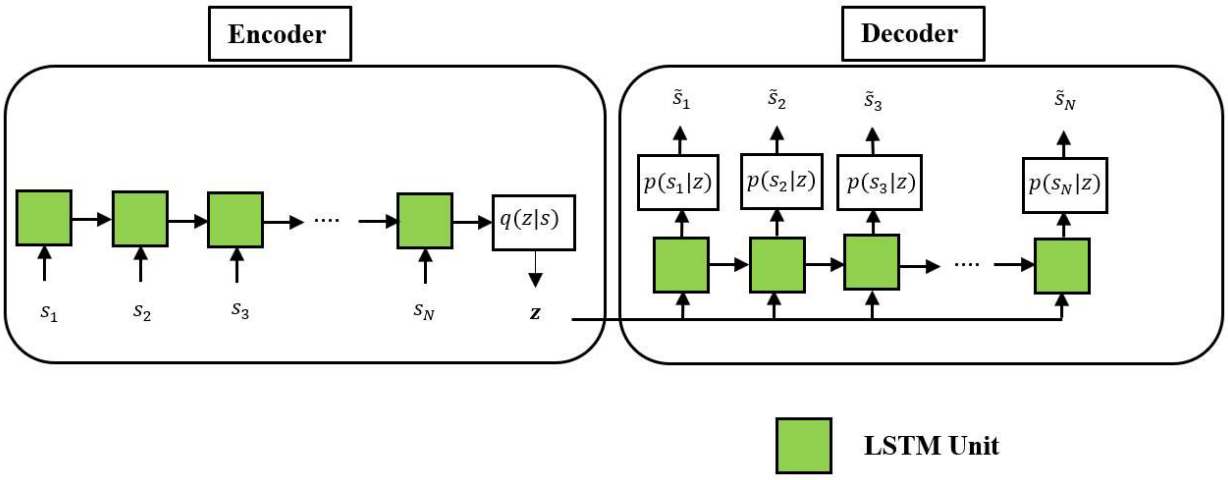

Figure 5-2: Illustration of the LSTM-variational autoencoder Architecture

The VAE encoder and decoder are both defined by long-short-term-memory (LSTM) units that are the same in number as the length of the input trajectory. This LSTM architecture allows for capturing the interdependence between sequential observations (short term) and the overall interdependence between all the points constituting the trajectory (long term). The latent variable layer contains the Gaussian distribution parameters (mean μ and standard deviation σ). The neural network is mirrored to re-estimate the observed trajectory states at each time step ($\tilde{s}_t$), with the latent variable z calculated as $z = \mu + \sigma * \epsilon,\ \epsilon \sim \mathcal{N}(0,1)$. This representation allows generation of samples from the network (after training) simply by sampling a cluster and a value for the random normal variable ϵ.

5.5 Model Testing

Cross-validation testing is undertaken using a testing dataset of 20% of the observed trajectories, randomly selected and removed from training. Simulations of the testing trajectories are made using the estimated model, with starting states equivalent to the test data. Aggregate distributions of both model variables (longitudinal distance, lateral distance, speed, speed difference, direction angle, direction angle difference, and deviation from the centerline) and

emergent variables (jerk and acceleration) are compared between simulated and observed trajectories to assess the accuracy and emergence of the model, and its ability to represent heterogeneous behaviour.

A two-dimensional Kolmogorov–Smirnov test is used to compare multidimensional parameter distributions (Fasano & Franceschini, 1987), with a 95% confidence threshold to reject the null hypothesis that the density distributions of paired state features in the simulated trajectories are drawn from the same distributions as the observed data. Kullback-Leibler divergence (KL divergence) is also used to compare bivariate distributions of state and action features between simulated and observed trajectory data. The KL divergence between observed $p(x)$ and simulated $q(x)$ data is a measure of the information lost when $q(x)$ is used to approximate $p(x)$, calculated:

$$D_{KL}(p(x)||q(x)) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)}$$

The information loss is measured in bits, a measurement unit for a distribution's entropy.

The model is also evaluated by comparison to two other cyclist simulation models from the literature, selected for their relevance to the modelling scope and representation of alternative approaches to cyclist simulation (Jiang, Hu, Wu, & Song, 2016; Zhao D. , et al., 2013). These models are implemented and calibrated to the same dataset, and then the simulated trajectories of all three models and observed cyclists are compared. The details of the comparison models, calibration, and performance are reported in the penultimate section of the paper, following the presentation of the proposed model results in the next section.

5.6 Results

The plots in Figure 5-3 give one realization of the estimated reward values over reward features (left), along with histograms of observed frequency of feature values in the dataset (right).

Higher reward values imply preferred states for cyclists, all else equal. The relationship between reward values and longitudinal distance shows a preference for longitudinal (following) distance peaking around 5 m. There is also a preference for lateral distances near zero (following directly behind the leading cyclist), and to be following toward the right versus toward the left (positive rather than negative lateral distance, at a given magnitude). The reward over varying deviation from the path centerline also clearly indicates a preference for staying on the right side of the path (as expected). A direction angle near zero (following the path) is preferred, with a preference to turn left (toward the centerline) rather than turn right (toward the right edge), based on the reward values at positive and negative direction angles of similar magnitude.

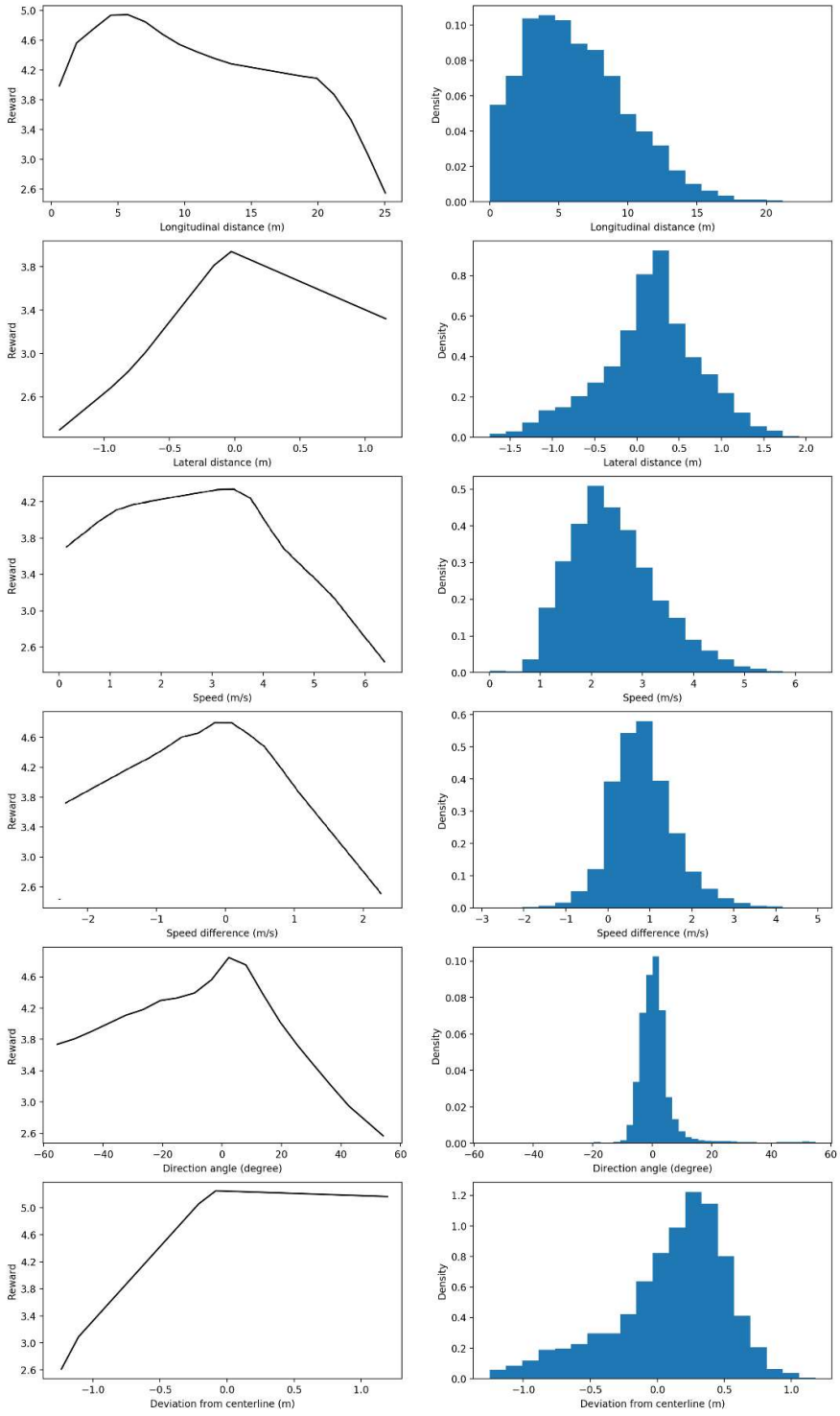

Figure 5-3: Reward function visualization with univariate reward feature distributions

Inferred speed preference peaks above 3 m/s (~12 kph), despite lower cycling speeds being more common due the hindering presence of other cyclists. Reward decreases much faster toward higher speeds than toward lower speeds, likely because the former requires more pedaling power, whereas the latter requires less and so is easier to accomplish. Cyclists also prefer to match the speed of a leading cyclist (speed difference near zero), but would prefer to be going slower (negative difference) rather than faster, which requires them to decelerate or overtake. A preference for low speed differences to leading cyclists is similar to results reported in Ma & Luo (2016). Other inferred preferences are not comparable to previous studies which only predicted longitudinal motion dynamics (Twaddle & Grigoropoulos, 2016; Zhao & Zhang, 2017).

Cross-validation distribution analysis results are given in

Table 5-1. KL divergence is low and the loss in information for the simulated versus observed data range from 5% to 14% across variables. Figure 5-4 gives observed and simulated bivariate density plots of state variables. Visual inspection reveals resemblance between the observed and simulated state variable distributions. For example, the speed difference tends to decrease to zero as longitudinal distance approaches zero. Two-dimensional Kolmogorov–Smirnov tests fail to distinguish between observed and simulated distributions at $p<0.05$ for all variable pairs.

Table 5-1: Cross-validation results comparing observed and simulated distributions.

Variable	KL divergence	Loss in information bits
Longitudinal distance	0.80	8%
Lateral distance	1.51	14%
Speed	0.96	9%
Speed difference	0.45	6%
Direction angle	1.67	15%
Direction angle difference	1.22	12%
Centerline deviation	0.38	5%
Acceleration range (per meter)	0.35	6%

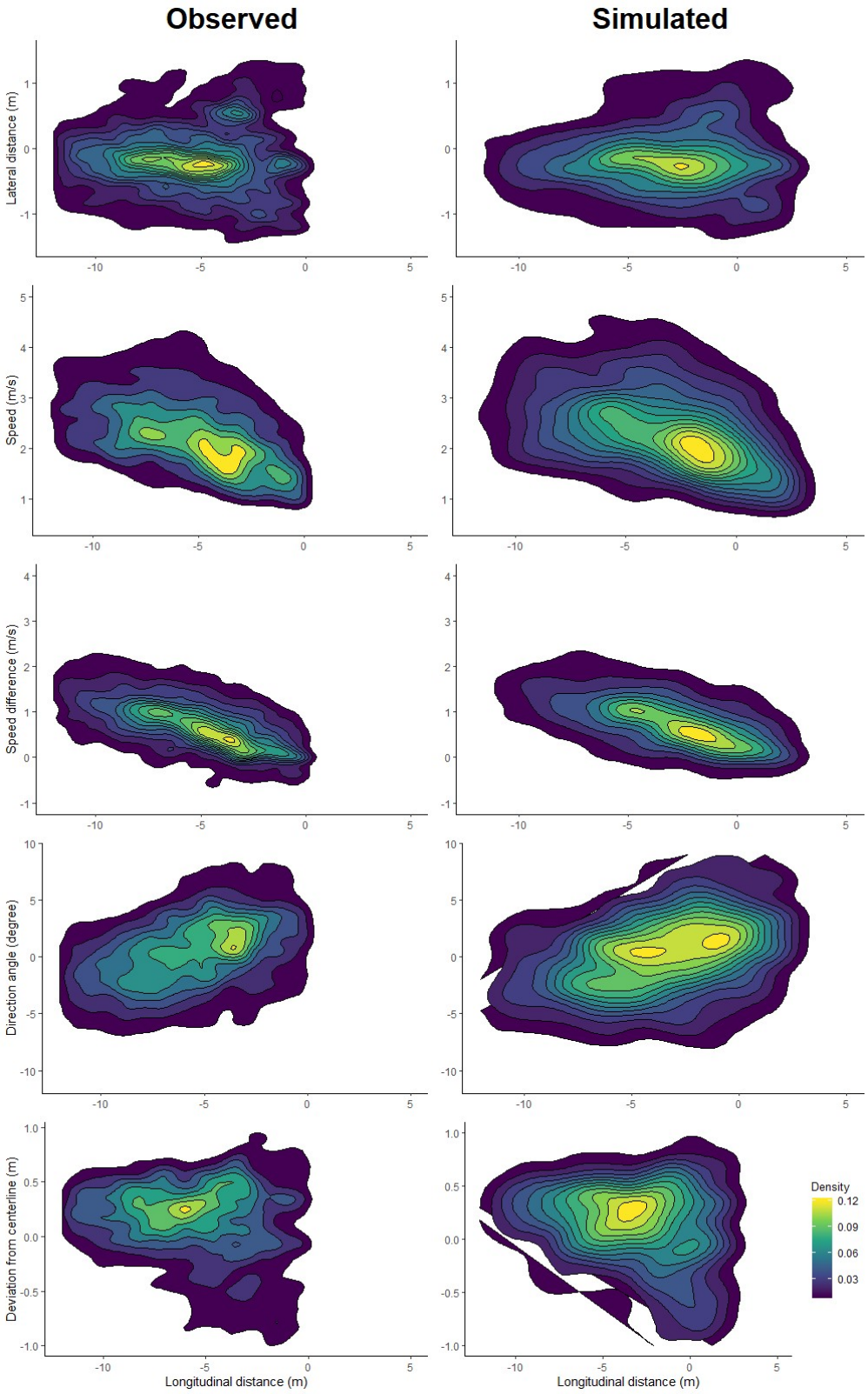

Figure 5-4: Bi-variate density plots of observed and simulated trajectories

To illustrate the modeled heterogeneity in cycling paths, Figure 5-5 shows an arbitrary observed trajectory and the distribution of 1000 trajectories simulated from the same initial state and environment dynamics. The observed trajectory lies within the simulated trajectory distribution. Variation in the simulated paths arises from the stochasticity in the agent policy and state transitions. Stochasticity in the model allows inspection of how heterogeneity varies along the path.

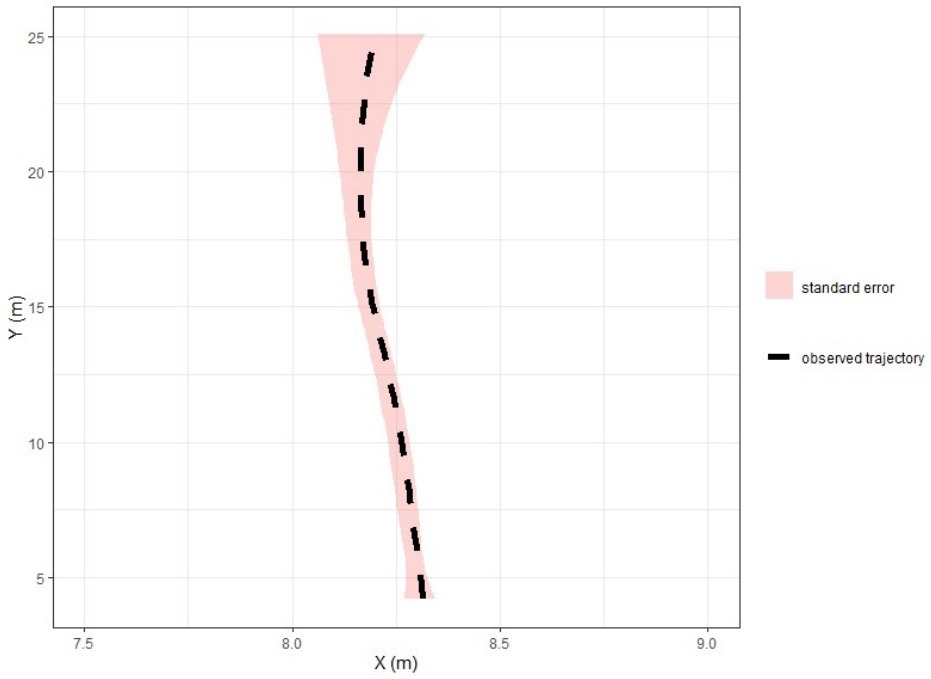

Figure 5-5: Example of the intra-cyclist variability captured by the model.

Overtaking maneuvers are the most dynamic behaviour to model on unidirectional cycling paths. Figure 5-6 gives a visual comparison of observed and simulated overtaking patterns as heatmaps of the relative locations of overtaking cyclists with respect to the leading cyclist's position. Both observed and simulated heatmaps show a higher density of overtaking maneuvers executed from the left of the originally leading cyclist. In addition, an "influence area" around the originally leading cyclist is observed as a lack of observations near the origin in both datasets, at

about ±1.6 m longitudinally and ±0.4 m laterally (increasing slightly at the overtaking position of $dy = 0$). These result shows that without explicitly including cyclist dimensions in the model, the agents faithfully reproduce the operating space of observed cyclists.

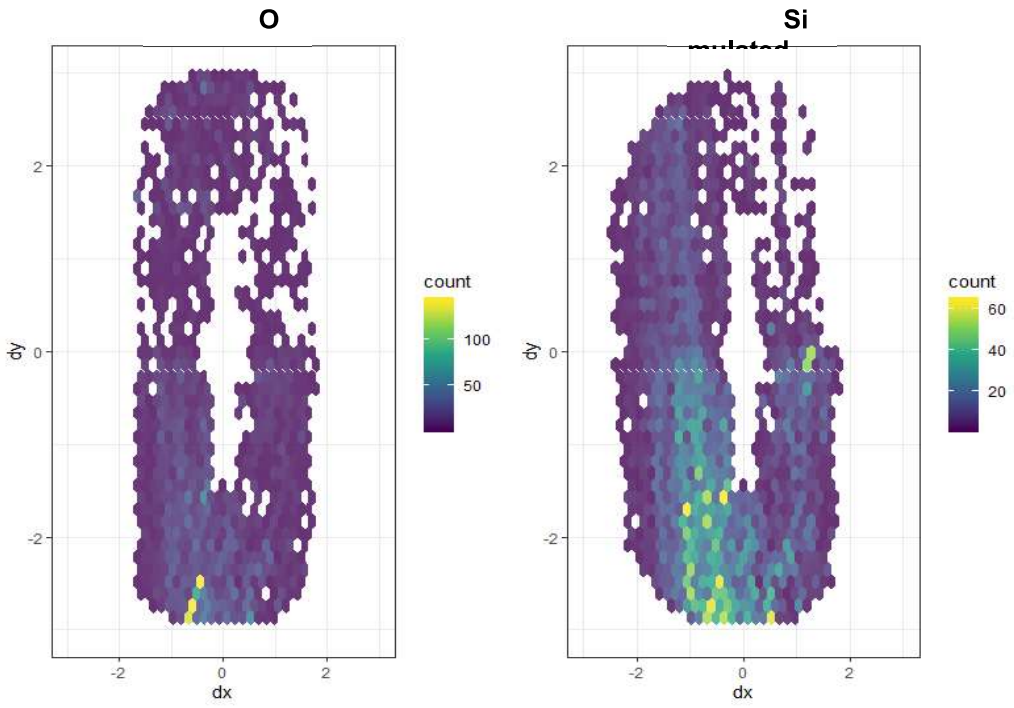

Figure 5-6: Locations of overtaking cyclists in observed and simulated data (in m, with respect to the cyclist being overtaken at $(dx, dy) = (0,0)$)

To get a sense of how the observed behaviour during overtaking maneuvers compares to literature. The results of this study are compared against (Khan & Raksuntorn, 2001), in which motion variables from video data of cyclists in following and overtaking interactions are extracted. In that study, the average lateral distance during overtaking was 1.78 m compared to about 1 m in our study (Figure 5-6). The difference could be due to the difference in cycling path width (3 m in Khan & Raksuntorn, 2001 study versus 2.2 m in our study).

Variable density functions in Figure 5-4 are compared against a similar study (Gavriilidou, Daamen, Yua, & Hoogendoorn, 2019) that attempted to follow an analytical approach to assess how the preferences of cyclists towards making certain actions changes with changing relative motion variables. In that study preference for lower deviation from the centerline was observed from the data and implemented in the model, which resembles what our model predicts (Figure 5-4). Also, in a recent study (Paulsen, Rasmussen, & Ank, 2019), in which the authors were trying to establish the effect of congestion of bicycle traffic on their microscopic behaviour, it was observed that cyclists prefer to move at higher speeds and speed differences when their longitudinal distances are higher, which can be seen in Figure 5-4.

5.7 Comparison to other models

To further assess the performance and benefit of the imitation learning model presented in this paper, comparisons are made to two other cyclist simulation models from the literature. These two models were selected because they have similar scope to the proposed model and represent alternative state-of-the-art microsimulation methods. The first model (Jiang, Hu, Wu, & Song, 2016) is based on the well-known cellular automata (CA) approach developed by Nagel and Schreckenberg (2002). The model was originally developed from an experiment in which multiple cyclists navigated an oval course. The second model (Zhao D. , et al., 2013) also uses the CA approach to model bicycle following and overtaking interactions, developed from video data of a separated bicycle path. Unlike the first model which only represents unidimensional following behaviour, the second model simulates both longitudinal and lateral movements.

The parameters of both comparison models were calibrated to the study dataset. The Jiang et al. (2016) model predicts the longitudinal position of a cyclist at each 1-second time step using the instantaneous speed. The instantaneous speed is updated each time step, influenced by the

longitudinal distances between successive cyclists, with a randomization parameter. The calibration parameters are the maximum cyclist speed v_{max}, the operating distance between cyclists d_{op}, and a third parameter d_c reflecting the effect of a leading cyclist's distance on the following cyclist's decisions. The Zhao et al. (2013) model updates both longitudinal and lateral positions of cyclists at each time step from instantaneous longitudinal and lateral speeds, which are influenced by the relative longitudinal and lateral positions of the following and leading cyclists. The model also has a randomization component to the speed updating equations. There is just one parameter to calibrate, which is the maximum speed v_{max}.

For this comparison, the parameters in the Jiang et al. (2016) model were calibrated to the study path data set by enumerating values in appropriate ranges of the calibration parameters and then calculating the root mean squared error (RMSE) between the x- and y- locations (m) in the simulated vs. observed data for each combination of parameter values. The calibration parameter values that yielded the lowest RMSE were $v_{max} = 5.2$ m/s, $d_{op} = 6$ m, and $d_c = 2$ m. The selected value of the randomization parameter p was 0.8, based on the original paper. The parameter of the Zhao et al. (2013) model was calibrated similarly with the result of a calibration value of $v_{max} = 4.8$ m/s. The selected value of the randomization parameter p for the Zhao et al. model was 0.45, based on the original paper.

Figure 5-7 gives the distributions of motion variables for the observed data and for the simulated output of the imitation learning model proposed in this paper and the two calibrated comparison models (at 1-second intervals, $N = 5645$). The distributions of longitudinal distance, speed, speed difference and lateral distance from each model are based on a single realization (simulation run) for all interactions in the study data set. Overall, the imitation learning model simulates the observed distributions of motion variables more accurately than the comparison

models. The Zhao et al. model has the poorest performance for longitudinal distance, and the Jiang et al. model has the poorest performance for speed. Both comparison models poorly represent the observed speed differences between leading and following cyclists. The Zhao et al. model produces too wide a range of lateral distances, while the Jiang et al. model does not represent the lateral dimension at all.

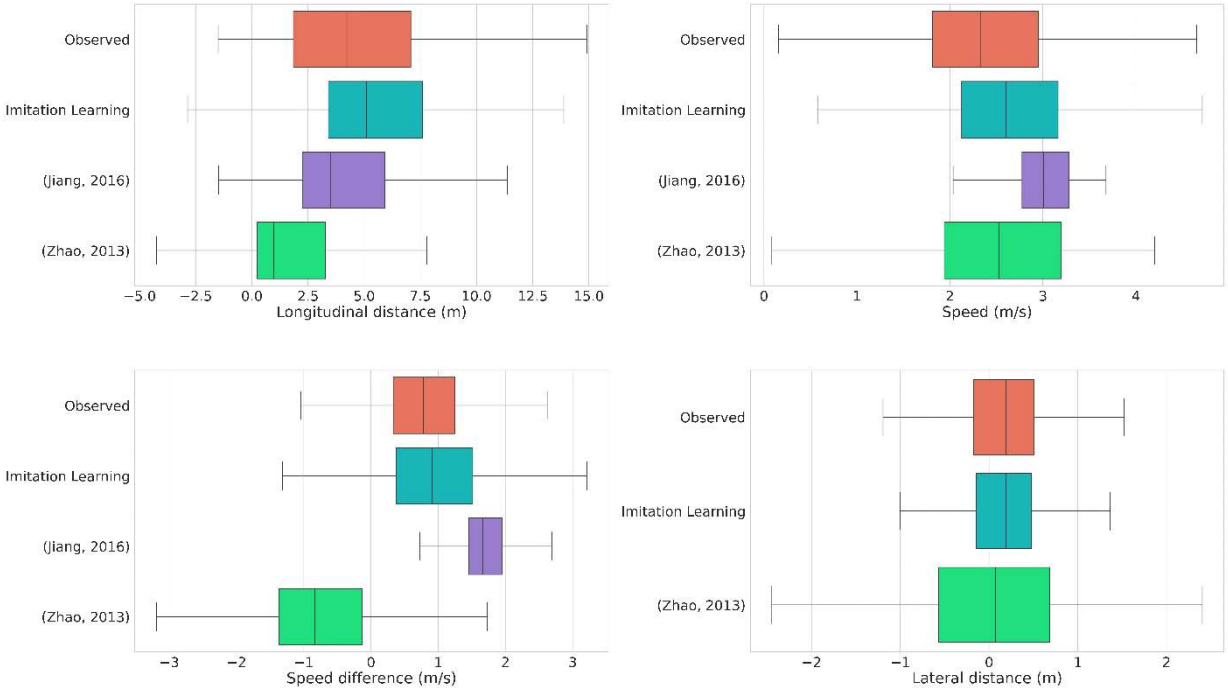

Figure 5-7: Motion variable distributions from observation and the three simulation models

In addition to these differences in model performance, there are conceptual advantages of the proposed imitation learning model. First, the proposed model represents the behaviour of cyclists in unconstrained and constrained states, and allows flexible transitions between those states. In contrast, the comparison models do not explicitly model unconstrained behaviour and instead assume constant or randomly fluctuating speed for “free-flow” cyclists, which further affects the simulated behaviour of the following or overtaking cyclists. The imitation learning

model also explicitly represents heterogeneity in both the environment and cyclist behaviour. In contrast, the comparison models introduce stochasticity simply with a random parameter in the speed updating process, which cannot represent the systematic differences in physical, psychological and other characteristics of cyclists.

To compare the stochastic performance of each model, 120 simulation runs for the entire study data set were executed using all three models (imitation learning, Jiang et al., and Zhao et al.). Figure 5-8 gives the resulting performance in terms of the mean absolute error (MAE) of the longitudinal distance variable for each simulated trajectory, compared to the observed data. Each point in the figure gives the MAE for a single model run on the entire dataset (i.e., the distributions in Figure 5-7 would yield a single point in each column); the jitter plot randomly positions points in the horizontal direction to prevent overlapping and improve visual interpretability. The imitation learning model has the most consistently high precision, with an MAE under 3 m for all 120 runs, compared to almost 6 m for the Jiang et al. model and almost 8 m for the Zhao et al. model. The difference in stochastic performance may be due to an overly simple representation of heterogeneity in the comparison models.

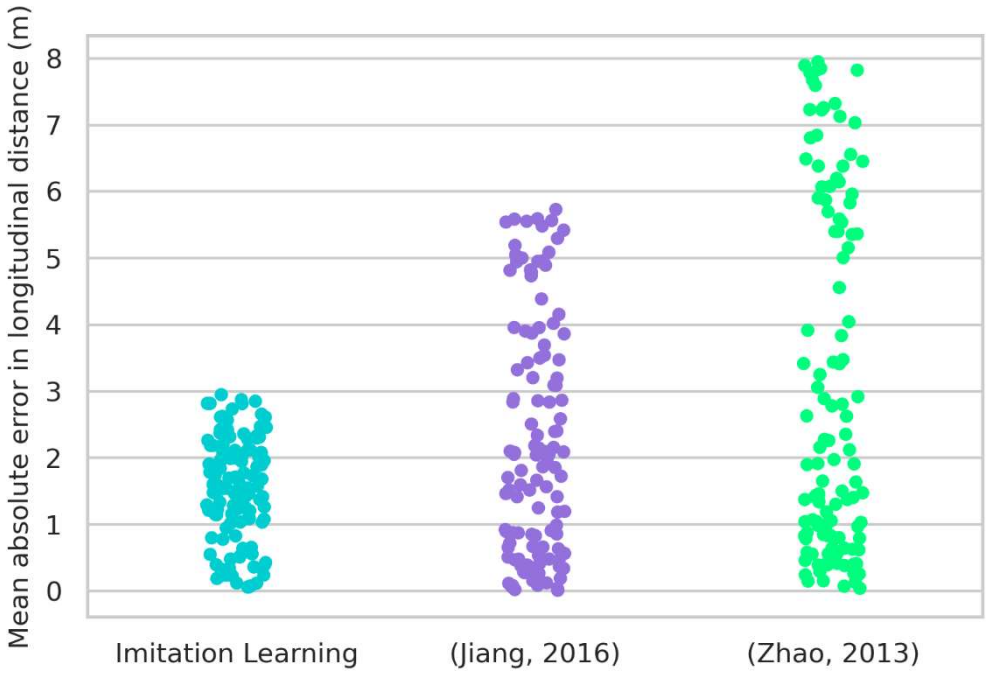

Figure 5-8: Distributions of MAE over 120 runs of the entire data set by each model

Chapter 6: Conclusion

6.1 Summary of findings and contributions

The results of the work presented in this thesis indicate that the threshold of longitudinal distance between constrained and unconstrained states in following interactions is 25 m, which equates to a 5 s time headway at 5 m/s average speed. This finding is supported by (Hoogendoorn & Daamen, 2016), in which the threshold between constrained and unconstrained cycling was estimated to be 4 s. In our study, the constrained longitudinal distance ranges from 1.5 m to 24 m with an average of 9 m; the minimum longitudinal distance is supported by the value measured in (Zhao & Zhang, 2017) of 2.0. The threshold of lateral distance between constrained and unconstrained states is 1 m, slightly higher than the value set in (Hoogendoorn & Daamen, 2016) of 0.7 m, based on the width of cyclist handlebars plus an additional shy-away distance. In overtaking interactions, the average speed difference in initiation state is 1.2 m/s, which is lower than the values measured in (Khan & Raksuntorn, 2001) and (Falkenberg, et al., 2003) of 2.6 and 3.6 m/s, respectively. In addition to using different samples, those two studies used subjective methods to determine the initiation stage of the overtaking interaction.

The sample size used in the analysis of overtaking interaction of 34 overtaking interactions with 2300 observation can be argued to be small. However, there is no clear standard in the literature for calculating the minimum sample size for clustering analysis. The minimum sample size should be correlated with the dimensionality of the problem (Dolnicar, 2002). Past studies provide methods for assessing the sample size in clustering analysis, such as (Formann, 1984), which mentions a rule to assess the number of clusters according to the number of variables used in the clustering as the minimum number of observations to be ($5*2^m$), where m is the number of

variables. According to this rule, the sample size used in this study is considered sufficient. In any study, more conclusive results could be obtained using a larger sample size.

The thesis proposes and tests the applicability of generative adversarial imitation learning for estimating continuous, non-linear, and stochastic policies and reward functions for cyclists in an agent-based simulation model. In addition, these advancements, the model is novel in predicting both longitudinal and lateral motion dynamics of interacting cyclists. Cross-validation of the simulation model indicates realistic generation of modeled and emergent variable distributions. The proposed model also outperforms two cyclist simulation models from the literature representative of state-of-the-art alternative approaches.

Another core contribution of this thesis is the development of the novel approach to agent policy estimation, with greater flexibility and fewer *a priori* assumptions than existing approaches. This research aims to advance microscopic modeling of cyclists by following an intelligent agent approach with a non-linear stochastic decision-making mechanism. The model differentiates between two distinct types of cyclists (constrained and unconstrained). Constrained cyclist behaviour and decision making are dependent on cyclist personal characteristics, environment variables, motion variables of the constrained (following) cyclist, and relative motion variables between the following and the leading cyclist. Unconstrained cyclist behaviour is only dependent on the cyclist's personal characteristics, environment variables, and own motion variables. The paper follows a generative modeling approach (as opposed to discriminative approaches) in modeling interacting cyclists' trajectories. Generative models are concerned with learning the internal multivariate distributions of model variables, while discriminative approaches identify the boundaries between behaviour schemes.

6.2 Limitations

There are several limitations to this study. The method used for clustering needs to be tested with other datasets from different contexts such as different facility types, road grades and cities. The trajectories extracted from the video data were short in length, so some interactions could have started in the field of vision of the video scenes and completed out of it. There are only two locations in the video data. Validation of the variable distributions and thresholds in different states is required to test the applicability and transferability to other locations. Other types of interactions such as between bi-directional cyclists and between cyclists and pedestrians should also be investigated. Cyclist behaviour in groups was not considered in following interactions. More research in cyclists' group following behaviour can reveal further behavioural characteristics. In addition, trajectories could not be extracted from the video data for 100% of cyclists, which might have biased the sample if the tracking failures were non-random.

Another key limitation in the presented work is model testing in a single case study location. The model application used short-distance observed trajectory data (around 25 m), which precluded analysis of longer spatial patterns. It is a challenge to collect detailed trajectory data over long distances from existing computer vision techniques. This limitation could be addressed in future research by creating a longer-coverage dataset from two or more synchronized cameras. Despite this limitation in the observed length of trajectories, the imitation learning approach that uses memoryless MDP representation allows for making inferences about behaviour even from short sequences of observations. The memoryless property means that agents use the estimated policies to take actions only based on the current state. Also, the large number of trajectories allows for observing wide distributions of motion variables that are used in the model (Figure 5-3).

Also, a limitation related to the video data is that it may not contain much variability in some of the factors expected to influence cyclist behaviour. Examples include cyclist socio-demographics, variables related to cycling effort such as road grade, wind, and mass, as well as factors relating to trip purpose and length, weather, and time of day. In this study, these factors would likely manifest in the variability of the stochastic policy estimation.

6.3 Future work

Future research can apply the distributions and thresholds of variables in different states to develop interaction rules specific to each state. These rules can describe the predicted change in acceleration and direction in each time step based on the perceptual variables. Other variables such as accelerations and yaw rates could also be tested for inclusion in clustering. Modeling other types of interactions will also be required for the development of improved bicycle microsimulation models. Other essential future research tasks include applying and testing the model in other locations, ideally with longer trajectories and augmented cyclist attributes (both of which would require enhanced data collection methods). Applying the model in other locations will allow further evaluation of different levels of model uncertainty and generalizability, with investigation of variability at the cyclist and location levels. Due to the unique nature of agent-based models, pattern-based validation methods should also be considered for further model validation. Finally, there are several key paths for future model development: most importantly extension to two-way paths and additional agent types representing pedestrians and other "multi-use path" users such as people on skateboards, e-scooters, and other devices. By enhancing the representation of cyclist heterogeneity in agent-based simulations, we believe that the proposed approach is a significant

methodological advancement toward more nuanced and sensitive analysis tools in active transportation.

Bibsliography

Abbeel, P., & Ng, A. (2004). Apprenticeship learning via inverse reinforcement learning. *Proceedings of the twenty-first international conference on Machine learning.*

Abbeel, P., & Ng, A. (2004). Apprenticeship learning via inverse reinforcement learning. *Proceedings of the twenty-first international conference on Machine learning.*

Abdou, M., Hamill, L., & Gilbert, N. (2012). *Designing and building an agent-based model." Agent-based models of geographical systems.* Dordrecht: Springer.

Aghabayk, K., Sarvi, M., & William , Y. (2015). A state-of-the-art review of car-following models with particular considerations of heavy vehicles. *Transport reviews, 1*(35), 82-105.

Aghabayk, K., Sarvi, M., & William, Y. (2015). A state-of-the-art review of car-following models with particular considerations of heavy vehicles. *Transport reviews, 1*(35), 82-105.

Ahmed, K. I. (1999). *Modeling drivers' acceleration and lane changing behaviour.* Massachusetts Institute of Technology.

Ahmed, K. I., & Ben-Akiva, M. (1996). Models of freeway lane changing and gap acceptance behaviour. *Transportation and traffic theory*(13), 501-515.

(2005). *AIMSUN User's Manual.* TSS-Transport Simulation System.

Bain, M., & Sammut, C. (1995). A framework for behavioural cloning. *Machine intelligence agents*.

Banfield, J. D., & Raftery, A. E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics*(49), 803-821.

Baster, B., Duda, J., Maciol, A., & Rebiasz, B. (2013). Rule-based approach to human-like decision simulating in agent-based modeling and simulation. *System Theory, Control and Computing (ICSTCC) 17th International Conference.*

Biagiotti, L., & Melchiorri, C. (2008). *Trajectory planning for automatic machines and robots.* Springer Science & Business Media.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1983). *Classification and regression trees.* Wadsworth, Belmont, California.

Buckley, D. J. (1968). A semi-poisson model of traffic flow. *Transportation Science, 2*(2), 107-132.

Carrignon, D. (2009). Assessment of the Impact of Cyclists on Heterogeneous Traffic. *TEC Magazine*, 323-325.

Celeux, G., Martin, O., & Lavergne, C. (2005). Mixture of linear mixed models for clustering gene expression profiles from repeated microarray experiments. *Statistical Modelling , 3*(5), 243-267.

Chandler, R. E., Herman, R., & Montroll, E. W. (1958). Traffic dynamics: studies in car following. *Operations research, 2*(6), 165-184.

Chopard, B. (2012). *Cellular automata modeling of physical systems.* New York: Springer.

COWI. (2012). *Micro Simulation of Cyclists in Peak Hour Traffc.* Copenhagen.

de Oliveira, Í. (2017). Analyzing the performance of distributed conflict resolution among autonomous vehicles. *Transportation Research Part B: Methodological, 96*, 92-112.

Djavadian, S., & Joseph, C. (2017). An agent-based day-to-day adjustment process for modeling 'Mobility as a Service'with a two-sided flexible transport market. *Transportation research part B: methodological, 104*, 36-57.

Dolnicar, S. (2002). A review of unquestioned standards in using cluster analysis for data-driven market segmentation. *Australian and New Zealand marketing academy conference (ANZMAC).* Melbourne.

Everitt, B. S., & Hand , D. J. (1981). *Finite mixture distributions.* London: Chapman and Hall CRC.

Falkenberg, G., Bonfranchi, T., Cosse, L., Draeger, W., Vortisch, P., Kautzsch, L., . . . Zimmermann, A. (2003). Bemessung von Radverkehrsanlagen unter. *Unterreihe Verkehrstechnik.*

Fasano, G., & Franceschini, A. (1987). A multidimensional version of the Kolmogorov–Smirnov test. *Monthly Notices of the Royal Astronomical Society, 225*(1), 155-170.

Fellendorf, M., & Vortisch, P. (2001). Validation of the microscopic traffic flow model VISSIM in different real-world situations. *Transportation Research Board 80th annual meeting.* Washington, D.C.

Formann, A. K. (1984). *Die latent-class-analyse: Einführung in Theorie und Anwendung.* Beltz, Weinheim.

Garrido, S., Moreno, L., & Abderrahim, M. (2006). Path Planning for Mobile Robot Navigation using Voronoi Diagram and Fast Marching. *International Conference on Intelligent Robots and Systems* . Beijing, China.

Gavriilidou, A., Daamen, W., Yua, Y., & Hoogendoorn, S. (2019). Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour. *Transportation research part C: emerging technologies*, 468-484.

Gazis, D., Herman, R., & Rothery, R. W. (1961). Nonlinear follow-the-leader models of traffic flow. *Operations research, 4*(9), 545-567.

Gipps, P. G. (1981). A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological, 2*(15), 105-11.

Götschi, T., Garrard, J., & Giles-Corti, B. (2016). Cycling as a part of daily life: a review of health perspectives. *Transport Reviews, 36*(1), 45-71.

Gould, G., & Karner, A. (2009). Modeling bicycle facility operation: Cellular automaton approach. *Modeling bicycle facility operation: Cellular automaton approach, 2140*, 157-164.

Gujarati, D. N. (2003). *Basic econometrics.* Mc-Graw Hill.

Guo, Y., Logan, H. L., Glueck, D. H., & Muller, K. E. (2013). Selecting a sample size for studies with repeated measures. *BMC medical research methodology, 13*(1).

Hamdar, S., Treiber, M., & Mahmassani, H. (2008). Modeling driver behaviour as sequential risk-taking task. *Transportation research record*, 208-217.

Heinen, E., Bert, V., & Kees, M. (2010). Commuting by bicycle: an overview of the literature. *Transport reviews, 30*(1), 59-96.

Heinen, E., Bert, V., & Kees, M. (2010). Commuting by bicycle: an overview of the literature. *Transport reviews, 1*(30), 59-96.

Hoogendoorn, S., & Daamen, W. (2016). Bicycle headway modeling and its applications. *Transportation Research Record : Journal of the Transportation Research Board 2587*, 34-40.

Hoogendoorn, S., & Daamen, W. (2016). Bicycle headway modeling and its applications. *Transportation Research Record : Journal of the Transportation Research Board*, 34-40.

Hussein, M., & Sayed, T. (2017). A bi-directional agent-based pedestrian microscopic model. *Transportmetrica A: Transport Science, 13*(4), 326-355.

Ismail, K., Sayed, T., & Saunier, N. (2008). A methodology for precise camera calibration for data collection applications in urban traffic scenes. *Canadian Journal of Civil Engineering, 40*(1), 96-104.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical review, 106*(4).

Jennings, N. (2000). On agent-based software engineering. *Artificial intelligence, 2*(117), 277-296.

Jennings, N. (2000). On agent-based software engineering. *Artificial intelligence, 117*(2), 277-296.

Jia, B., Li, X.-G., Jiang, R., & Gao, Z.-Y. (2007). Multi-value cellular automata model for mixed bicycle flow. *The European Physical Journal B, 56*(3), 247-252.

Jiang, R., Hu, M., Wu, Q., & Song, W. (2016). Traffic Dynamics of Bicycle Flow: Experiment and Modeling. *Transportation Science*, 998-1008.

Jiang, R., Jia, B., & Wu, Q.-S. (2004). Stochastic multi-value cellular automata models for bicycle flow. *Journal of Physics A: Mathematical and General, 37*(6).

Khan, S., & Raksuntorn, W. (2001). Characteristics of passing and meeting maneuvers on exclusive bicycle paths. *Transportation Research Record: Journal of the Transportation Research Board*(1776), 220-228.

Khan, S., & Raksuntorn, W. (2001). Characteristics of passing and meeting maneuvers on exclusive bicycle paths. *Transportation Research Record: Journal of the Transportation Research Board 1776*, 220-228.

Leutzbach, W., & Wiedemann, R. (1986). Development and applications of traffic simulation models at the Karlsruhe Institut fur Verkehrwesen. *Traffic engineering and control, 5*(27).

Levine, S., & Koltun, V. (2012). Continuous inverse optimal control with locally optimal examples. *Proceedings of the 29th International Conference on Machine Learning (ICML)*. Edinburgh, Scotland, UK.

Levine, S., Popovic, Z., & Koltun, V. (2010). Feature construction for inverse reinforcement learning. *Advances in Neural Information Processing Systems.*

Li, L., & Li, X. (2019). Parsimonious trajectory design of connected automated traffic. *Transportation Research Part B: Methodological*.

Liang, X., Baohua, M. A., & Qi, X. U. (2012). Psychological-physical force model for bicycle dynamics. *Journal of Transportation Systems Engineering and Information Technology, 12*(2), 91-97.

Liang, X., Baohua, M. A., & Qi, X. U. (2012). Psychological-physical force model for bicycle dynamics. *Journal of Transportation Systems Engineering and Information Technology, 2*(12), 91-97.

Liu, J., & Zhou, X. (2016). Capacitated transit service network design with boundedly rational agents. *Transportation Research Part B: Methodological, 93*, 225-250.

Lucas, B. D., & Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision. *International Joint Conference on Artificial Intelligence*, (pp. 674-679).

Ma, X., & Luo, D. (2016). Modeling cyclist acceleration process for bicycle traffic simulation using naturalistic data. *Transportation research part F: traffic psychology and behaviour, 40*, 130-144.

Ma, X., & Luo, D. (2016). Modeling cyclist acceleration process for bicycle traffic simulation using naturalistic data. *Transportation research part F: traffic psychology and behaviour*(40), 130-144.

Mahajan, B., & Marbate , P. (2013). Literature review on path planning in dynamic environment. *International Journal of Computer Science and Network*.

Michon, J. A. (1985). A critical view of driver behaviour models: what do we know, what should we do? *Human behaviour and traffic safety*, 485-524.

Mohammed, H., Bigazzi, A. Y., & Sayed, T. (2019). Characterization of bicycle following and overtaking maneuvers on cycling paths. *Transportation Research Part C: Emerging Technologies, 98*, 139-151.

Mohammed, H., Sayed, T., & Bigazzi, A. (2019). Toward agent-based microsimulation of cyclist following behaviour: Estimation of reward function parameters using inverse reinforcement learning. *98th Annual Meeting of the Transportation Research Board.* Washington D.C.

Nagel, K., & Schreckenberg, M. (2002). Cellular Automata Approach to Three-Phase Traffic Theory. *Physica A: Mathematical and General, 35*(47), 9971-10013.

Ng, A., & Russel, S. (2000). Algorithms for inverse reinforcement learning. *International Conference on Machine Learning.*

Paulsen, M., Rasmussen, T., & Ank, O. (2019). Fast or forced to follow: A speed heterogeneous approach to congested multi-lane bicycle traffic simulation. *Transportation research part B: methodological*, 72-98.

Pelosi, M., Alfò, M., Martella, F., Pappalardo, E., & Musarò, A. (2015). Finite mixture clustering of human tissues with different levels of IGF-1 splice variants mRNA transcripts. *BMC bioinformatics*(16).

Plekhanova, V. (2003). *Intelligent agent software engineering.* IGI Global.

Plekhanova, V. (2003). *Intelligent agent software engineering.* IGI Global.

PTV. (2011). *VISSIM 5.40 user manual.* Karlsruhe, Germany .

Rakha, H., & Wang, W. (2009). Procedure for calibrating Gipps car-following model. *Transportation Research Record: Journal of the Transportation Research Board*(2124), 113-124.

Ratliff, N., Bagnell, A., & Zinkevich, M. (2006). Maximum margin planning. *Proceedings of the 23rd international conference on Machine learning (ICML)*, (pp. 729-736).

Ratliff, N., Silver, D., & Bagnell, A. (2009). Learning to search: Functional gradient techniques for imitation learning. *Autonomous Robots, 27*(1), 25-53.

Saunier, N., & Sayed, T. (2006). A feature-based tracking algorithm for vehicles in intersections. *Third Canadian Conference on Computer and Robot Vision.* Quebec City, Quebec, Canada.

Savitzky , A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedure. *Analytical Chemistry, 36*(8), 1627-1639.

Schulman, J. L. (2015). Trust region policy optimization. *International conference on machine learning.*

Scrucca , L., Fop , M., Murphy , T. B., & Raftery , A. E. (2016). mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal, 1*(8), 2055-233.

Sud, A., Anderson, E., Curtis, S., Lin, M., & Manocha, D. (2007). Real-time path planning for virtual agents in dynamic environments. *IEEE Virtual Reality Conference.* Charlotte, North Carolina, USA.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning.* MIT press.

Tang, T.-Q., Rui, Y.-X., Zhang, J., & Shang, H.-Y. (2018). A cellular automation model accounting for bicycle's group behaviour. *Physica A: Statistical Mechanics and its Applications, 492* , 1782-1797.

Taylor, D. B., & Mahmassani, H. S. (1998). *Behavioural models and characteristics of bicycle-automobile mixed-traffic: planning and engineering implications.* University of Texas, Austin.

Therneau, T., & Atkinson, B. (2018). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-13. Retrieved from https://CRAN.R-project.org/package=rpart

Tomasi, C., & Kanade, T. (1994). Detection and Tracking of Point Features. *IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 593-600).

Twaddle, H., & Grigoropoulos, G. (2016). Modeling the Speed, Acceleration, and Deceleration of Bicyclists for Microscopic Traffic Simulation. *Transportation Research Record: Journal of the Transportation Research Board, 2587* , 8-16.

Twaddle, H., Schendzielorz, T., & Fakler, O. (2014). Bicycles in urban areas: Review of existing methods for modeling behaviour. *Transportation Research Record: Journal of the Transportation Research Board*(2434), 140-146.

Twaddle, H., Schendzielorz, T., & Fakler, O. (2014). Bicycles in urban areas: Review of existing methods for modeling behaviour. *Transportation Research Record: Journal of the Transportation Research Board 2434*, 140-146.

Wei, Y., Avci, C., Liu, J., Belezamo, B., Aydin, N., Li, P., & Zhou, X. (2019). Dynamic programming-based multi-vehicle longitudinal trajectory optimization with simplified car following models. *Transportation Research Part B: Methodological*.

Wolfram, S. (1983). Statistical mechanics of cellular automata. *Reviews of modern physics, 55*(3).

Wolfram, S. (1984). Cellular automata as models of complexity. *Nature, 311*, 419-424.

Wolfram, S. (1985). Twenty problems in the theory of cellular automata. *Physica Scripta* .

Xue, S., Jia, B., Jiang, R., Li, X., & Shan, J. (2017). An improved Burgers cellular automaton model for bicycle flow. *Physica A: Statistical Mechanics and its Applications, 487*, 164-177.

Zaki, M., & Sayed, T. (2016). Automated cyclist data collection under high density conditions. *IET Intelligent Transport Systems*, 361 –369.

Zaki, M., Sayed, T., & Cheung, A. (2013). Computer Vision Techniques for the Automated Collection of Cyclist Data. *Transportation Research Record: Journal of the Transportation Research Board*, 10-19.

Zhao, D., Wang, W., Li, C., Li, Z., Fu, P., & Hu, X. (2013). Modeling of passing events in mixed bicycle traffic with cellular automata. *Transportation Research Record: Journal of the Transportation Research Board 2387*, 26-34.

Zhao, D., Wang, W., Li, C., Li, Z., Fu, P., & Hu, X. (2013). Modeling of passing events in mixed bicycle traffic with cellular automata. *Transportation Research Record: Journal of the Transportation Research Board*(2387), 26-34.

Zhao, D., Wang, W., Li, C., Li, Z., Fu, P., & Hu, X. (2013). Modeling of Passing Events in Mixed Bicycle Traffic with Cellular Automata. *Transportation Research Record*, 26-34.

Zhao, Y., & Zhang, H. (2017). A unified follow-the-leader model for vehicle, bicycle and pedestrian traffic. *Transportation Research Part B: Methodological, 105*, 315-327.

Zhao, Y., & Zhang, H. M. (2017). A unified follow-the-leader model for vehicle, bicycle and pedestrian traffic. *Transportation Research Part B: Methodological, 105*, 315-327.

Ziebart, B., Maas, A. L., Bagnell, A., & Dey, A. K. (2008). Maximum entropy inverse reinforcement learning. *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence.*