UNIVERSITY OF THE PHILIPPINES

Doctor of Philosophy in Physics

Damian N. Dailisan

Modeling Transport: Different Aspects of Urban Mobility Research

Dissertation Adviser:

May T. Lim, Ph.D.

National Institute of Physics

University of the Philippines Diliman

Date of Submission:
December 2020

Dissertation classification:
P

This dissertation is not available to the public. Please ask the library for assistance.

National Institute of Physics
College of Science
University of the Philippines
Diliman, Quezon City

ENDORSEMENT

This is to certify that this dissertation titled **Modeling Transport: Different Aspects of Urban Mobility Research**, prepared and submitted by Damian N. Dailisan in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics, is recommended for acceptance.

MAY T. LIM, Ph.D.
Dissertation Adviser

JOSE PERICO H. ESGUERRA, Ph.D.
Dissertation Reader

RONALD S. BANZON, Ph.D.
Dissertation Reader

The National Institute of Physics endorses acceptance of this dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics.

WILSON O. GARCIA, Ph.D.
Director
National Institute of Physics

The dissertation is officially accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics.

GIOVANNI A. TAPANG, Ph.D.
Dean
College of Science

Office of the Associate Dean for Mentoring Academic Progress and Advancement

GRADUATE OFFICE

REPORT OF MASTERAL/DOCTORAL EXAMINATION

Name of Candidate:	Damian N. Dailisan
Degree Sought:	Ph.D. Physics
Title of Thesis/Dissertation:	Modeling Traffic: Different Aspects of Urban Mobility Research
Schedule and Place of Master's/Dissertation Examination:	Dec 16, 2020. 4:00 PM via Zoom. Link: https://up-edu.zoom.us/j/2532175689

ACTION OF MASTER'S/DOCTORAL EXAMINATION PANEL

	PANEL MEMBER	FOR ACCEPTANCE	FOR REJECTION	PROVISIONAL
1.	May T. Lim, Ph.D. (ADVISER)	_____	_____	_____
2.	Jose Perico H. Esguerra, Ph.D. (READER)	_____	_____	_____
3.	Ronald S. Banzon, Ph.D. (READER)	_____	_____	_____
4.	Reinabelle C. Reyes, Ph.D. (EXAMINER)	_____	_____	_____
5.	Varsolo C. Sunio, Dr. Eng. (EXAMINER)	_____	_____	_____

PANEL DECISION: PASSED (✓) FAILED ( ) PROVISIONALLY PASSED ( )

CONDITIONS, IF ANY: _____

Submitted by:

Jose Perico H.Esguerra, Ph.D
Chairman, Examination Panel

Wilson O. Garcia, Ph.D.
Chair, Graduate Committee

Date: 16 December 2020

Date: _____

Publications

Parts of this work have appeared in the following publications:

- Rubio, L. J. M., **Dailisan, D. N.**, Osorio, M. J. P., David, C. C., and Lim, M. T. (2019). *Modeling the residential distribution of enrolled students to assess boundary-induced disparities in public school access.* PLOS ONE, 14(10), 1–15.
- **Dailisan, D. N.**, and Lim, M. T. (2020). *Crossover transitions in a bus–car mixed-traffic cellular automata model.* Physica A, 557, 124861.

Acknowledgments

First of all, to **Dr. May Lim**, for all of our discussions, traffic related or not. You have taught me a lot about how to think critically and analyze different phenomena around us. Thank you for your patience in mentoring me from my undergraduate to doctoral studies.

To my thesis readers, **Dr. Jose Perico Esguerra** and **Dr. Ronald Banzon**, as well as my examiners **Dr. Reinabelle Reyes** and **Dr. Varsolo Sunio**. Thank you for you comments and insights, which have improved the quality of this work.

To all the institutions that have shared data that made this thesis possible. To the **Philippine Statistics Authority**, the **National Mapping and Resource Information Authority**, and the **Land Management Bureau** for the census and barangay shapefile data. Special thanks to **Ardie Orden** for sharing the barangay shapefile data with us. To the Department of Transportation, through Dr. Varsolo Sunio for allowing access to the RTRS data. And to **SakayPH**, through **Dr. Noriel Tiglao** for sharing their GTFS data.

My brief research visit to UC Berkeley helped broaden my perspectives in research. Thank you to **Prof. Alexandre Bayen** and his lab for hosting me during my 2-month stay in Berkeley. Thank you to **Gabriel Gomes**, for sharing your expertise on traffic models. Our discussions motivated my analysis of public transportation at the scale of Metro Manila. Thank you to **Qijian Gan** for our discussions that helped me understand traffic from the perspective of traffic engineers.

To the **CX Team**, for being my second set of mentors, friends, and companions. To **Chester** for all of the valuable insights and discussions, as well as all of the well deserved sanity breaks (a.k.a. meryenda) back when the pandemic was not a thing. To **Louie** and **Ivan**, for all of the research experiences we shared in the project, and as CX team members. And the rest of the team, **Jona**, **JP**, **Cyd**, **Jejo**, **Joseph**, **Loraine**, **Cri-cri**, **Charles**, **Kenneth**, **Kelvin**, **Mark**, **Elmo**, and **Olyn**. I wish you all the best in

research and your future careers!

To the **IPL Staff** for inspiring me to do research work in Instru. All of you who research on not only interesting fields, but also dedicate your work to understanding and improving the Philippine situation. Thank you for creating an enabling environment where brilliant minds are able to thrive.

To my NIP PhD batchmates **Alfred**, **Reggie**, and **Ritz**. Together, we survived the stress and hardships involved in juggling research, acads, and teaching. There will be more challenges ahead of us, but I believe we can all succeed in our fields and specializations.

Finally, thank you to my family, for everything.

Funding Agencies

My Ph.D. journey would not have been financially possible without the agencies whose funding helped bring this research to life. I would like to thank the following funding agencies:

- Commission on Higher Education, through the Philippine–California Advanced Research Institutes: Data Analytics for Research and Education [IIID-2016-006].
- The University of the Philippines Office of the Vice President for Academic Affairs through the Emerging Interdisciplinary Research [C 06-013.1].

ABSTRACT

Modeling Transport: Different Aspects of Urban Mobility Research

Damian N. Dailisan
University of the Philippines, 2020

Adviser:
May T. Lim, Ph.D.

We analyzed traffic in Metro Manila using a range of modeling tools and perspectives: from evaluating traffic using traditional speed and flow metrics, measuring passenger wait times, modeling demand, to analyzing the system using a network perspective. We modeled vehicular traffic and accounted for bus–passenger interactions. The model exhibits platooned to non-platooned (free-flow and congested) states; platoons dissolve by either a decrease in passenger arrival rates or an increase in bus densities. The critical passenger arrival rate at which platoons dissolve is an exponential function of vehicle density. At low bus densities, closer stops induce platooned states, which reduces system speeds and increases wait times of passengers. We proposed a data-driven method to estimate traffic demand using establishment demand, granular residence, and road network data. This used a gravity-model to estimate OD trip demands, but the smaller scale of Metro Manila yields poor estimates of trip demands between zone-pairs when compared to similar approaches used for the US. Lastly, we evaluated public transportation in Metro Manila through a hybrid network–GIS perspective that allowed us to use a traveler-centric approach. Our analysis of Origin–Destination (OD) and Boarding–Alighting (BA) surveys highlighted transport-deficient zones. In 2020, the LTFRB proposed a new set of public transport routes with reduced route lengths. Route design for cities presents challenges, as routes should ideally improve transport coverage and trip frequency, without worsening the commutes of passengers. We compared these new routes with the old routes before the route modernization and found a reduction in the number of trips completed in 2-or-less rides (62.1% vs. 96.1%), and resulted in walks longer than 500 m for 24.7% of passengers. Our analysis shows that these proposed routes inadvertently result in a degradation of the rider experience.

Significance

Phase transitions in systems are characterized by a rapid change of properties of the system in response to small changes in certain parameter. Freeway traffic models exhibit a phase transition from free-flow to congestion at a critical density. This phase transition implies that congestion can be addressed by reducing vehicle densities below the critical density, but is impractical due to the current volume of traffic in cities. By accounting for interactions between public transportation and commuters, we find a different crossover transition from platooned to non-platooned states, and this occurs at densities above the phase transition; this means that it is possible to improve the current state of traffic by managing the supply and demand side of public transportation.

Modeling traffic is a many step process. One of the key components in modeling traffic is to know the demand for transportation between different pairs of locations (origins and destinations) in a city. Traditional methods of obtaining this are usually done by surveys, which are labor intensive and time consuming. An alternative approach is to use models of transportation and migration that take inspiration from physics models. The gravity model (inspired by Newton's Law of Gravitation) is such an example, which relates the flow of people between an origin and destination as proportional to the product of the population of the two locations, and inversely proportional to the distance between them. Our algorithm can help estimate demand by leveraging the availability of online data, in the absence of more expensive surveys that occur infrequently. With the right models, city planners can be guided on how transportation demand, establishments, vehicles, road configurations, and policies that impact traffic congestion.

Preamble

Transportation drives economic activity in urban areas; it enables people and goods to move between various locations within cities. Over the years, economic growth and developments in urban areas have resulted in increased demand for transportation. This increase in demand also brings problems to urban areas, such as congestion, traffic accidents, and even pollution. Transportation studies aim to solve the above by addressing inefficiencies in transportation systems. Studying transportation spans several scales:

- Individual: Commuters need to get from an origin to a destination.
- Transport Network: Commuters plan trip routes and choose between available transport modes (This can be a choice between personal transport, public transport, or even a combination of the two.)
- Traffic Dynamics: The collective use of roads by commuters determines traffic flow.

The first two provide an idea of the traffic demand on the transportation system, while the last allows us to understand how the resulting demand on roads shape traffic flow. These three scales are highly interdependent – high traffic demand on roads results in congestion, which causes commuters to adjust route choices accordingly.

This work is divided into two parts that cover the three scales identified above. The first half of this work delves into a traffic model that allows us to incorporate driving behavior and public transportation in traffic simulations. We focus on the dynamics of the model that result in crossover transitions, which hint at opportunities to reduce congestion.

The second half of this work discusses the public transportation system of Metro Manila. We propose a data-driven gravity model approach to generate trip distributions from origins to destination using establishment data. Finally we analyze various datasets to characterize and identify transport deficient areas and transportation hubs, and evaluate the current use of the public transportation system of Metro Manila.

Table of Contents

List of Figures and Tables

Chapter 1

Traffic Modeling

1.1 Introduction

Public transportation is universally acknowledged as a fundamental component in solving traffic congestion. Together with rail systems, buses form the backbone of medium- to long-haul modes of people transport. Since the interaction of buses and passengers introduces complex behavior in transportation systems, more so in cities that do not have designated stops, several models have been proposed to study bus traffic [1–5]. In these models, a delay in the arrival time of a bus leads to more passengers waiting for the bus, which then leads to further delays. Succeeding buses find fewer waiting passengers leading them to catch up to the delayed bus. Thus, buses form platoons (or bunches). These models also show a transition from the platooning state to a non-platooned state with increasing bus density.

The tendency of buses to form platoons poses problems for public transport. In an ideal scenario, an efficient transport system would try to maintain equal time intervals between vehicle arrivals. However, an equal headway configuration of vehicles is unstable [6], more so with buses [1]. Since the instability is inherent to the interaction between public transport vehicles and

Parts of this work appear in "Dailisan, D. N., and Lim, M. T. (2020). Crossover transitions in a bus–car mixed-traffic cellular automata model. Physica A, 557, 124861. https://doi.org/10.1016/j.physa.2020.124861"

passengers, approaches to maintain equal headways should consider both traffic and passenger behavior [7].

Bus route models [1–3] omit interactions between buses and other vehicle types. Yet we expect that these vehicular interactions play a key role in the dynamics of traffic flow — buses making curbside stops impede traffic flow outright [8], while small perturbations in vehicle speeds can induce congestion [9, 10]. Even a single bus in two-lane mixed traffic alters traffic states and jam transition densities [11]. As such, it is also critical to decision makers when curbside stops have to be replaced by bus bays to alleviate congestion [12]. Using a modified comfortable driving model, Yuan et al. [13] focused on system performance to show a dependence of the system capacity on the number of bus stops. They also found a gradual transition from platooned to non-platooned states. However, they did not explore the interplay of passenger arrival rates and vehicle densities extensively nor considered passenger waiting times.

It is important to understand transitions in traffic models with public transportation in mind. Slower moving buses and the associated platoon formation can result in the perception that buses cause congestion. Even a few undisciplined bus drivers can often cause traffic jams when their large vehicles straddle multiple lanes [14, 15]. Such situations, as well as a lack of understanding of these systems, can lead to proposals by policymakers to ban buses to alleviate congestion. An example of which would be the proposal to ban provincial buses (which account for less than 3% of traffic volume [16]) from Metro Manila's highways in order to alleviate congestion in one of the city's major thoroughfares [17, 18]. A Melbourne study found that is bus services are removed, up to 30% of bus users would shift to cars — this defeats the purpose of eliminating buses to alleviate congestion [8]! Knowledge of the factors that influence transitions can inform decisions, avoiding policies that do more harm to the current state of traffic congestion.

There is an urgent need for solutions to our country's traffic woes. Traffic

costs estimates in Metro Manila alone amount to PhP2.4 billion daily [19]. The government's approach to solving traffic has been mainly about building more roads, replacing aging jeepneys with newer ones, and removing buses from the main thoroughfares of the metro. While there are efforts to augment our lacking mass transit system with the MRT-7 and LRT-2 extensions, and MRT-3 rehabilitation projects, the quality of the public transportation experience has noticeably gotten worse through the years. Our poor perception of public transport has led to a private-car driven mindset, which leads to worsening congestion.

This work aims to create a traffic simulation model that incorporates public transportation and scenarios commonly found in the Philippine context. We create a modified Nagel-Schreckenberg (NaSch) model with buses and cars as primary vehicle types. Buses can stop at arbitrary positions on the road to pick up waiting passengers.

We focus on the state transitions in the model, and its dependence on vehicle densitiy and passenger arrival rate. We show how the interplay between vehicle density and passenger arrival rates affect the crossover transition from non-platooned to platooning states. Both single- and multi-lane cases, as well as mixed traffic scenarios, are studied. Lastly, we look at the effects of the placement of bus stops on traffic flow, as well as passenger waiting times.

1.2 Traffic Models

Greenshields et al. performed the first systematic study of traffic in 1933 when they empirically measured the relationship between vehicle speed and density [20, 21]. They found a linear relationship between speed and density, which can be written in the form $v(\rho) = v_{\max}\,(1 - \rho/\rho_{\text{jam}})$. Using the from the flow-density relationship given by

$$q(x,t) = v(x,t)\rho(x,t), \tag{1.1}$$

the Greenshields fundamental diagram took the form of $q = v_{\max}(\rho - \rho^2/\rho_{\text{jam}})$. This fundamental diagram, predicts that traffic exists in two states: free flow ($\rho < \frac{\rho_{\text{jam}}}{2}$) and congestion ($\rho > \frac{\rho_{\text{jam}}}{2}$).

With the increased availability of instruments and sensors that can collect traffic data, recent measurements show that the speed-density relationship is not linear [6, 22]. Since then, various models have been proposed to explain traffic phenomena.

1.2.1 Macroscopic models

Macroscopic traffic models describe traffic from a fluid dynamics perspective by relating the flow of vehicles to the density using the equation

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0. \tag{1.2}$$

These models use the fundamental diagram as an ansatz to solve the differential equations. The earliest know macroscopic model is Lighthill-Whitham-Richards (LWR) model [23, 24]. Here, they assumed that the flow is solely a function of density, that is $q(x,t) = q(\rho(x,t)) = \rho v(\rho(x,t))$. Solutions from macroscopic models predict shock waves (rapid changes in the density profile) which propagate opposite traffic flow [23–26]. These models have also shown the instability of a uniform-density configuration of traffic, tending to form density waves [25]. They have even been extended to describe heterogeneous traffic [27, 28], and use to fluid dynamic equations (like the Navier-Stokes equation) [25, 29].

While successful at reproducing real traffic properties, macroscopic models work best on freeways, and are unsuited to studying city traffic. Incorporating driver behavior would involve a need for defining and controlling parameters which dictate the behavior of vehicles. Using a macroscopic way of defining traffic does not allow us to explicitly control driving behavior in traffic models.

1.2.2 Microscopic models

Microscopic traffic models treat traffic as a system on interacting particles where the dynamics is modeled as a response to some stimuli [9, 30–32]. The simplest microscopic traffic model is the car-following model which can be written as

$$\ddot{x}_j(t+T) = \frac{1}{\tau}\left(\dot{x}_{j+1}(t) - \dot{x}_j(t)\right), \tag{1.3}$$

where T is some delay, and $\frac{1}{\tau}$ is the sensitivity to the difference in velocities of the current (j) and leading car $(j+1)$. With these models, it is easy to see the exact vehicle dynamics and incorporate particular driver characteristics such as reaction time. Traffic instability in these models results from an *over-deceleration effect* [33], wherein overreactions and finite reaction times of drivers result in waves of speed reduction in traffic flow. While fundamental relation cannot be obtained analytically from the equations governing these systems, ensemble averages still yield experimental fundamental diagrams that agree with empirical measurements.

These types of models can be used to model city traffic, and is at the heart of traffic simulation software such as SUMO [34], Aimsun [35], PTV Vissim [36], and others. Microscopic models describe instabilities of traffic well, but become computationally expensive due of the number of particles needed to be modeled as continuous non-linear differential equations [37].

1.2.3 Cellular automata

Cellular automata (CA) aims to provide a computationally efficient alternative to microscopic traffic models. First CA models for traffic were proposed by Nagel and Screckenberg, and use discretized space and time while still modeling traffic at the vehicular level [22, 38, 39]. Well defined state transitions encode driving rules, while additional transitions can easily be added to encode complex driving behavior; these circumvent the complexity of using differential equations to describe each vehicle.

Simplification by discretization allows CA models to simulate traffic at faster than real time speeds, while describing macroscopic phenomenon [22]. The Nagel–Schreckenberg (NaSch) model describes the real freeway traffic feature of free flow at low road densities followed by a transition to jamming at a critical density [39, 40]. The NaSch model can easily be modified to incorporate geometric constraints such as median U-turns [41]. The abstraction of CA also makes it simple to introduce rules that change the behavior of vehicles — one can model the effect of differences in driving behavior by changing how agents interact with each other. Cellular automata models have been used to study lane changing [14, 15], heterogeneity [13, 14, 42], bus dynamics [1, 5, 12, 13], and U-turns [41], among others. When a vehicle can only be in two states, namely stationary and moving at $v_{max} = 1$, the NaSch model reduces to the well-studied Totally Asymmetric Single Exclusion Process (TASEP), which is a particle hopping model wherein particles only move in one direction with a fixed probability [43].

1.2.4 Modified Nagel-Schreckenberg model

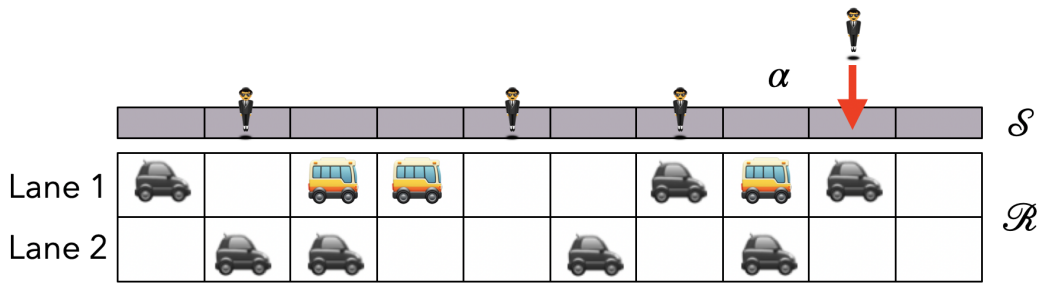

Figure 1.1: Diagram of road $\mathcal{R}$ and sidewalk $\mathcal{S}$ models. Pedestrians can spawn on an empty cell of $\mathcal{S}$ with probability α. In multi-lane scenarios, buses are limited to the lane adjacent to the sidewalk.

Our model combines elements from the Nagel-Schreckenberg (NaSch) [14, 22] model and Bus Route Model (BRM) [1]. The two components of this hybrid model are the road and sidewalk (Fig. 1.1). The road model $\mathcal{R}$ has

N lanes of length L sites, with periodic boundary conditions. A vehicle can interact with passengers (buses), or ignore passengers (cars). A sidewalk model $\mathcal{S}$ with one lane of length L sites is updated synchronously with the road model. States for the ith vehicle are lane l^i, position x^i, and speed $0 \leq v^i \leq v_{\max}$. We denote sites occupied by cars as $\mathcal{R}(x,l) = 1$ and those occupied by buses as $\mathcal{R}(x,l) = 2$. The road lane adjacent to the sidewalk is denoted as $l = 1$. Passengers occupy the sidewalk if $\mathcal{S}(x) = 1$. Similar to the BRM, we impose that $\mathcal{R}(x, l = 1) = 2$ and $\mathcal{S}(x) = 1$ cannot occur simultaneously, i.e. passengers cannot spawn when a bus occupies the adjacent road cell.

Realizations of the model involve assigning vehicle density ρ, passenger arrival rate α, and bus to vehicle ratio f_B. Vehicle states are updated in random sequential order at each timestep t ($\delta t = 1.65\,\text{sec/step}$) following these rules:

R1: Acceleration: $v^i_{t+1} \leftarrow \min\left(v^i_t + 1, v_{\max}\right)$

R2: Bus Loading: $v^i_{t+1} \leftarrow 0$ if $\mathcal{S}(x^i_t) = 1$; $\mathcal{S}(x^i_t) = 0$

R3: Lane Change: $l_i \leftarrow l_i \pm 1$ with probability p_l if $v^i_{t+1} > \Delta x^i$

R4: Deceleration: For cars, $v^i_{t+1} \leftarrow \min\left(v^i_{t+1}, \frac{\Delta x^i}{\Delta t}\right)$.

For buses,

$$v^i_{t+1} \leftarrow \begin{cases} \min\left(v^i_{t+1}, \frac{\Delta x^i}{\Delta t}\right), & \text{if } \sum_{k=1}^{\lfloor \frac{v_{\max}}{a_{\text{dec}}} \rfloor} (v_t - k a_{\text{dec}}) > \frac{\Delta x_{\text{pass}}}{\Delta t} \\ \min\left(v^i_{t+1}, \frac{\Delta x^i}{\Delta t}, \max(v_t - a_{\text{dec}}, a_{\text{dec}})\right), & \text{if } v_t + a_{\text{dec}} \geq \frac{x}{\Delta t} \geq a_{\text{dec}} \\ \min\left(v^i_{t+1}, \frac{\Delta x^i}{\Delta t}, \frac{\Delta x_{\text{pass}}}{\Delta t}\right), & \text{otherwise.} \end{cases}$$

R5: Random Slowdown: $v^i_{t+1} \leftarrow \max\left(v^i_{t+1} - 1, 0\right)$ with probability p_{slow}

R6: Forward Movement: $x^i_{t+1} \leftarrow x^i_t + v^i_{t+1}\Delta t$

The headway Δx^i is the number of empty cells ahead of vehicle i while the passenger headway Δx_{pass} is the distance from the bus to the closest passenger. The notation $\lfloor x \rfloor$ denotes the floor operator. The implementation of **Deceleration** for buses ensures deceleration rates are limited to a_{dec} when

picking up passengers [44, 45]. The first criterion lets a bus ignore passengers that appear when the bus is passing too fast to safely decelerate. The second criterion allows a bus to anticipate stopping, and decelerate safely. The slowdown probability p_{slow} parametrizes non-acceleration or overreactions to decelerating. For simplicity, both vehicle types have similar parameters $v_{\max}$ and p_{slow}, with the only difference between the two being the additional interaction with pedestrians[1] for buses. Additionally, buses have infinite passenger capacity, though we include a short discussion on the effect of finite passenger capacities in Sec. 1.3.1. For cars, **Bus Loading** is skipped. Vehicles may change lanes with probability $p_l \in \{0, 1\}$ with equal chances of moving left or right. Vehicle i attempts to move into adjacent lanes to avoid decelerating, but it may only change lanes if the target lane is empty and it satisfies the safety criteria. It is safe to change lanes when the trailing vehicle in the target lane can safely decelerate without crashing into vehicle i. The safety criteria can be written as $v^i + (x^i - x^j) \geq v^j - a_{\text{dec}}$, where j denotes the trailing vehicle. In the context of overtaking vehicles, **Random Slowdown** is skipped when a vehicle successfully changes lanes. Otherwise, the original NaSch rules (**R1, R4, R5, R6**) are followed. Through these rules, interactions of vehicles and pedestrians give rise to complex dynamics of traffic flow.

We measure the time averaged flow q of the system, defined as

$$q = \frac{1}{T} \frac{\sum_{t=1}^{T} \sum_i v_t^i \Delta t}{L}, \tag{1.4}$$

where v_t^i is the speed of the i^{th} vehicle at time t. The average speed of vehicles in the system is then $\bar{v} = \dfrac{q}{\rho}$.

We allow for a transient simulation time of $T_\tau = 3000$ timesteps so as not to include transient behavior in the data [41]. For all realizations, measurements span $T = 3000$ timesteps with fifty trials for each set of parameters ρ,

[1] For the sake of brevity and since a passenger waiting for the bus to arrive occupies sidewalk space, we use the terms pedestrian and waiting passenger interchangeably.

Table 1.2: Parameters and their values used in the simulation. In this model, each cell is 5.5 meters long and a timestep is 1.65 seconds.

Symbol	Description	Typical Values	Model Value
L	Length of the road	100 m (city), 10 km (highway)	500, 5120
v_{max}	Maximum speed	20 - 120 km/h (residential, highway)	5 (60 km/h)
f_B	Fraction of buses	0.03 [16]	0 - 1
p_{slow}	Slowdown probability	0.01 - 0.5 [22, 28, 41, 46]	0.1
p_l	Lane change probability	–	0 (bus), 1 (car)
a_{dec}	maximum deceleration	$4\,\text{m/s}^2$	2
ρ	Vehicle density	0 - 180 vehicles/km	0.02 - 0.98
T_τ	Transient time	–	3000
T	Measurement time	–	3000

α, and f_B. A complete list of parameters used in our simulations, with their corresponding calibration factor and real-world values, is given in Table 1.2.

1.3 Results and Discussion

In Sec. 1.3.1, we focus on the interplay of changing passenger arrival rates, fraction of buses, and transitions for the single lane case. Section 1.3.2 extends the model to multiple lanes, with buses being limited to the first lane. Section 1.3.3 discusses placement of bus stops, and the effect of different bus fractions on the waiting time of passengers.

1.3.1 Single lane case

If all vehicles on the road are buses ($f_B = 1$), we expect different behaviors at the extreme values of α. For the case $\alpha \to 0$, passengers do not arrive. With no passengers to pick up, buses do not slow down, and we recover the original NaSch model. In the NaSch model, we expect a phase transition

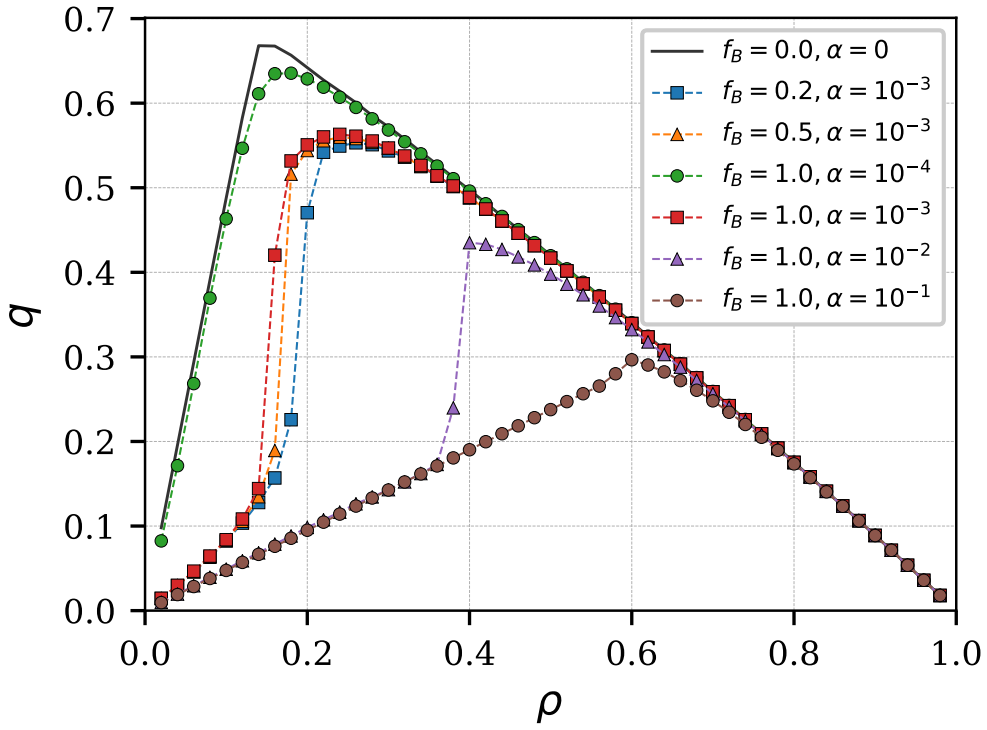

Figure 1.3: Fundamental diagram for the single lane case. As density increases, jumps in q are observed. These jumps occur at different densities, which appear to have a dependence on α.

to occur at $\rho_{\text{crit}} = \frac{1}{v_{\max}+1}$, which is the maximum allowable density where vehicles have enough headway to avoid slowing down [22, 47]. For the case of $v_{\max} = 5$, $\rho_{\text{crit}} \approx \frac{1}{6}$. Thus, increasing the density beyond ρ_{crit} induces a phase transition from free-flowing traffic to congestion. On the other extreme, as $\alpha \to 1$, buses stop after every other timestep. High values of alpha guarantee that there will be a passenger waiting for a bus at the next cell. Thus, at low densities, we expect buses in this system to move at an average speed of $\bar{v} = 0.5$ sites per timestep. A transition occurs at $\rho \approx \frac{2}{3}$, which corresponds to the critical density for $v_{\max} = 0.5$.

We verify the existence of these two limiting cases (Fig. 1.3). Two competing trends appear to govern the dynamics of the system: (1) bus inter-

actions with passengers which tend to drive the average speed of the system to $\bar{v} \to 0.5$; and (2) the dynamics of the NaSch model without buses ($\alpha \to 0$ or equivalently, $f_B = 0$). For the values of $f_B = 1.0$, $\alpha = [10^{-3}, 10^{-2}, 10^{-1}]$, we observe that for low densities, buses picking up passengers is the dominant behavior. However, increasing the density of buses results in jumps in throughput in the fundamental diagram. These jumps indicate that the dominant behavior of the system has shifted towards the dynamics of the fundamental NaSch model (Fig. 1.3, $f_B = 0$). We also observe that the density values at which these jumps occur shift to higher densities, as the passenger arrival rates are increased.

This interplay between α and ρ suggests that low values of density allow for passengers to accumulate on different sites of the sidewalk, which forces buses to make more frequent stops. On the other hand, increasing the density of buses not only accounts for the pick-ups of these passengers, but also prevents the accumulation of passengers in the system. An interesting consequence is that the average speed of the system actually increases.

We can look at the interplay between ρ and α in another way, by varying α for fixed values of ρ. We observe that the crossover from non-platooned to platooned states depends on α (Fig. 1.4a). For a fixed density, once α becomes sufficiently large, passengers accumulate in the system, which results in more frequent stops for buses. This sets the formation of a platoon of vehicles, while at the same time allowing for more passengers to accumulate ahead of the platoon.

Figure 1.5 illustrates the crossover from either a free flow or congested phase to a platooned phase for two different values of α. A platoon develops from an initial non-platooned configuration as arrival rates are increased above some critical value α^*. For a system with a density of $\rho = 0.4$, we find that $\alpha^* \approx 10^{-1.91}$ (Fig 1.4a). In the case of $\alpha = 10^{-2}$, platoons do not form in the system at all. However, when we look at the case of $\alpha = 10^{-1.8}$, a platoon develops after 600 timesteps, despite having the same initial configuration for

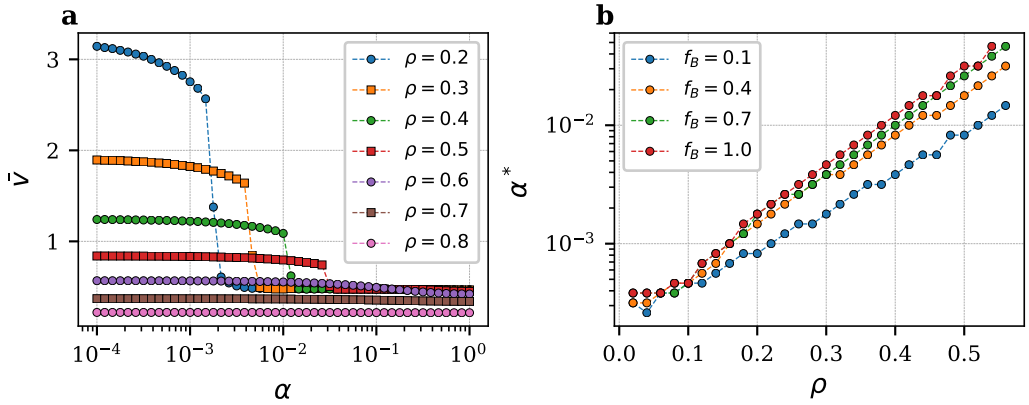

Figure 1.4: (a) Comparisons of speed as a function of the passenger arrival rate α for different densities ($f_B = 1$). Sudden drops in vehicle speeds as α is increased indicate a crossover from non-platooned to platooned flow. However, no clear transition is observed for $\rho \geq 0.6$. (b) The crossover transition involves a critical value α^* for different densities and bus fractions. The value of α^* increases monotonically with ρ.

the case $\alpha = 10^{-2}$. As passengers accumulate in the gaps between buses, the decrease in the speed of buses creates larger gaps. This feedback loop drives the formation of platoons and also results in longer passenger waiting times.

For the case of mixed traffic (Fig. 1.3, $\alpha = 10^{-3}$), the transition from a platooned to a non-platooned phase occurs at lower densities with increasing f_B. This observation is still consistent with the notion of the transition from platooned to a non-platooned phase due to increasing density of buses. As the effective bus density scales with f_B, higher densities are needed to compensate for lower f_B. Thus, reducing f_B shifts the crossover transition to higher densities.

We can make two observations of our single lane model. First, we see that there exists a $\rho^*(\alpha)$ responsible for a crossover behavior from a platooned phase to either a free-flow or congested phase. When $\rho < \rho^*$, a low f_B gives more time for passengers to accumulate in the gaps between buses. At the same time, buses will tend to clump together, creating larger gaps. Both

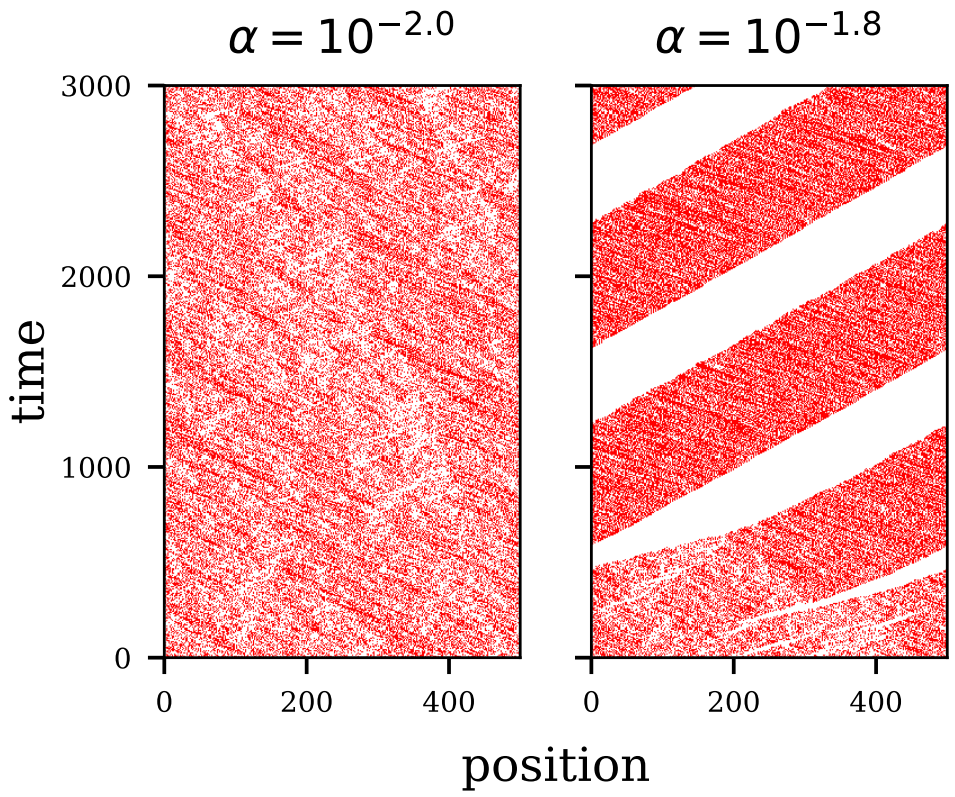

Figure 1.5: Spatio-temporal diagrams of the system of buses ($f_B = 1$) near the transition point of $\rho = 0.4$, for low passenger arrival rate ($\alpha = 10^{-2}$, left) and high passenger arrival rate ($\alpha = 10^{-1.8}$, right). Increasing α induces a transition from non-platooned to platooned flow.

processes aid in the formation of platoons. When $\rho > \rho^*$, the gaps between buses are sufficiently small to substantially reduce passenger accumulation that cause a cascade of slowdowns. Thus, smaller or no platoons are formed, and we obtain a non-platooned phase.

Secondly, the crossover from a platooned to a non-platooned phase is also determined by $\alpha^*(\rho)$. For arrival rates $\alpha < \alpha^*$, the system would be found in the non-platooned phase, while for $\alpha > \alpha^*$ the system will exhibit platooning. The dependence on ρ is only up to $\rho \approx \frac{2}{3}$, beyond which the density-dependent dynamics of the NaSch dominates the system. Figure

1.4b highlights the interplay between α^* and ρ. The crossover transition value appears to scale as $\alpha^* \sim \exp(\rho + c)$. We observe the scaling only until $\rho = 0.56$, beyond which the platooning and congestion effects are difficult to distinguish. Unlike the phase transition from free flow to congestion which only occurs at low densities, this crossover can occur at densities above ρ_{crit}. While the phase transition from free flow to congestion is undoubtedly an important aspect in the study of transport and vehicular traffic, the densities involved are typically low densities. Since the problem of congestion in city traffic involves densities above ρ_{crit}, this particular phase transition is rather insignificant when it comes to policy intervention since controlling vehicle volume is difficult. However, the crossover transition of passenger–bus interactions is present for a larger range of densities. The key to managing traffic flow at densities above ρ_{crit} must then lie in avoiding the formation of platoons.

The behavior of buses in our simplified model has underlying assumptions that are absent in real traffic. Buses in our model have infinite capacities, do not wait for passengers, and our model does not have passenger drop-offs. What happens when we relax these assumptions?

From infinite to finite bus capacities. With each bus having a passenger capacity c, the "leading bus" gets filled up and eventually stops picking up passengers. Said bus breaks away from the platoon that it leads, leaving the next bus as the new "leading bus". The notion of a "leading bus" in this case is not specific to a single bus, as any unfilled bus can be the "leading bus" and cause a platoon. A spatio-temporal diagram of our model with different capacities shows this phenomenon (Fig. 1.6).

If the passenger volume eventually exceeds aggregate bus capacity, all buses will be full after some time T_{cap}. In the case of Fig. 1.6, $T_{\text{cap}} \approx 3000$. At that moment, the platoon dissolves, and we recover the dynamics of the NaSch model. While this speeds up the movement of vehicles and passengers on board the buses, this also means that those who have yet to board will

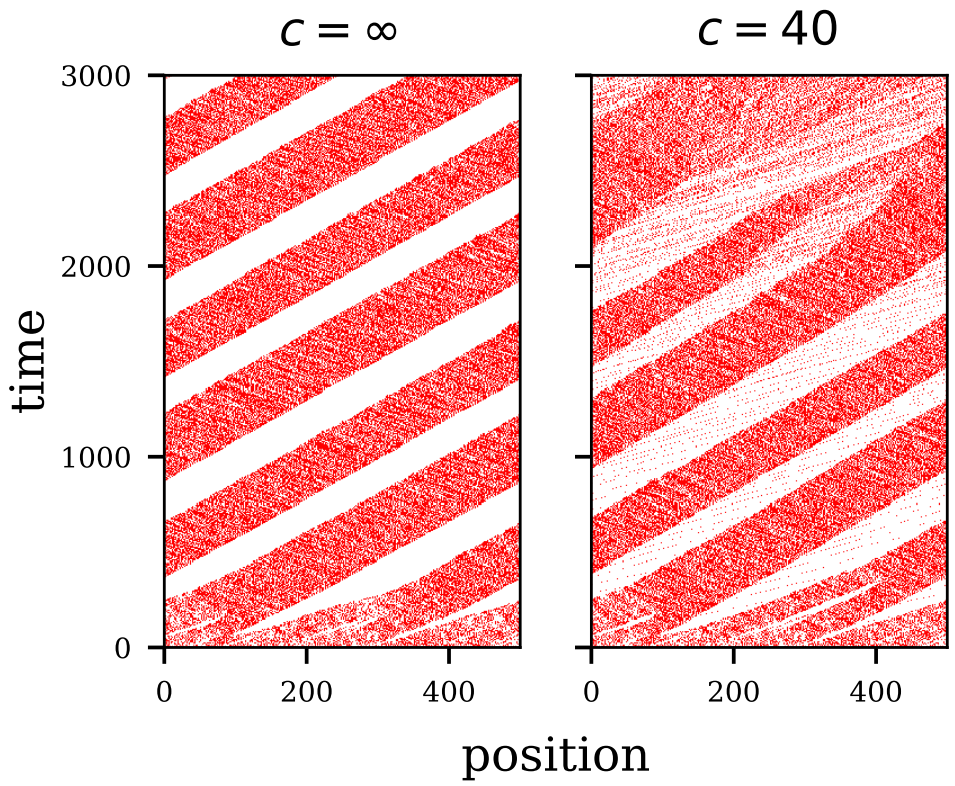

Figure 1.6: Spatio-temporal diagrams of the system of buses ($f_B = 1, \alpha = 10^{-1.6}, \rho = 0.4$) with different passenger capacities c. The system with finite capacity ($c = 40$) platoons like our simplified model ($c = \infty$) for some time $T_{\text{cap}} \approx 3000$ until the platoon dissolves as most buses get filled. Small dots breaking away from the platoon are buses have just reached their capacity.

have to wait indefinitely for a bus. We find that T_{cap} scales linearly with the passenger capacity c, for all densities and passenger arrival rates (Fig. 1.7).

From no-waiting to waiting for passengers. Allowing buses in our model to stop for a duration T_{wait} is analogous to waiting for passengers or making scheduled stops. In effect, this slows down the movement of the platoon (Fig. 1.8a). In the limit that α is sufficiently large, we can calculate the average speed of the platoon as $\bar{v} = \frac{1}{T_{\text{wait}}+1}$. In the case of our simplified model, $T_{\text{wait}} = 1$ and we have $\bar{v} = 0.5$.

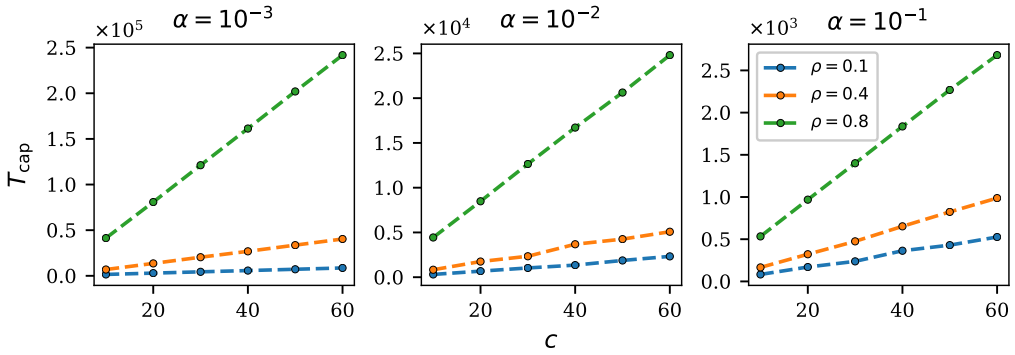

Figure 1.7: The time it buses to reach capacity T_{cap} as a function of the passenger capacity c for different combinations of α, ρ. T_{cap} scales linearly with the passenger capacity for all combinations of density and passenger arrival rate.

Aside from slowing down the platoon, increasing the stop duration of buses shifts the crossover transition to higher densities. We find that for the same densities, having shorter waiting times cause their transitions to occur at higher values of α (Fig. 1.8b). This ties into the mechanism of platoon formation, as when buses wait longer at stops, the average speed of the platoon decreases. For the same pedestrian spawning rate α, this decrease in platoon speeds results in more pedestrians accumulating ahead of the platoon.

Adding drop-offs. Our description of the model does not explicitly let passengers alight. While the term *picking up passengers* is useful in creating a mental picture of what goes on in the model, a better interpretation is that these are stopping events as a result of passengers boarding and alighting the bus. Instead of explicitly modeling multiple passengers boarding (and alighting) buses at the sidewalk, buses can be filled (and emptied) by multiple passengers at a single stop. This works in the regime where the net flux of passengers do not fill up buses.

When buses are close to capacity, explicitly modeling drop offs influences the dynamics of buses. In Fig. 1.6, we show that as buses reach their

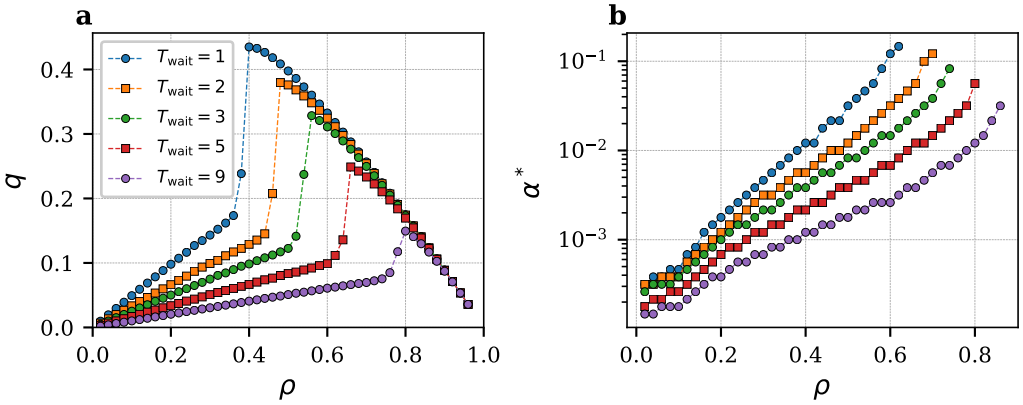

Figure 1.8: (a) Fundamental diagram for different bus waiting times T_{wait}. Waiting for passengers for a longer duration reduces the average speed of buses in the system, and shifts the crossover transitions to higher densities. (b) The critical value α^* vs. densities for different bus waiting times. Increasing T_{wait} changes the slope of $\log \alpha^*$ vs. ρ.

capacities, they break away from the platoon, and can end up at the tail of a different platoon. If no passengers alight, the bus catches up with the tail of the next platoon. On the other hand, if a passenger alights within the route, the bus once again becomes a potential platoon starter. In this way, a bus can alternate between a platoon starter, and a platoon trailer, depending on the distribution of drop-off points of passengers. If passenger drop-offs are spread uniformly throughout the route, we expect to see several smaller platoons instead of a single large platoon. Note that some commute routes would have a concentration of pickups at the start of the route, and drop-offs at the end of the route. This situation would resemble the dynamics of buses at capacity, with faster transit speeds as stops mid route are minimized.

1.3.2 Multi-lane traffic

We now extend the single-lane model to two lanes. In this model, we restrict buses to drive along the first lane, while cars can move freely on any lane. As buses stop to pick up passengers, they block the movement of vehicles

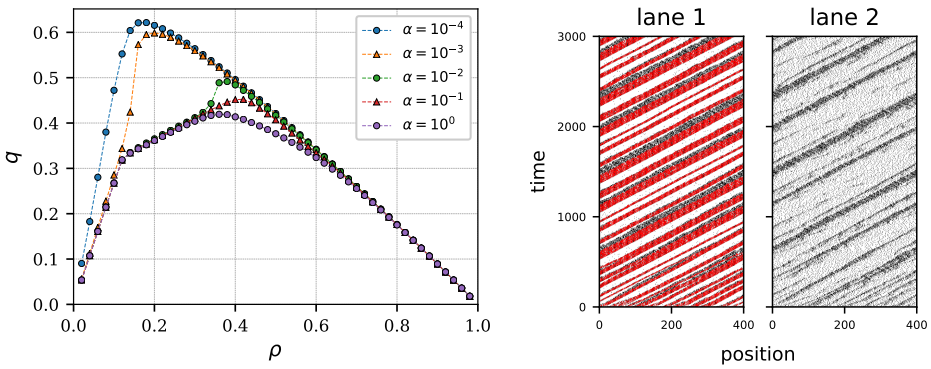

Figure 1.9: (Left) Fundamental diagram of multi-lane traffic. Jumps in q are not as obvious as in the single lane case, but are still observed for $\alpha = [10^{-3}, 10^{-2}]$. (Right) Spatio-temporal diagram for $\rho = 0.3, \alpha = 10^{-1}$, with cars (black) and buses (red). Coupling of both lanes occurs at the trailing end of platoons due to cars changing lanes.

on the bus lane. Lane changing allows cars to overtake stopped or slow-moving buses, though the maneuver can cause congestion on the adjacent lane. Buses cause bottlenecks, which greatly reduce the flow of vehicles in the system. These bottlenecks are similar to work zone scenarios [48], with the key difference being that the flow reduction is temporary and that the bottleneck is non-stationary.

In the case of multiple lanes, we see three regions in the fundamental diagram for the case of $\alpha > 10^{-3}$ (Fig. 1.9). The first two regions are from $0 < \rho < 0.12$ and $0.12 < \rho < 0.66$. In the single lane case, when α is sufficiently high, a bus moves slowly as it picks up passengers, which obstructs all other vehicles behind it. In the multiple lane case, cars can overtake the slow-moving buses and prevent the platoon formation ($0 < \rho < 0.12$).

For the second region ($0.12 < \rho < 0.66$), traffic flow in the two lanes gets coupled due to cars changing lanes. As platoons form behind slow-moving buses, they impede the movement of cars in the first lane. When these cars switch lanes, they also slow down vehicles in the second lane (see spatio-

temporal diagram, Fig. 1.9). On the other hand, we also see that these cars can utilize gaps between platoons. Platoons in this multi-lane model do not appear as a single clump of closely spaced vehicles followed by a large gap of space. Instead, many smaller platoons form, with overtaking cars utilizing the gaps between platoons. However, this region does not exist for the case of $\alpha \leq 10^{-3}$, as the arrival rates of passengers is not high enough to cause the platoon formation. Thus, in the absence of platoons, the speeds of buses and cars are similar, and we recover dynamics similar to the NaSch model.

The remaining region ($\rho > 0.66$) marks the dissolution of the platoons in the bus lane. In this region, the density has increased enough such that both lanes move at similar speeds. As a result, the effect of bus pickups is indistinguishable from the congested dynamics of the NaSch model. We also observe this phase in the case of a crossover transition occurring when $\rho < 0.66$ (such is the case when $\alpha = [10^{-2}, 10^{-3}]$).

In this multi-lane scenario, stop durations increase when buses are allowed to wait for passengers. This slow-down has a two-fold effect: it reduces overall platoon movement; and it impedes the adjacent lane as more cars move out of the bus lane.

1.3.3 Periodic stops

In the previous sections, passengers may board and alight on any segment of the road. Such a scenario is common in cities like Manila and Jakarta, whereas having designated bus stops is a more common occurrence world-wide. As such, it is natural to take accessibility into account when designing stop locations. Closer spacing between stops can improve accessibility for commuters. But when spaced too closely, stops slow down transit travel speeds [49].

The decentralization of bus franchise operators in cities like Manila makes it challenging to plan for different aspects of a transit system. In routes that

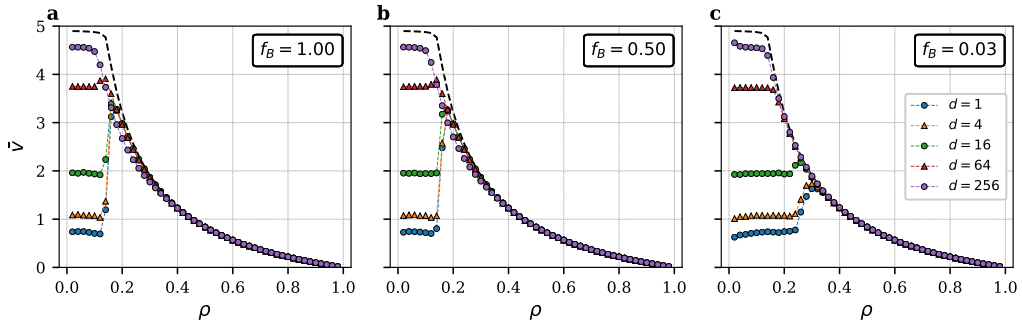

Figure 1.10: Speed vs. density plots for a single-lane model with varying station separation distances $d \in [1, 4, 16, 64, 256]$. Black dashed lines are for the baseline NaSch model with no buses. Spacing stops further apart increases speeds of vehicles on the system by mitigating platoon formation.

do not have sufficient demand, buses will be infrequent. Uncertainty in arrival times of buses in such cases can lead to either long waiting times, as well as a reduction of demand as commuters search for a more reliable alternative. Implementing bus timetables can help manage passenger expectations and establish the reliability of the transit system.

In comparing realizations with different distances between designated stops d, we assumed the same total number of passengers. Thus, we fixed the value of the pedestrian arrival rate for the entire system $A = \alpha \left(\frac{L}{d}\right)$ per timestep. With fewer stops (larger d), the arrival of passengers at each stop, α, rises commensurately. We ran simulations for $d \in [1, 4, 16, 64, 256]$, and fixed $A = \alpha_{\max} \frac{L}{d_{\max}}$, such that when $d = 256$, $\alpha = 0.1$. A maximum value of $d = 256$ is specifically chosen for our cell size of 5.5 m, as this corresponds to a 1.408 km distance between stops. We choose this based on the assumption that at most, a pedestrian would be forced to walk 704 m for a destination in between two stops. To account for our choice of $d_{\max}$, we use $L = 5120$ and set $A = 2$ for simulations in this section. For comparison, we will look at the cases ranging from all buses ($f_B = 1$) to a low bus volume, Philippines-inspired case ($f_B = 0.03$) [16].

First, we look at the single-lane case of an all-bus system. Increasing the distances between stops improves the overall speed of the system for $\rho < 0.12$ (Fig. 1.10a and 1.10b). The inflection point at $\rho \approx 0.12$ corresponds to the crossover transition from platooning to a non-platooned phase. Configurations with $d \leq 64$ exhibit an inflection at $\rho \approx 0.12$, a crossover transition which is absent in farther-spaced stops ($d = 256$). We observe similar behavior in the $f_B = 0.03$ case (Fig. 1.10c), where system speeds increase with station separation distances, but with two key differences from the all-bus ($f_B = 1$) case. Only $d = 1$ and $d = 4$ have crossover transitions that occur at higher densities, and the case of $d = 256$ does not have slower system speeds than $d \leq 64$ for all densities.

We attribute the slowdown of vehicles in the case of $d = 256, f_B = 1$ to the high arrival rates of passengers, which form slow-moving jams near the stops where buses have a high probability of stopping. Due to our model design, buses do not wait for passengers at stops. Since the distance (and time) headways between succeeding buses are short, the high arrival rates of passengers would make succeeding buses stop at the station. If a bus can wait for multiple passengers, buses behind it would not have to stop often, which would improve the overall system speeds.

In general, we expect the average number of buses that stop to board passengers should be the same for all realizations of d. Allocating larger separation distances between stops reduces the frequency of stops made by buses. This allows buses to maintain their top speeds for longer, and results in a higher system throughput.

System speeds, however, only paint half of the picture. Travel begins when the pedestrian leaves his origin, not at the moment the pedestrian boards a vehicle. Thus, it is necessary to include the time spent waiting for public transport. We find that the spacing of stops matters most in medium to low density regimes ($\rho < 0.2$ for $f \in [0.25, 0.5, 1]$, $\rho < 0.32$ for $f_B = 0.03$, Fig. 1.11). Having stops close to each other increases waiting

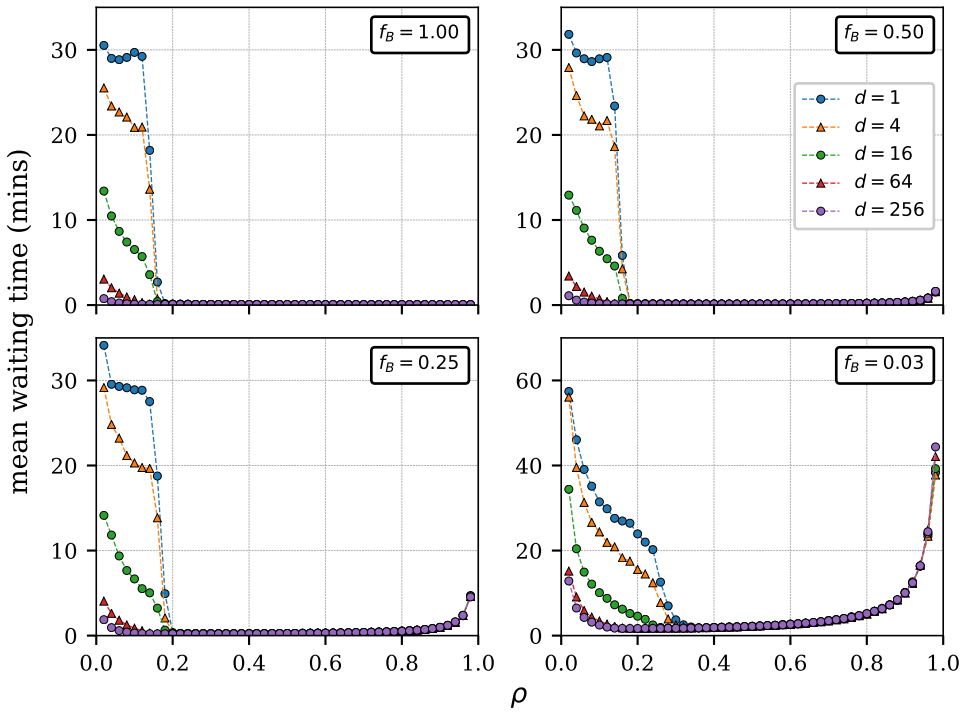

Figure 1.11: Mean waiting times at various densities for different station separation distances d. Station distances affect waiting times at middle to low density values ($\rho < 0.32$). Longest waiting times can be found at low densities (all f_B), and at high densities for mixed traffic ($f_B < 1$).

times of passengers, more so at low densities. Because of platooning, bus delays lead to more waiting passengers, which in turn causes further delays. Thus, waiting times can reach more than 30 minutes when buses are allowed to stop anywhere ($d \in [1, 4]$). Increasing stops separation therefore results to faster system speeds and shorter waiting times.

For low densities, passenger waiting times increase with decreasing system density. Infrequent bus arrivals force passengers to wait longer. Spacing stops close together exacerbate platooning effects — even spacing of bus headways are not maintained, and stopping occurs more frequently. This also leads to inefficiencies, as buses may arrive just after the preceding bus finished

boarding. In situations where public transport is deficient by way of long waiting times, having a bus timetable would help. A communicated schedule allows passengers to adjust their arrival at the station accordingly, effectively lowering expected waiting times.

For higher densities ($\rho > 0.16$), waiting times become constant for the all-bus case (Fig. 1.11, $f_B = 1$). Furthermore, the wait is just one timestep: the pedestrian gets to ride the bus one timestep after getting to the sidewalk. While this looks good from the point of view of passengers, it is quite inefficient since some buses fail to pick up passengers. We set the net pedestrian arrival rate $A = 2$, but the number of buses in the system for $\rho > 0.18$ is more than 921. Even if our model does not take into account bus passenger capacities, we can already see that reducing the number of buses may improve system speeds without loss of service reliability.

In mixed traffic settings, waiting times do not remain constant with increasing vehicle density. At sufficiently high road densities, waiting times also increase. Increased waiting times at higher road densities are an indication of insufficient supply of buses. The lack of buses, coupled with the slow speeds of vehicles on the road, lead to prolonged waiting times as congestion gets worse (Fig. 1.11, $f_B = 0.03$).

Allowing passengers to board and alight vehicles at any point is an attractive idea. The convenience of being able to go directly from your origin to your destination is the value proposition of owning personal vehicles or using taxis, ride-sharing services, and forms of para-transit such as tricycles, pedicabs, and *habal-habal*, which are prevalent in the Philippines. This mentality bleeds into other forms of public transportation such as buses, Jeepneys, and AUVs, which do not have designated stops. Our work supports the notion of planning out stops that are spaced further apart, but we must also take into account the added time and effort to the commute of a passenger whose stops would lie in between two stops.

The growing popularity of ride-sharing services also causes problems. Al-

though these have infrequent stops, their point-to-point nature is similar to a system of buses that do not have well spaced stops. Entire fleets of these vehicles can end up clogging roads and side-streets while they wait for their next booking. If left unchecked, the sheer number of ride-sharing vehicles waiting for passengers can end up negating any benefits from a well designed public transport system.

1.4 Summary

We created a hybrid model that combines features found in the Nagel-Schreckenberg and Bus Route Models. In this model, we observed three phases of traffic: homogeneous free flow, homogeneous congestion, and platooning. While phase transitions in the regime of $0 < \alpha < 1$ were absent, we observed crossovers characterized by $\rho^*(\alpha)$ and $\alpha^*(\rho)$. This crossover is distinct from the phase transition described by ρ_{crit}, which marks the transition from homogeneous free flow to homogeneous congested traffic.

Our simulations show that decreasing passenger arrival rates, and increasing bus densities induces a crossover from platooning to homogeneous phases. Moreover, we found an exponential scaling relation between the critical passenger arrival rate and density. The platooning crossover transition was observed up until $\rho \approx \frac{2}{3}$, which marks the phase transition from free flow to congestion for $v_{\max} = 1$.

The model also tackles bus stop placement and its impact on transit speeds and passenger waiting times. Spacing out stops helps greatly in improving transit speeds, as this minimizes the negative impact of platooning and congestion near bus stops. From a waiting commuter's perspective, spacing stops closer together increases waiting times. Both of the above-mentioned effects are relevant at low to medium densities ($\rho < 0.32$). These show that increased accessibility and convenience associated with the ability to hail a bus at any part of the road comes at the cost of overall system

efficiency and reliability. Lastly, our simulations on bus stop placement show the worst thing that can happen to a transit system is reduced availability of buses at high vehicle densities; this is when transit speeds are slowest, and waiting times of commuters are highest.

Our use of a simple model comes with some caveats; we do not take into account multiple passenger arrivals on sidewalk cells, finite bus capacities, and alighting scenarios. In the limit that buses are not being filled to capacity, waiting times measured are an upper limit, since all subsequent pedestrians arriving at a cell will always have shorter waiting times than the first pedestrian present. When buses are at capacity, waiting times are expected to be longer in general. This model might also underestimate the impact of bus stops on traffic flow, as buses will realistically wait for more than 5 seconds (or more than 3 simulation timesteps) for boarding and alighting passengers.

Chapter 2

Public Transportation in Metro Manila: Travel Characteristics and Trip Distributions

2.1 Introduction

As the population in urbanized areas grow, transportation infrastructure becomes increasingly burdened by the sheer number of vehicles, which inevitably leads to the collapse of the system, resulting in widespread traffic jams [50]. Cities need to develop a good public transportation system to manage the ever growing volume of traffic in a city, prevent traffic jams, and cater to the transportation needs of commuters. Doing so requires an understanding of how people utilize public transportation to go about their daily routines in cities.

While the traffic models in Chap. 1 describe the dynamics of vehicles on the road, they do not explain how vehicles end up on the road in the first place. Vehicles serve as modes of transportation in getting people from one place to another. Modeling the demand of traffic is an important aspect of the transportation model, and allows us to predict how roads will be used. Traffic demand is usually measured by performing Origin–Destination (OD) demand surveys. An OD demand survey gives an idea of the transportation needs of

citizens, and usually serve as the basis of traffic models. Traditional methods of obtaining these matrices include roadside and household surveys, which are expensive and have limited data collection capability, and are quickly outdated by ongoing urban development.

As such, there have been studies that attempt to use the increased availability of data sources to infer OD demands from data sources. One such approach involves Boarding–Alighting surveys which reflect the locations that people use public transportation systems. These are typically collected en masse through smart card systems, from which riders travel patterns can be extracted [51, 52]. Data from smart card systems can also be used to derive an OD estimate [53–55]. Other approaches also involve the use of call detail records (CDR) from telecommunications data [56–58], and geolocated Twitter posts [59].

While these data-based methods of obtaining Origin–Destination matrices improve over the traditional methods, some sampling bias may be expected in the context of the country the method is being applied to. In the Philippines, which has 113 million mobile cellular subscriptions as of 2016 [60], texting dominates calls as the mode of communication. Internet usage estimates in 2015 is at 41.4 million people (40.6% of the population) [60], but the expensive rates for mobile internet data also means that using platforms such as Twitter as a basis for an Origin–Destination matrix fails to capture the segment of the population that is unable to afford mobile internet data. And while the adoption of the Beep card smart ticketing system for the three rail lines and a handful of bus routes in Metro Manila is a potential source of smart ticket data, the current lack of coverage for many routes limits the usefulness of this data source. Finally, the use of data from social media platforms like Twitter could only represent those commuters with access to the internet.

The Department of Transportation commissioned the Metro Manila Route Transit Rationalization Study [61, 62] (RTRS) to understand how the cities

public transport routes were being utilized, as well as to propose new routes for the city. As a part of this study, RTRS conducted OD and BA surveys. Traffic analysis will typically divide a city into traffic zones, and OD demands give the flow of people between two zones, at different times of the day, and which modes of transportation was used. In the case of RTRS, the BA dataset are collected manually by surveyors riding bus lines, but lacks end-to-end tracking of commuters — this makes the analysis of travel patterns much more difficult. Both the OD and BA datasets also suffer from limitations brought about by the survey methodology (the survey was conducted only on PUBs (buses) and PUJs (jeepneys)) and misses out on traffic patterns from other modes of transportation.

In my MS thesis work [63], I explored the possibility of estimating road usage by using locations of road establishments and their proximity to points along roads to quantify a "potential" value for these roads. We have extended this idea to create a model of the spatial residential distribution of students, based solely on school enrollment data and total population census at the barangay level [64]. Since the problem of transport usually starts from the residence (origin) to points of interest on the map, such as schools, malls, and commercial spaces (destinations), leveraging available data sources can potentially provide a robust, up-to-date source of OD data in the absence of more rigorous data collection methods.

This work explores the public transportation data of Metro Manila. We measure differences in the OD demands and BA survey values in different traffic zones to identify public transport deficient zones and transit hubs. We also compare the differences between driving and the public transportation network, and show how commutes can change with the implementation of new proposed routes. Using a gravity model approach, we estimate the origin demands from empirical destination demand data in Metro Manila and show that this provides a good estimate of the zone-to-zone OD demands. Finally, we use the gravity model to generate OD demands using school enrollment

and barangay-level census data to capture a comprehensive OD demand that is not constrained to those that ride PUBs and PUJs.

2.2 Methods

2.2.1 School enrollment dataset

The Department of Education (DepEd) dataset used in this work includes school coordinates, population broken down by year level from 2012-2016 for all public schools, and school IDs. We use the 2016 school enrollment data in our analysis. Schools without location data, erroneous coordinates, or misassigned DepEd regions were removed. A total of 758 public elementary and secondary schools within Metro Manila were used in this analysis. These schools account for 1,825,943 students that are within the public school (Grades 1-12) system.

2.2.2 Population dataset

Barangay-level shape and population data for 2015 from the Philippine Statistics Authority [65] were also used for visualization and data cleaning. The population data is aggregated by sex and age; this makes it easy to narrow down the portion of the population for which we are estimating an OD for. We filter the population from ages 6–23 (Grade 1 to 5th year college) to encompass traffic due to educational establishments. Out-of-school youths and those that do not proceed to tertiary education are unaccounted for in our analysis.

2.2.3 RTRS dataset

The Metro Manila Road Transport Rationalization Survey (RTRS) study was conducted in 2013 by the Department of Transportation (DoTr) that sought to create an robust inventory of transit supply and passenger demand

data. The goal of the project was to develop an evidence based approach of optimizing Metro Manila's transit network based entirely on the passenger demand of the region [61]. A follow-up study in 2015 (aptly named RTRS2) expanded upon the data collected by the RTRS [62]. The entire RTRS study covers Metro Manila and provinces in its immediate vicinity (Fig. 2.1). This work uses three datasets from the RTRS study: the public transportation routes, an Origin-Destination (OD) demand and Boarding and Alighting (BA) survey.

Origin-Destination and Boarding-Alighting surveys

An OD demand survey aims to measure the number of commuters whose trips start and terminate at origin-destination pair, regardless of transportation mode. The raw survey data were processed as follows [61]:

- For each route, a Public Transportation Frequency survey was conducted to determine the frequency of vehicles operating on the route. The surveys also noted the number of passengers in each vehicle.
- The total number of passengers along the route is given by multiplying the number of vehicles and the number of passengers per vehicle.
- Under the assumption that the OD survey respondents are representative of the rest of the day, survey respondents are weighted such that the weighted count of the number respondents is the total number of passengers along the route.

We can express the survey process of the RTRS in the following way: the total demand for an origin–destination pair can be written as

$$T_{od} = \sum_{\text{route} \in od} \text{weight}_{\text{route},od} (\text{survey count})_{\text{route}}, \tag{2.1}$$

where the weight used for routes is

$$\text{weight}_{\text{route},od} = \frac{(\text{num vehicles})(\text{num passenger per vehicle})}{(\text{num respondents})_{\text{route}}} \tag{2.2}$$

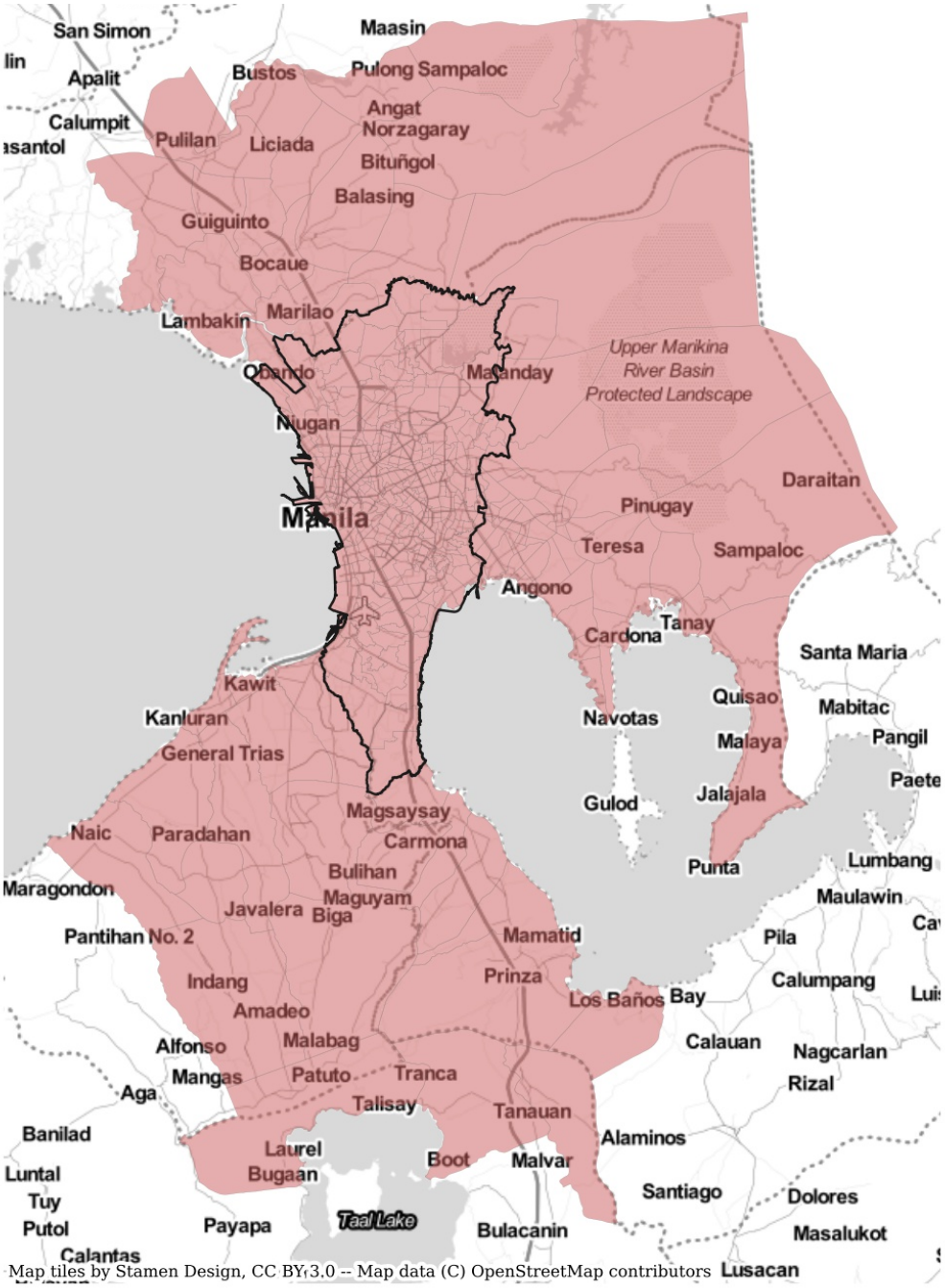

Figure 2.1: Map of all the zones used in the RTRS study. We focus on Metro Manila (black border) as it has well mapped roads and public transportation routes.

Ideally, the transportation demand T_{od} between an origin zone o and a destination zone d should include all modes of transportation, such as by private vehicle, Public Utility Vehicle (PUV), and walking. However, due to the limitations in data collection by the RTRS survey, the transportation demand reflected by the OD survey only accounts for PUJ and PUB movements.

The BA survey is a collection of geolocated points where commuters board or alight public transport. This includes 685 Public Utility Jeepney (PUJ) and 84 Public Utility Bus (PUB) routes. The raw survey data were processed similarly to the OD demand surveys:

- For each route, a Public Transportation Frequency survey was conducted to determine the frequency of vehicles operating on the route. The surveys also noted the number of passengers in each vehicle.

- Under the assumption that the BA survey respondents are representative of the rest of the day, survey respondents are weighted such that the sum of the product of weighted counts of the number respondents and frequency of buses along the route is the total number of passengers in each vehicle on the route.

Again, we can express the survey process of the RTRS in the following way: the boarding (or alighting) counts, denoted as X, at a stop can be written as

$$X = \sum_{\text{route}} \text{weight}_{\text{route},BA}(\text{survey count})_{\text{route}}, \tag{2.3}$$

where the weight used for routes is

$$\text{weight}_{\text{route},BA} = \frac{\text{num. pass. per vehicle}}{\sum(\text{respondents})(\text{bus freq.})}. \tag{2.4}$$

Both datasets are divided into 6 time periods: 06:00-9:00, 09:00-12:00, 12:00-15:00, 15:00-17:00, 17:00-19:00, and 19:00-22:00. Our analysis in this

work focuses on the 06:00-9:00 time period, which corresponds to when students typically begin their commute to schools. This time period also corresponds to the morning rush hour.

To simplify the analysis, the RTRS study divides Metro Manila and its immediate neighbors (Cavite, Laguna, Rizal, Bulacan) into 483 traffic analysis zones. Of the 483 zones, 381 are within Metro Manila. Each geolocated datapoint for the OD demand and BA surveys is associated with the zone it lies in.

Comparison of empirical OD and BA data

Comparison of the OD and BA datasets are done at the zone level. For a zone i, we compare the zonal origin demand $\mathcal{O}_i = \sum_d T_{id}$ with the zonal boarding counts $\mathcal{B}_i$, and the zonal destination demand $\mathcal{D}_i = \sum_o T_{oi}$ with the zonal alighting counts $\mathcal{A}_i$. These numbers however are sensitive to area variations of zones. To account for area sensitivity, we need to rescale the zonal aggregates by a quantity that also scales with the area of a zone.

We define the zonal vehicle capacity $\mathcal{C}$ as the maximum number of vehicles that can fit on the road network of the zone. Road networks obtained from OpenStreetMaps (OSM) [66] include roadways with the following tags: `motorway`, `trunk`, `primary`, `secondary`, `tertiary`. These combination of tags best captures roads used by public transportation routes, but include roads that are not covered by public transportation. For each roadway, OSM provides information on the length l and number of lanes n of the roadway. In cases where the number of lanes is not explicitly tagged, it is assumed that the road has one lane. We estimate the zonal vehicle capacity as $\mathcal{C}_i = \rho_{\text{jam}} \sum_{j \in i} l_j n_j$, where $\rho_{\text{jam}} = 166.67\,\text{veh/km}$ assuming that in a traffic jam, a vehicle takes up 6 m of space on the road. The zonal vehicle capacity includes roads within 50m from the borders of a zone.

2.2.4 Trip distribution

Trip distribution models are used to predict the transportation demand T_{od} (also known as the OD demand), which is the number of trips that occur between each pair of origin and destination zones. Predictions are calculated using the predicted number of trips coming from an origin and arriving at the destination (also known as *trip production* and *trip attraction* in the traffic engineering literature).

Gravity model

The gravity model gets its name from Newton's Law of Gravitation, which is of the form

$$F_{ij} = G\frac{m_i m_j}{r^2}, \tag{2.5}$$

and states that the attractive force between two masses i and j increase proportionately to their masses, and inversely as the distance between them. In a similar vein, the gravity model often takes the general form

$$T_{ij} = \frac{m_i^\alpha m_j^\beta}{f(r_{ij})}, \tag{2.6}$$

where the masses m_i represent populations, and the cost function $f(r_{ij})$ is chosen to fit the data. This model states that two zones are likely to have increased population movements proportionate to their populations, and inverse to the cost of movement between them. In transportation modeling, values of $\alpha = \beta = 1$ are frequently chosen, leaving the parameter to the cost function. Gravity models have seen widespread use from migration and mobility [67, 68], economics [69], communications [70], to international trade [71].

Modified gravity model

Using residential distribution data (census) and occupancy of establishments (such as schools, malls, etc), we generate an OD estimate for Metro Manila.

Using the same method we used for school data[1] [64], we modify the gravity model to estimate the flow of people from their place of residence to their workplace. We define the inferred transportation demand T_{od} between a pair of origin o and destination d as

$$T_{od} = kP_dP_o\phi_{od}(r_{od}), \tag{2.7}$$

where k is a coupling strength, P denotes the population of the origin and destination, and $\phi_{od}(r)$ is some distance dependent function. In our case, we know for certain the demand at our destinations, which corresponds to the number of enrolled students in the school. Thus, we have the constraint $P_d = \sum_o T_{od}$. Setting $k = \frac{1}{\sum_o P_o\phi_{od}}$ satisfies this constraint, and with a bit of rearrangement of terms, we have

$$T_{od} = P_d\frac{P_o\phi_{od}(r_{od})}{\sum_o P_o\phi_{od}(r_{od})}. \tag{2.8}$$

The inferred demand can therefore be interpreted as redistributing the population at the destination to their origins, relative to the total resident population at the origin, weighted by some distance dependent cost. This work uses a distance dependent function of the form

$$\phi_{od}(r) = \min\left\{1, \left(\frac{r}{r_c}\right)^{-\alpha}\right\}. \tag{2.9}$$

In general, the cost function $\phi_{od}(r)$ may be a function of travel distance, travel time, apartment and transportation costs. For this work, we choose r to be the distance between the origin-destination pair. We use actual road distances along the road network (rather than using the euclidean distance between the origin and destination). This form of the cost function states that individuals are less likely to choose destinations that are much further

[1]Part of this work appears in "Rubio, L. J. M., Dailisan, D. N., Osorio, M. J. P., David, C. C., and Lim, M. T. (2019). *Modeling the residential distribution of enrolled students to assess boundary-induced disparities in public school access.* PLOS ONE, 14(10), 1–15. https://doi.org/10.1371/journal.pone.0222766"

from where they are. Setting r_c incorporates the preference for locations that are reasonable close for the commuter and also ensures that the cost function remains finite at small distances.

2.2.5 Calculation of route travel characteristics

The OD surveys lack the actual route itineraries of commuters. We have to infer these itineraries ourselves to obtain an idea of the typical travel distances of commutes in an origin-destination pair. This work considers two types of transportation networks: *driving* and *public transportation*.

We represent these networks as a graph [72]. A graph $G(V, E)$ is defined as the set of vertices $V = \{v_1, v_2, \ldots, v_n\}$ and edges $E \subseteq V \times V$. Each edge is denoted by a pair of vertices (u, v) where $u, v \in V$. Graphs can be undirected ($(u, v) \equiv (v, u)$) or directed ($(u, v) \not\equiv (v, u)$). Nodes u and v are considered neighbors if there is an edge (u, v) connecting them.

Driving network

The *driving* network represents the routes available to commuters that use personal modes of transport (e.g. cars, taxis, ride-hailing apps). In this network, commuters can use any road infrastructure that is available. We obtained this road network through OSM and (for the sake of computational efficiency) narrowed it down to these set of OSM `highway` keys: `primary`, `trunk`, `motorway`, `secondary`, `tertiary`. This network can be represented as a directed graph, where vertices represent road intersections, while the edges represent actual roads themselves.

Public transportation network

The *public transportation* network represents routes available to commuters when they use public modes of transport (e.g. buses, jeepneys, trains). These modes have predetermined routes, and in most countries (but not the Philippines), designated stops. Route and stop information were obtained through

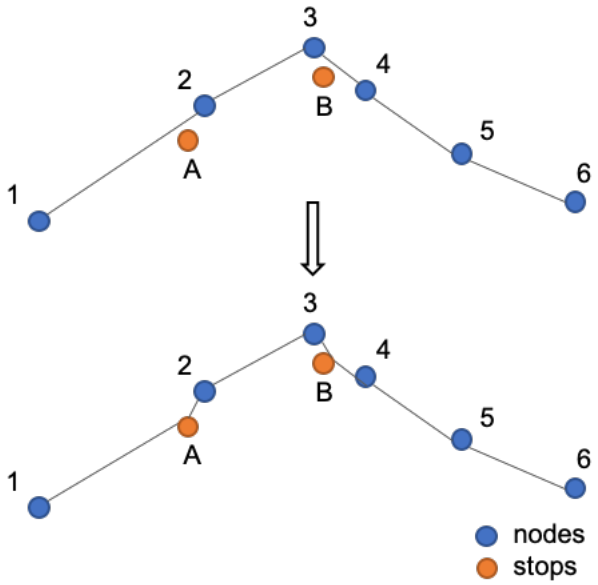

Figure 2.2: Consider a set of shape points $\{1, 2, 3, 4, 5, 6\}$ and stop points $\{A, B\}$. We match the stops to the shape by inserting them between the endpoints of the closest segments. This gives us a full sequence of $[1, A, 2, 3, B, 4, 5, 6]$.

General Transit Feed Specification (GTFS) feeds[2] obtained from the Department of Transportation (DoTr).

The RTRS GTFS feed contains a file that encodes routes as as a sequence of geolocated points that together form the `shape` or the route traversed by the PUV. A separate file contains information on the stops, which are associated with a `shape` id. We employ a *stop-to-shape* matching algorithm in order to insert the stop points into their corresponding shape point sequence (see Fig. 2.2). The `shape` is broken up into segments represented by subsequent shape point pairs. Stops are then inserted between the endpoints of the nearest line segment to form the new shape–stop point sequence.

Unlike the *driving* network graph, which was constructed exactly as how the real physical road network is configured, we need to employ an additional layer of abstraction when constructing the directed graph representation of

[2]https://developers.google.com/transit/gtfs

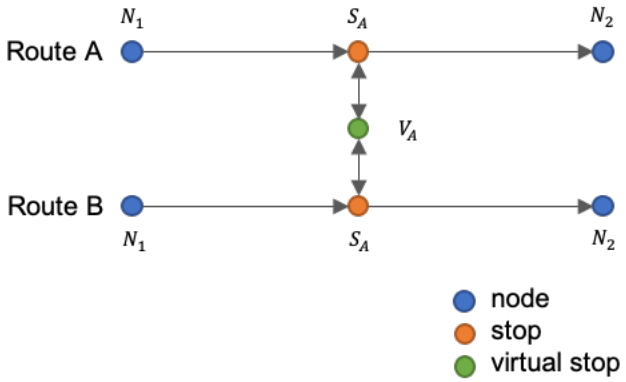

Figure 2.3: It is possible that nodes from two routes A and B share the same geographical locations. These are treated as separate vertices, and a virtual stop node V_A is added to connect the common stop nodes S_A.

the *public transportation* network. In the graph representation of this network, stops are used as vertices and portions of the route between subsequent stops are used as edges. Traversing this graph is analogous to traveling along the PUB/PUJ route. Route segment lengths and names were also stored as edge metadata. We then added an additional *virtual stop* vertex with a bi-directional connection to each stop; this edge represents *boarding* and *alighting* of a commuter (see Fig. 2.3). Virtual stops are also connected by an edge to other virtual stops within a distance of $R_{\text{walk}} = 500\,\text{m}$ from each other; this represents transfers to a different route does not share a stop with the commuter's current route. These *virtual stops* are a necessary abstraction used to find the shortest transit path later on.

Shortest path between an OD-pair

With the graph representations of the *driving* and *public transportation* road networks, we connected source (origin) and sink (destination) nodes to the graphs. Source and sink nodes were connected to the nearest vertex on the graph: intersections for the *driving* network and the *virtual stops* for the *public transportation* network. Shortest paths were calculated using Dijk-

stra's algorithm [73] on a weighted graph. The weights of edges $w(u, v)$ serve as path costs, and is typically taken to be the lengths of the road sections. Dijsktra's algorithm finds the shortest path from an origin node o to a destination node d in a graph. The way the algorithm works is outlined as follows:

1. Initialize a distance measurement of $x(o) = 0$ and $x(v_i) = \infty$ for all other nodes $v_i \neq o$. Mark all nodes as part of the unvisited set.

2. Set the node with the smallest d value in the unvisited set as the current node.

3. From the current node u, update $x(v)$ for all neighboring nodes u as $x(u) \leftarrow \min(x(u) + w(u, v), x(v))$. When $x(u)$ changes, we keep track of the node v that caused the change. Each of the neighboring nodes u is then removed from the unvisited set.

4. Repeat steps 2 and 3 until the destination d has been visited, or if the node with the smallest value in the unvisited set is $x = \infty$ (the destination is unreachable.)

Once Dijkstra's algorithm is finished, $x(d)$ is the shortest path length. The shortest path can be obtained by tracing back all of the nodes that led to d.

While the use of Dijkstra's algorithm on the *driving* network is straightforward, this isn't the case for the *public transportation network*, as the shortest path could contain unnecessary route transfers. To resolve this, we introduce two costs associated with boarding and walking transfers. The boarding cost discourages unrealistic hopping between two routes that have a common direction; this can also be interpreted as a flat minimum cost the commuter incurs (such as the minimum fare when looking for the cheapest routes), independent of distance travelled. We define the boarding cost as a distance:

$$w_{\text{board}} = \text{transit speed} \times \text{transfer time}, \tag{2.10}$$

where we use a transit speed of 30 kph and a transfer time of 5 mins, which translates to 2.5 km (for reference, the minimum fare is Php 9.00 for first 4 kms for PUJs, and Php 13.00 for the first 5 kms for PUBs). We also define the walking transfer cost as

$$w_{\text{walk}}(u, v) = C_{\text{walk}}||\vec{x}_u - \vec{x}_v||, \tag{2.11}$$

which just multiplies the distance between two virtual stops u and v by a factor of $C_{\text{walk}} = 5$. Setting this value of C_{walk} means we assume that commuters would prefer to walk up to 500 m than to board another PUV (this also eliminates shortest paths constructed by jumping through multiple PUV routes.) These two costs result in a more realistic travel itinerary for a commuter on the *public transportation* network.

From the shortest path travel itinerary, we can count the number of transfers in a commuting route. The number of transfers is a useful metric for evaluating the efficiency of a transportation network. We define a transfer as an instance of a change in mode of transportation; this can also be measured by subtracting 1 from the number of unique route names that appear in the public transportation itinerary. Additional transfers can incur added costs, takes up time, and affects the perceived comfort and convenience of a commute.

2.3 Results and Discussion

This section is organized as follows: we discuss generating estimates for the trip distributions of Metro Manila in Sec. 2.3.1. First, we use the gravity method to estimate origin demands from empirical destination demands. We then use school enrollment and census data to generate a completely synthetic estimate of the trip distribution attributed to public elementary and secondary schools in Metro Manila.

We then analyze the travel characteristics of public transportation in

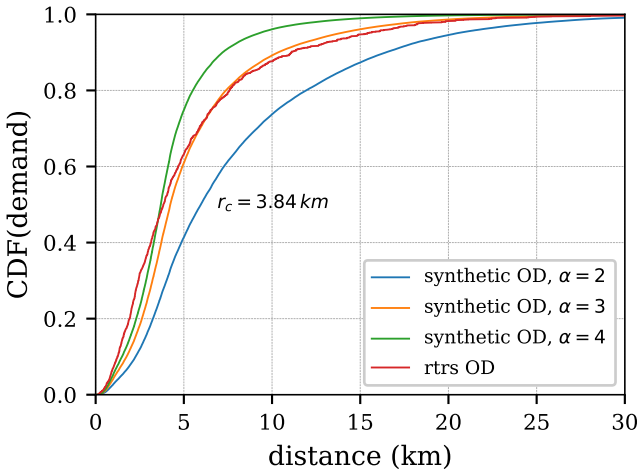

Figure 2.4: Distance distributions of empirical and synthetic demands.

Metro Manila in Sec. 2.3.2. We first compare Origin–Destination (OD) and Boarding–Alighting (BA) data to identify transport deficient zones. This is followed by an analysis of the driving network against the public transportation network in Metro Manila.

2.3.1 Travel Distributions

Generating synthetic trip distributions

The Origin Destination (OD) demands provide a count of the number of trips between an origin and destination pair. Using the aggregated zonal destination demands, we want to see if a gravity model can estimate the OD demands of the RTRS OD dataset. That is given D_i at zone i, we want to find N_{oi} (via Eq. 2.8) for all origins o, and all zones. For this, we will take the origins as barangays.

The gravity model however involves a two-parameter cost kernel (Eq. 2.9) that uses a cutoff distance r_c and a scaling exponent α. To find reasonable estimates of r_c and α from the data, we measured the distance distribution function of the empirical (RTRS) data (Fig. 2.4). From the distance dis-

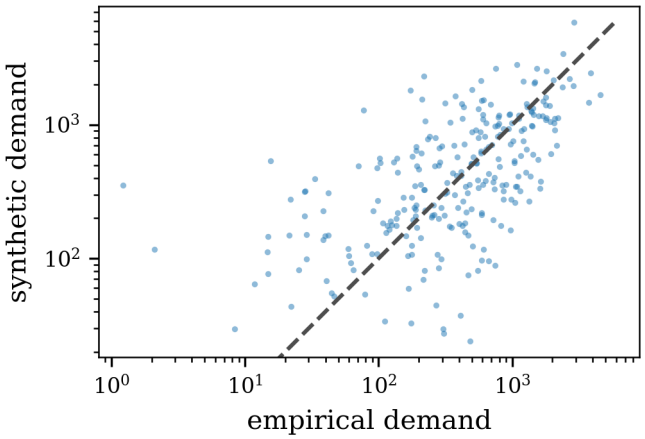

Figure 2.5: Correlation of empirical and synthetic zonal demand. Points represent zones, and the dashed line has a slope of one.

tribution, we will arbitrarily set r_c as the median distance of all trips. We then generate synthetic OD demands for three different values of the scaling parameter: $\alpha = 2$, $\alpha = 3$, and $\alpha = 4$. We find that using $r_c = 3.84\,\text{km}$ and $\alpha = 3$ generates a distance distribution which is similar to that of the RTRS data.

How does this synthetic OD compare to our empirical OD? A correlation plot comparing the aggregate origin demands at zones shows that the synthetic aggregate demand has some correlation to the empirical aggregate demand (Fig. 2.5 and 2.6); the synthetic method was able to roughly estimate the zonal origins from population data and zone distance, albeit with a noticeable amount of fluctuation. For individual trips however, empirical and synthetic demands appears to lack any correlation (Fig. 2.7). To further explore this lack of correlation, we look at how trips were distributed to different zones. Taking the destination zone with the highest demand, we visualize how the trip demands are distributed among the origin zones and compare these distributions for the empirical (RTRS) data and synthetic (gravity method) estimates (Fig. 2.8). While the synthetic estimates that the trip origins are generally close to the distribution, we see from the em-

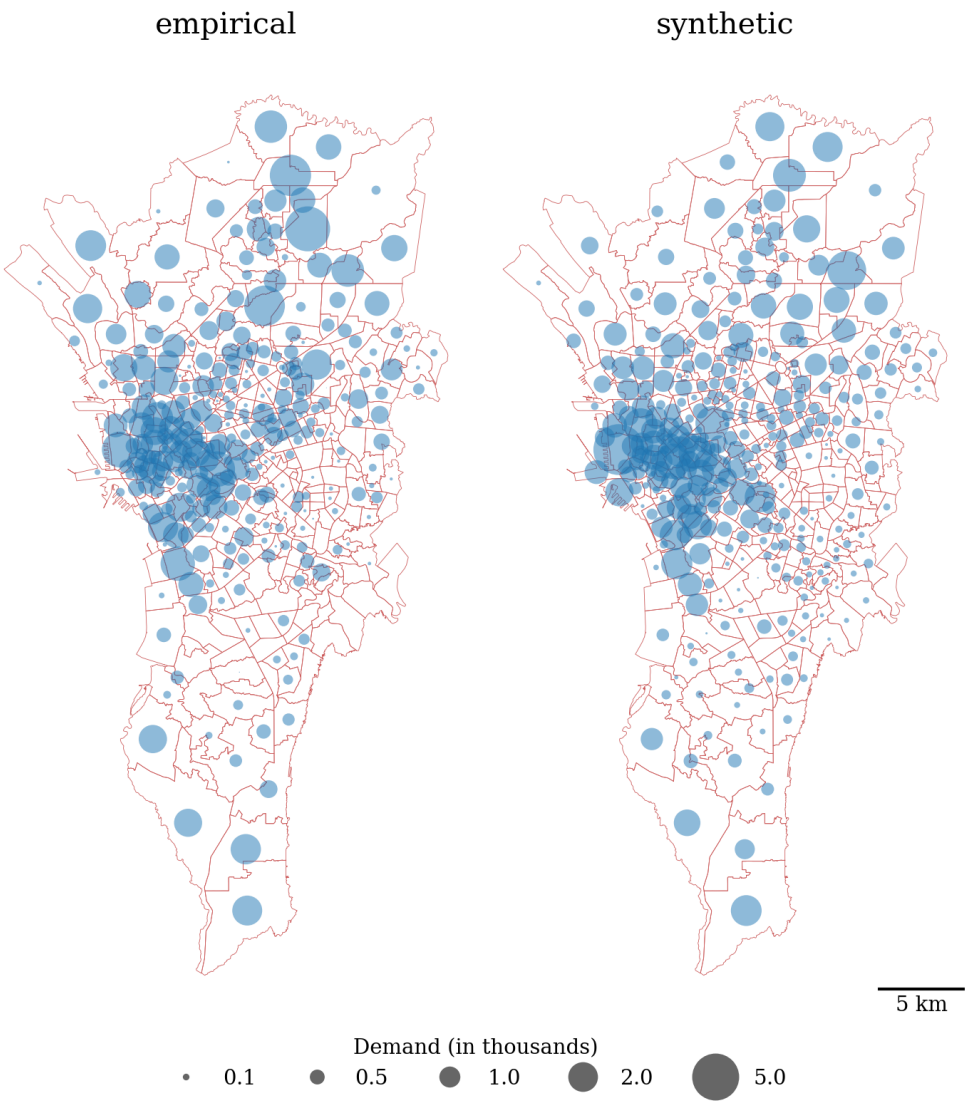

Figure 2.6: Maps of empirical and synthetic demand. Size of bubbles denote the demand in each zone.

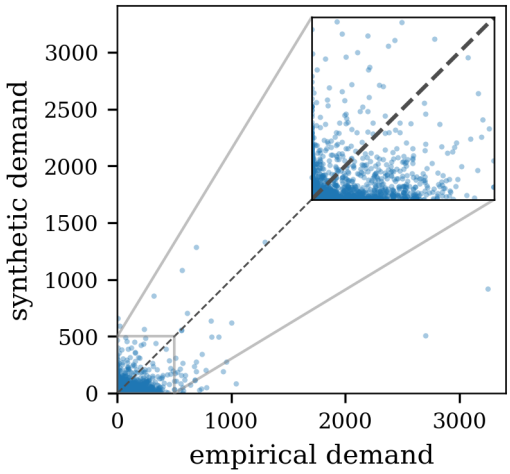

Figure 2.7: Correlation of empirical and synthetic trip demands. Points represent a trip between an OD pair. The inset shows demands of up to 500.

pirical demands that there are several trips that come from zones that are much further from the destination zone.

The gravity method describes the movement of large populations (typically between countries and large cities, down to the county level for the US); the granularity of individual barangays and buildings themselves with population sizes may be insufficient to approximate average behavior. This could explain why the gravity method estimates the aggregate trip demands between zones much better than at individual OD pairs. As we inspect the estimated demand for individual trips, we observe that there are instances when the empirical demands are located further away from the destination zone. This shows possible flaw in our implementation of the gravity method; we assumed that r_c and α (which can be interpreted as how far commuters are willing to travel) is uniform for all possible destination establishments (which are schools). Although we are working within a single type of purpose of travel, this can still be fairly heterogeneous since `Education` could be a mix of public and private schools, elementary, secondary, and tertiary

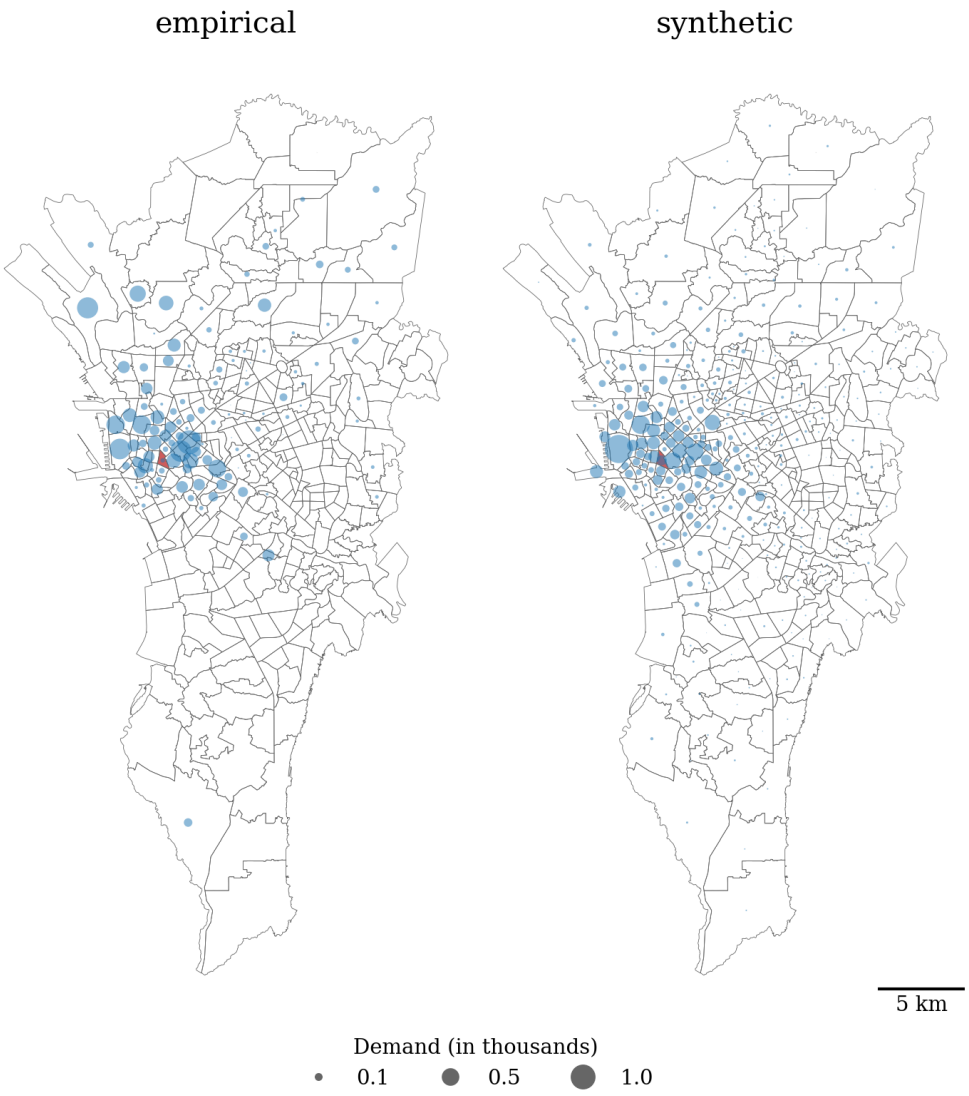

Figure 2.8: Origin demands for the zone with the highest destination demand (in red). The synthetic method estimates demand to be close to the destination, but the empirical data shows that there are zones far from the destination with non-zero demand.

level schools. For the case of origins, the choice of residence may also be constrained by living costs. Our model does not explicitly account for heterogeneity in financial status of individuals in the population. These factors may force individuals to tolerate endure longer commutes just to get to their destination, which suggests it may be better to use a distribution of r_c and α values, rather than a single number for the entire population.

The gravity model also assumes that population size acts as "mass"; there proportionately more travel between zones with high working (destination) and residential (origin) populations. We used census data to inform the model of the residential populations, but this comes with a caveat: census data does not necessarily reflect the actual residency of the population. This is a particularly important point for `Education` tagged trips, as students that reside in dormitories and boarding houses are not accurately captured by census data. The census usually captures the permanent addresses of citizens, even if a portion of the population have present addresses tied to short-term rentals. In fact, other problems emerge from the distinction between present and permanent addresses, an example of which is the distribution of the social aid in the time of COVID-19, as non-residents faced difficulty in getting social aid from local government units.

Simini et al. previously demonstrated that, on average, the gravity model is able to predict empirical OD demands between counties in the US (Fig. 2.9). In contrast to our zones, which are based on barangay delineations, counties are much larger in scale — comparable to provinces of the Philippines in terms of land area. Although counties are analogous to provinces in terms of land area, this is not necessarily a one-to-one comparison; there could be differences in how the population is distributed within these two areas. The OD demands observed in the work of Simini are also three orders of magnitude (up to 10^6 as compared to 10^3 in Metro Manila). It is possible that the data for Metro Manila is insufficient to demonstrate reasonable predictions using the gravity model.

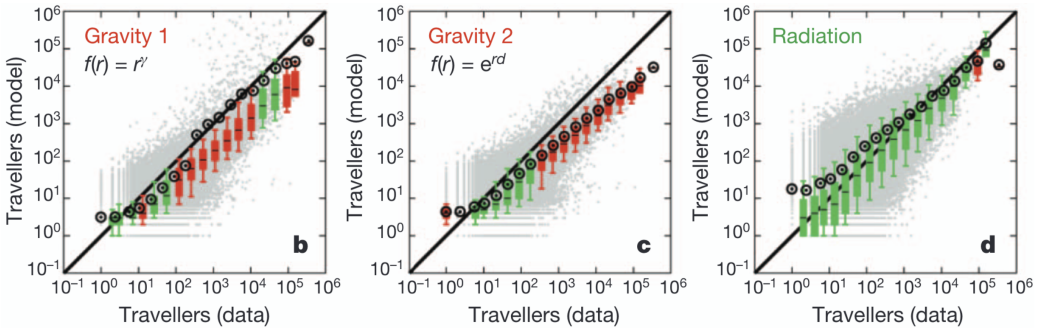

Figure 2.9: Comparison of empirical and modeled OD demands using other forms of cost function $\phi(r)$ for the gravity model, as well as a different model. This study uses OD trip demands between US counties (grey dots), and spans several orders of magnitude more than our analysis of Metro Manila (up to 10^6 as compared to 10^3 in Fig. 2.7). Figure taken from the work of Simini et al. [74].

Comparing destination demand with school enrollment

The OD demands measured by RTRS are tagged with the purpose of travel. We can filter out all demands that are tagged with `Education` to isolate travel to schools from the dataset. Furthermore, we take the OD demands for the 06:00-09:00 time period only. Students account for 184,004 (10.8%) of all commuter movement in this time period. However, compared to the number of enrolled students in the DepEd public school enrollment data, which is 1,822,398 from Grades 1-12, the RTRS data suggest that only 501,530 (27.5%) of all enrolled students actually use public transport in the course of an entire day. Figure 2.10 compares the zonal demands in the RTRS dataset with the zonal aggregates of school enrollment; clearly, the survey methodology of the RTRS does not capture the demographic of students from Grades 1-12, as reported by the Department of Education. This indicates either student commuters are under sampled due to some survey bias, or that public school students used modes of transportation other than PUBs and PUJs.

Since the enrollments differ from the commuter movements by an order of magnitude, we divide the enrollment data by a factor of 10 in all comparisons.

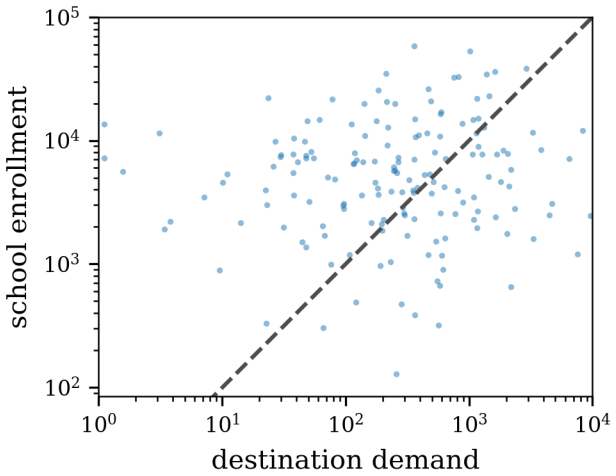

Figure 2.10: Correlation of destination demand and school enrollment. Values are rescaled by the total resident population of zones.

We find that there is little to no correlation between the travel demand measured by RTRS and the actual locations of schools; most `Education` related travel terminates in Manila (Fig. 2.11). While there is no way to determine the age demographic of the commuters surveyed by the RTRS, we speculate that these commuters are predominantly university students.

Estimate trip demands from school enrollment

We now use the gravity model to estimate the OD demand associated with primary and secondary schools from the DepEd dataset. The locations of schools and their total enrollment serve as the destination demand, and barangay centroids as the origin points. From the known school enrollment figures of students, we estimate how these students of different schools are distributed throughout the barangays in Metro Manila; this also corresponds to the origin demands.

Again, we use the empirical RTRS OD demands as a basis for the values of our model parameters, and find that with $r_c = 3.84\,\mathrm{km}$ and $\alpha = 3$, we obtain a synthetic OD demand with similar distance distributions as the

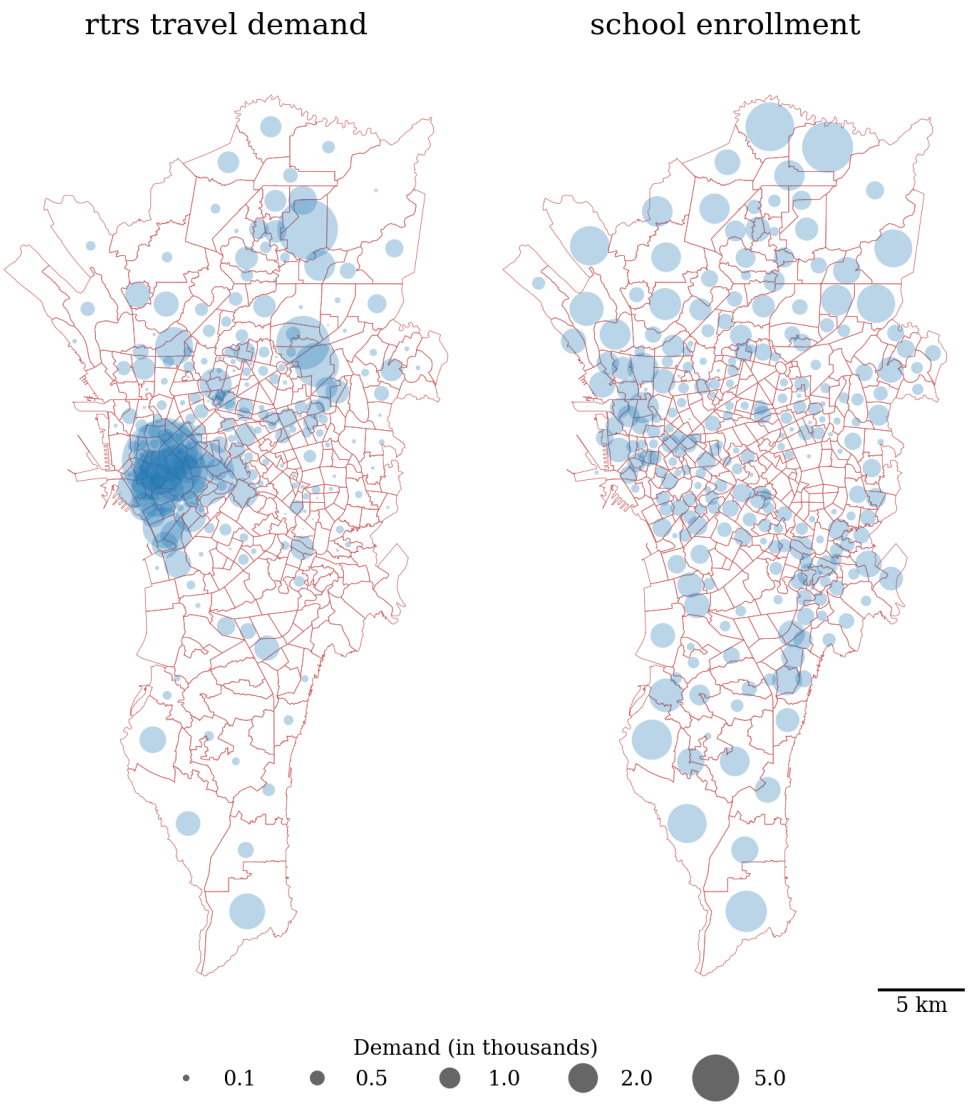

Figure 2.11: Comparison maps of destination demand and school enrollment. School enrollment demand is divided by 10 (since enrollments differ from destination demands by one order of magnitude.)

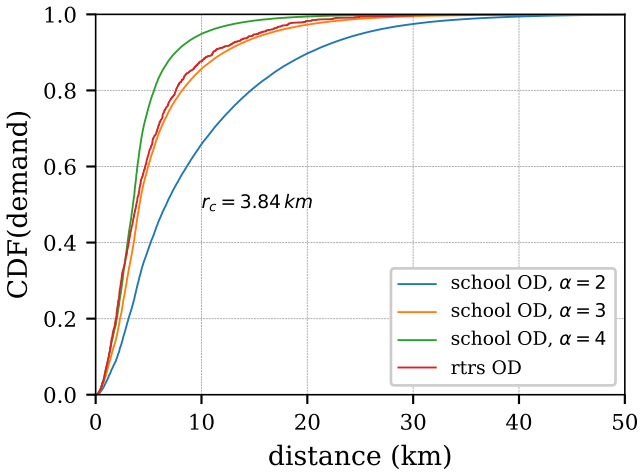

Figure 2.12: Distance distributions of empirical and synthetic school demands.

RTRS data (Fig. 2.12). We find that the gravity model estimates only 12.3% of movements are intra-zonal (movements within the same zone.) While small adjacent zones could fall within the $r_c = 3.84$, this still leaves the other half of students ($\sim 900,000$) needing some form of transportation. If the design process of public transportation neglects these students, their transportation needs will need to be fulfilled by other modes that may be less efficient than mass public transport.

What this tells us is that the RTRS survey missed out on a significant portion of the transportation demand of students in Metro Manila. These are the students that walk to school, ride tricycles, or are dropped off by their guardians via motorcycles or even cars. Even if we assume that intra-zonal movements account for students who walk or use tricycles/pedicabs, this leaves $\sim 1,600,000$ students that use some form of transportation. The RTRS study's sampling of commuters — those that use PUBs and PUJs — highlighted the utilization of the public transportation system, but prevented them from obtaining a comprehensive picture of the transport demand in Metro Manila. In sampling only the traffic demand of commuters already

using public transportation, they failed to capture the rest of the population that are not adequately served by the existing system, who then find alternative ways to meet their transportation needs. This ultimately hinders the development of a reliable public transit system that attracts commuters.

The gravity method may be unable to accurately estimate individual trip demands, but it provides a reasonable overview of traffic demands between zones. While this work focuses on school related trip demands, the gravity method is not limited to schools. Any type of establishment (which correlates with different types of travel purposes) can be used, provided that the occupancy of an establishment is known. Schools make a good test case since the already capture this data annually (UP does it every semester via CRS); government agencies too already have an avenue to collect such information when businesses register annually. However, different establishments and activities may have different associated r_c and α values that may have to determined.

2.3.2 Travel Characteristics

Comparison of empirical OD and BA data

We define the scaled difference between two counts X and Y as

$$\Delta(X, Y)_i = \frac{X_i - Y_i}{\mathcal{C}_i}. \tag{2.12}$$

To compare the zonal origin demand with boarding counts, this gives us $\Delta(\mathcal{O}, \mathcal{B})_i = \frac{\mathcal{O}_i - \mathcal{B}_i}{\mathcal{C}_i}$. Similarly, for zonal destination demand and alighting counts, we have $\Delta(\mathcal{D}, \mathcal{A})_i = \frac{\mathcal{D}_i - \mathcal{A}_i}{\mathcal{C}_i}$. The magnitude of Δ reflects the additional road infrastructure needed to accommodate the disparity, assuming one vehicle to one commuter; $\Delta = 1$ would need double the current capacity (by adding additional road infrastructure, or increasing vehicle occupancy) of the zone to accommodate the needs of commuters. This scaled difference shows the disparity between demand and actual usage of different zones in

the city. Zones with $\Delta > 1$ are 12.0% (10.6%) of zones, and 8.49% (8.70%) for $\Delta < -1$, for origin–boardings (destination–alightings) (Fig. 2.13).

For both the origin–boarding and destination–alighting pairs, a more positive Δ can be attributed to zones that use modes outside the coverage of the BA survey, such as private vehicles, taxis, tricycles, and walking. This is an indication of the lack or insufficient access to public transportation within these zones. While the OD survey does not capture private vehicle usage, residents in zones that are underserved actually have an incentive to own personal vehicles in order to have more convenient commutes. Through our comparison, we find that several underserved zones coincide with residential areas, such as Signal Village in Taguig, Batasan Hills, Ayala Heights, UP Village, Project 6, New Manila, and Valle Verde (Fig. 2.14). In certain cases, such as Project 6 and UP Village, public transportation lies right outside the borders of the zone, but the inner parts of the zone are not part of public transportation routes.

Conversely, areas with more negative Δ are transit zones, where commuters switch lines to continue on to their final destinations. Some of these transit zones are associated with endpoints and intersections of rail lines (North Avenue, Aurora Blvd, Taft Avenue, Recto stations of the MRT and LRT), or passenger terminals (FTI in Taguig, SM Mall of Asia). Other zones along EDSA that show up as transit zones are Guadalupe, SM Makati, and Ortigas Center which contains Robinsons Galleria and SM Megamall. However, we find that presence of terminals do not necessarily make a zone a transit hub. Transit hubs are areas that are associated with congestion, as public vehicles that load and unload passengers stop on the curb, which impedes traffic flow. In the case of transit hubs that also act as terminals, improper infrastructure design also limits the capacity of such terminals, which in turn causes a buildup of traffic on adjacent roads. Lastly, while the BA survey cannot track commuters at the individual level, high incidence of transit zones could indicate an inconvenient and inefficient commute (exceeding

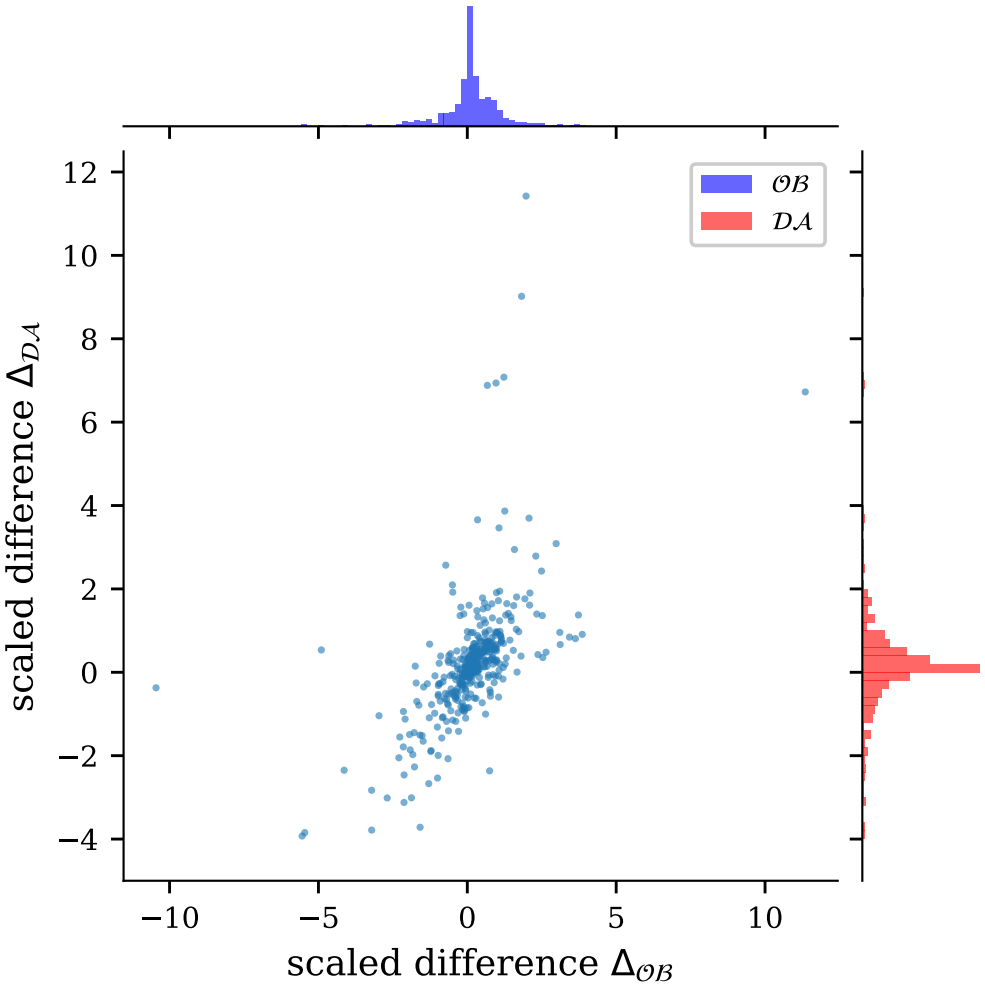

Figure 2.13: Joint distribution of the scaled difference of origin demand and boarding counts ($\Delta_{\mathcal{OB},i}$) and destination demand and alighting counts ($\Delta_{\mathcal{DA},i}$) of zones.

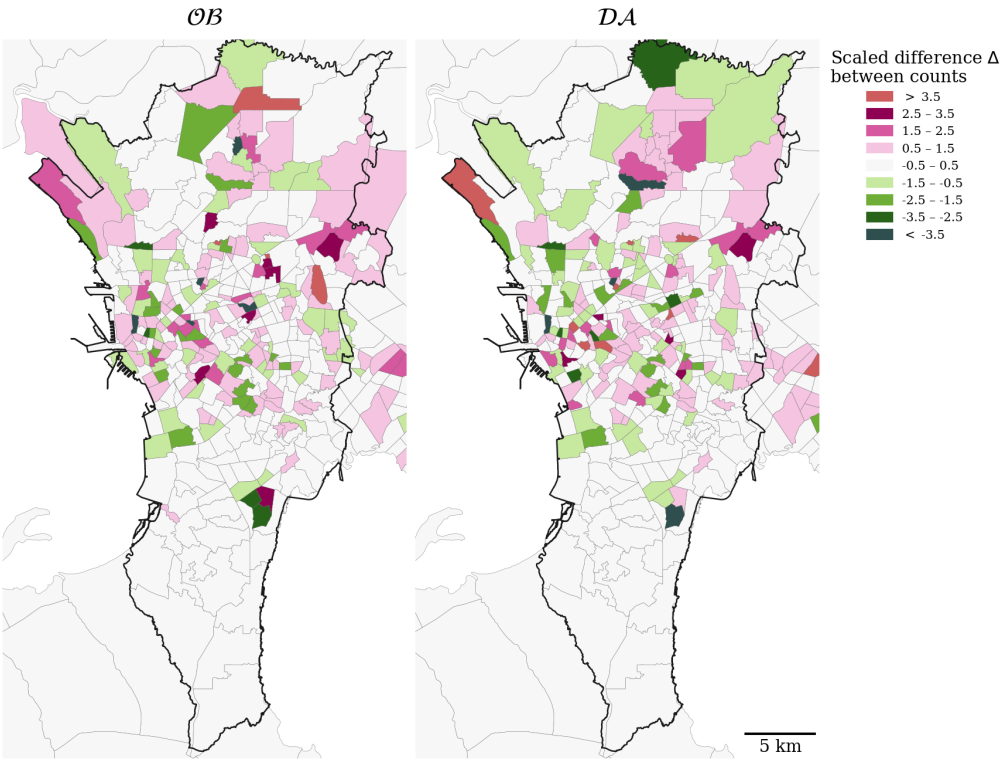

Figure 2.14: Scaled difference of origin demand and boarding counts ($\Delta_{\mathcal{OB},i}$, left) and destination demand and alighting counts ($\Delta_{\mathcal{DA},i}$, right). Pinker zones mean its constituents are unable to use public transportation to start/end their trips (underserved), while greener zones have a lot of changes between transit lines (transit hubs). Values of $|\Delta| > 3.5$ are omitted from the range, and shown in dark green and red colors.

two transfers), particularly in context of the Philippine climate.

Zones with $\Delta > 1$ are areas that should be prioritized for new public transportation routes. These comprise 81 zones with an excess of 175,226 and 226,607 trips affected at their origins and destinations, respectively. As most of these underserved zones are associated with residential areas, placing accessible public transportation option is one of the ways private car use can be curbed.

Zones with $\Delta < -1$ on the other hand are at higher risk of infection in an outbreak of disease, with 55 zones affected. These zones can serve as a gathering point of people from various origins and destinations, and can serve to spread the disease more rapidly. With the aggregated datasets provided, it is difficult to ascertain how many individuals are at risk in these zones; however there are 566,841 boardings and 555,703 alightings in these zones. People inevitably gather in these zones as a consequence of our transport infrastructure, which in turn endangers the population as the risk of spreading the infection increases. Existing facilities in some transit hubs like PUV terminals and train stations are regularly overcrowded, makes social distancing difficult, and decreases the effectivity of any form of community quarantine.

Driving vs. Public Transportation

Would you rather drive a car, or use public transportation (PT) in your commute? This basic question highlights an important aspect of transportation. We know that traffic is primarily caused by high densities of vehicles and that PT is more efficient at moving people than private vehicles. And yet private cars make up 67% [16] of traffic along radial and circumferential roads in Metro Manila.

How does commuting using the public transportation system compare to driving? We use data from the 523 Public Utility Jeepney (PUJ) and Public Utility Bus (PUB) routes to characterize PT in Metro Manila. The trip distance distributions of these two modes show that trip distances are shorter

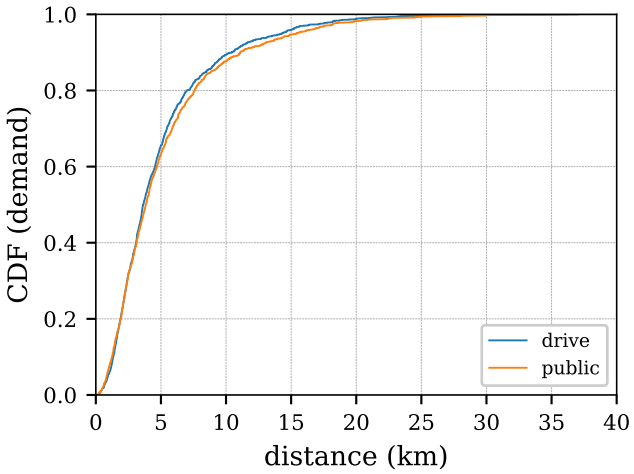

Figure 2.15: Trip distance distributions of the drive and public networks. The public network has slightly longer trip distances ($r_c = 3.84\,\text{km}$) compared to driving ($r_c = 3.65\,\text{km}$).

when driving as compared to using PT (as expected); however, we also find that the difference is marginal, with a median trip distance of 3.84 km for PT, and 3.65 km for driving (Fig. 2.15). While this does not take into account that PT will be slower as well due to stops along the route, this shows us that PT routes are almost as direct as driving.

The trip distances alone do not account for the fact that trips may require traversing multiple routes. Using the public transportation routes included in the GTFS feeds of Metro Manila, we calculated the travel itineraries of trips between all OD pairs. The travel itineraries include the start and end legs of the commute (which are assumed to be via paratransit modes: walking, tricycles, pedicabs, etc.), PUB/PUJ routes taken, and transfers. However, we exclude the start and end legs from the count of number of rides, since we cannot determine which modes are available to commuters with the given data. We find that 73.2% of all trips in the RTRS survey can be achieved in 1-or-less rides, while 96.1% can be done in 2-or-less rides (again, excluding the start and end legs, see Fig. 2.16). A count of zero rides means that the

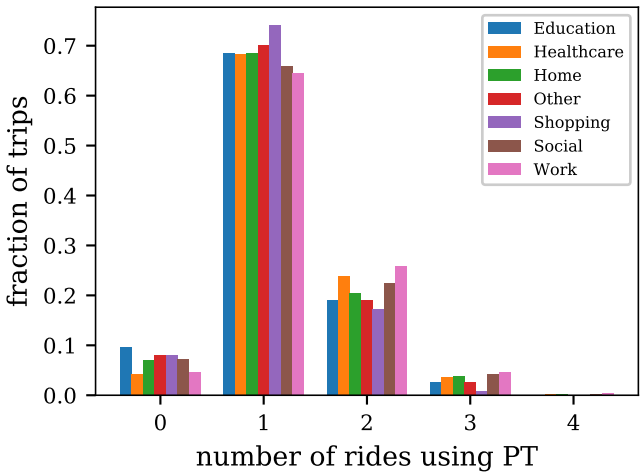

Figure 2.16: Distribution of the number of boardings using public transportation (PT). Distributions show similarities regardless of purpose of travel, with more than 73.2% of all trips achieved in one-or-less ride.

trip could be achieved without using PT through paratransit modes. This observation remains consistent throughout all purposes of travel covered by the RTRS survey. For comparison, a 2004 study on the service area of the Miami-Dade Transit Agency (MDTA), which covers 777 km^2 (Metro Manila: $\sim$ 620 km^2) and has 83 train routes and 81 bus routes, had 14.3% trips achievable in 1-or-less rides, and 55.13% in 2-or-less rides [75]. We also find that more boardings are attributed to longer trips, with trips shorter than 20 km consist 94.7% of all trips with 2-or-less rides (Fig. 2.17).

Despite the abundance of direct routes available to commuters, we know (from anecdotal experience) that public transportation in Metro Manila has a lot of room for improvement. Evaluating the public transport system based on trip travel times (which correlates with distance as well as vehicle speeds) and number of transfers alone (which relates to cost and convenience) overlook the fact that there are other factors that affect the quality of a commute. The sheer challenge of getting rides — evidenced by long lines for the MRT and LRT, and commuters that spill onto the roads while they attempt to

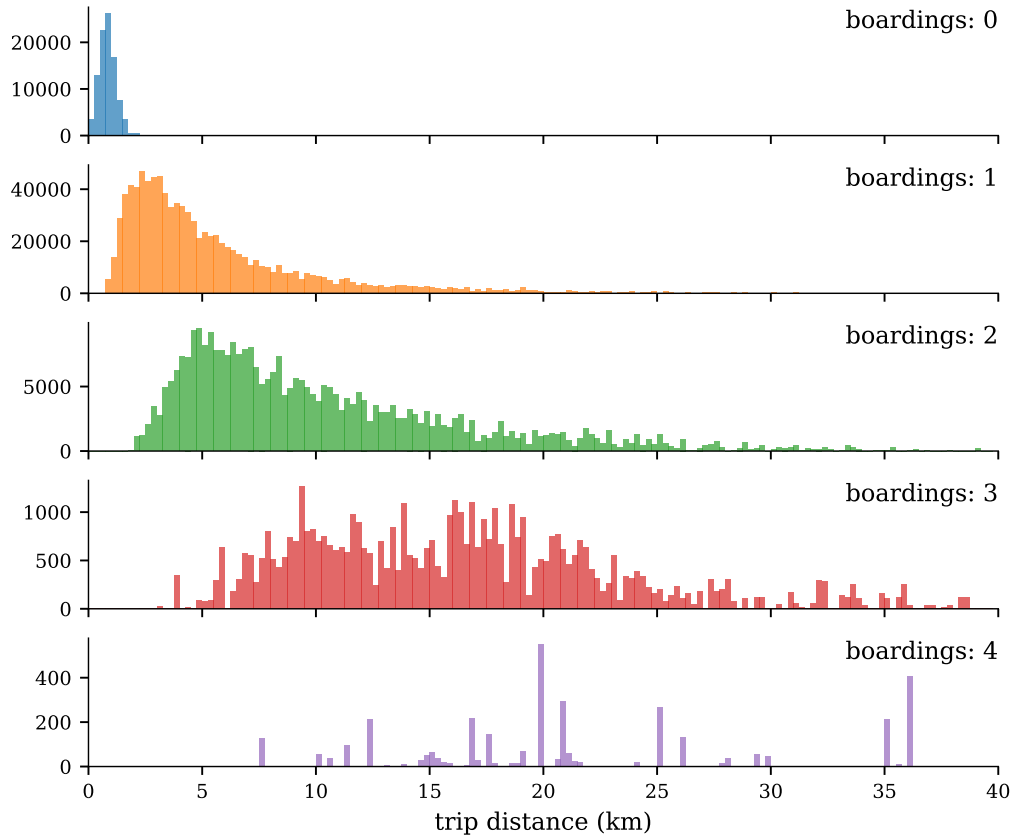

Figure 2.17: Distance distributions, broken down by the number of boardings for the trip. More boardings are more likely to be associated with longer trips.

hail rides — is part of the everyday hell commuters must endure. Other factors such as infrastructure, personal attributes, and comfort are also key considerations in mode choice [76]. Rizada et al. demonstrated that the typical commuter is exposed to high temperatures, uneven walking surfaces, and high humidity [77]. These less than ideal conditions must also be addressed by the government to encourage commuters to shift to public transportation.

Route modernization

Public transportation systems should be dynamic and adapt to the needs of commuters over time. The Land Transport Franchise and Regulatory

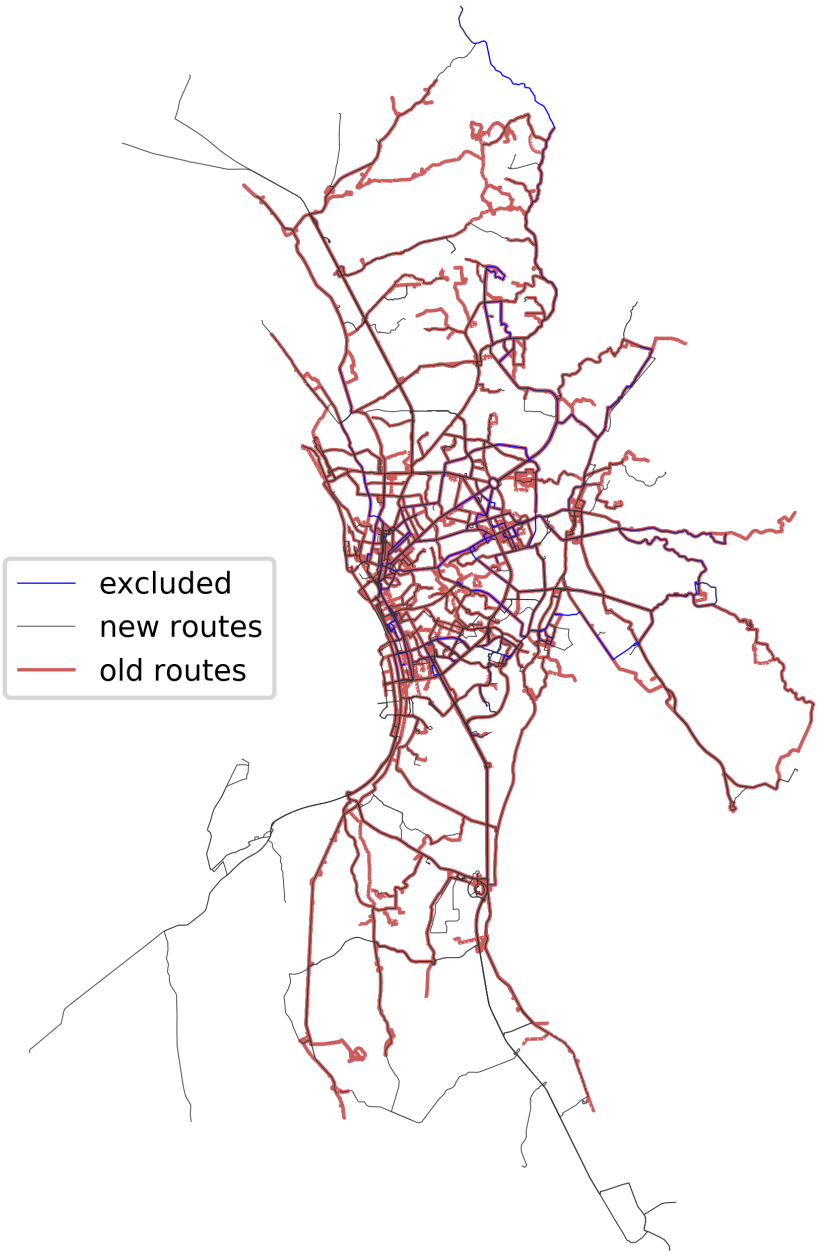

Figure 2.18: Comparison of new and old PT routes. We have excluded some routes in our analysis as these routes were broken up into smaller routes in the source shapefiles.

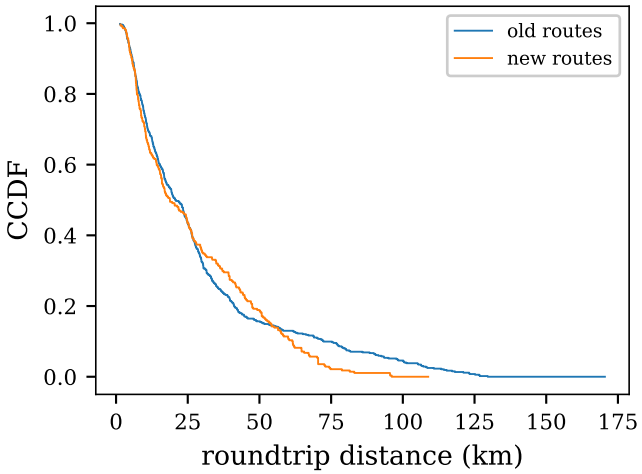

Figure 2.19: Roundtrip length distributions of the old and new public transportation routes. The revised routes reduced are shorter than the old routes.

Board (LTFRB) regulates public transportation by setting fixed endpoints, but routes may generally self-organize over time. Over the years, a combination of regulation and self-organization of resulted in a lot of redundant routes, multiple operators, and an inefficient PT system where drivers of PUVs compete with each other in picking up passengers.

The LTFRB have proposed new routes[3] to overhaul and consolidate routes in the public transportation system of Metro Manila. With the shutdown of public transportation in light of COVID-19, the agency has begun to implement the new bus routes and effectively halt the operations of a lot of the traditional jeepney routes that ply Metro Manila. There are 286 proposed routes (see Tab. 2.20) for PUBs, PUJs, Point-to-Point buses, and UV Express, which consolidates the 523 PUB and PUJ routes that ply Metro Manila just before the COVID-19 pandemic (Fig. 2.18). These new routes are also shorter than the old routes (Fig. 2.19). The proposed overhauled routes cover new areas outside Metro Manila, but also cut service in some

[3]data downloaded as of August 24, 2020. currently available at https://gist.github.com/temetski/856a6d012e5bf0a9b6a1bb2cb17774d8

Table 2.20: List of issuances from the LTFRB authorizing new routes, broken down by mode. Data updated as of August 24, 2020, from https://ltfrb.gov.ph/issuances/memorandum-circulars/ and and https://ltfrb.gov.ph/issuances/board-resolutions/.

Issuance	PUB	P2P	PUVMP	Old PUJ	UVE
BR 2020 No. 124	0	0	1	0	0
MC 2020-019	18	0	0	0	0
MC 2020-023 A	0	0	2	0	0
MC 2020-023 B	0	0	1	0	0
MC 2020-023 C	0	0	8	0	0
MC 2020-023, Annex A	0	0	15	0	0
MC 2020-023, Annex B	0	0	6	0	0
MC 2020-023, Annex C	0	0	8	0	0
MC 2020-025	0	33	0	0	47
MC 2020-026	0	0	0	49	0
MC 2020-029	0	0	0	15	0
MC 2020-029 A	0	0	0	15	0
MC 2020-035	0	0	0	0	4
MC 2020-036	0	0	0	14	0
MC 2020-037	0	0	0	26	0

parts of Metro Manila, Cavite, Laguna, and Bulacan.

The overhauled routes are currently in a shapefile format, and do not have stops defined (as in the GTFS feeds of the current PT system). As a workaround, we generated stops in regularly spaced intervals (250 m) by interpolating the route shapes; we had to exclude 24 routes (see Tab. 2.21) that could not have the stop placements interpolated as the route was not a contiguous shape. We believe it is reasonable to assume regularly spaced stops for a densely populated city.

We compare the new public transportation network against the routes contained in the RTRS as a baseline. We observe changes in trip distances; trip distances increased for 40.9% and decreased for 34.4% of commuters. The remaining 24.7% of commuters had trips that could not be calculated by our shortest path algorithm, which indicates that the overhauled routes

Table 2.21: List of excluded routes by issuance and type.

Issuance	City Bus	PUVMP	Traditional
BR No. 124, Series of 2020	0	1	0
MC 2020-019	13	0	0
MC 2020-023, Annex B	0	2	0
MC 2020-023, Annex C	0	1	0
MC 2020-029	0	0	2
MC 2020-036	0	0	5

had stops that were no longer within $R_{\text{walk}} = 0.5\,\text{km}$ of each other. This does not necessarily mean the trip is impossible with the overhauled routes — increasing R_{walk} to 1 km results in 0.3% cut trips, with 57% of trips shortened and 42% of trips lengthened — this highlights the added inconvenience to nearly a quarter of commuters who now have to walk more than 0.5 km to complete at least one transfer in their itinerary.

The configuration of the overhauled routes increased of the number of boardings required to complete trips, although the trip distance distributions for 4-or-less boardings look similar to the old system (see Fig. 2.17 for the old routes and Fig. 2.22 for the overhauled routes). As a result of the consolidation of redundant routes, 37.9% of commuters have 3-or-more boardings in contrast to 3.87% under the current PT routes. There are commuters that will need to add additional transfers to their commutes, which costs additional fares and wait times.

While we highlight the need to account for the additional burden to commuters, it remains to be seen how the modernization of PUV routes will impact the overall levels of traffic congestion in Metro Manila. It is difficult to directly compare the old and overhauled routes solely based on these figures, as it is possible that the overhaul of routes could also lead to less vehicles and better utilization of seats in PUVs, which could result in faster trips and reduced trips overall.

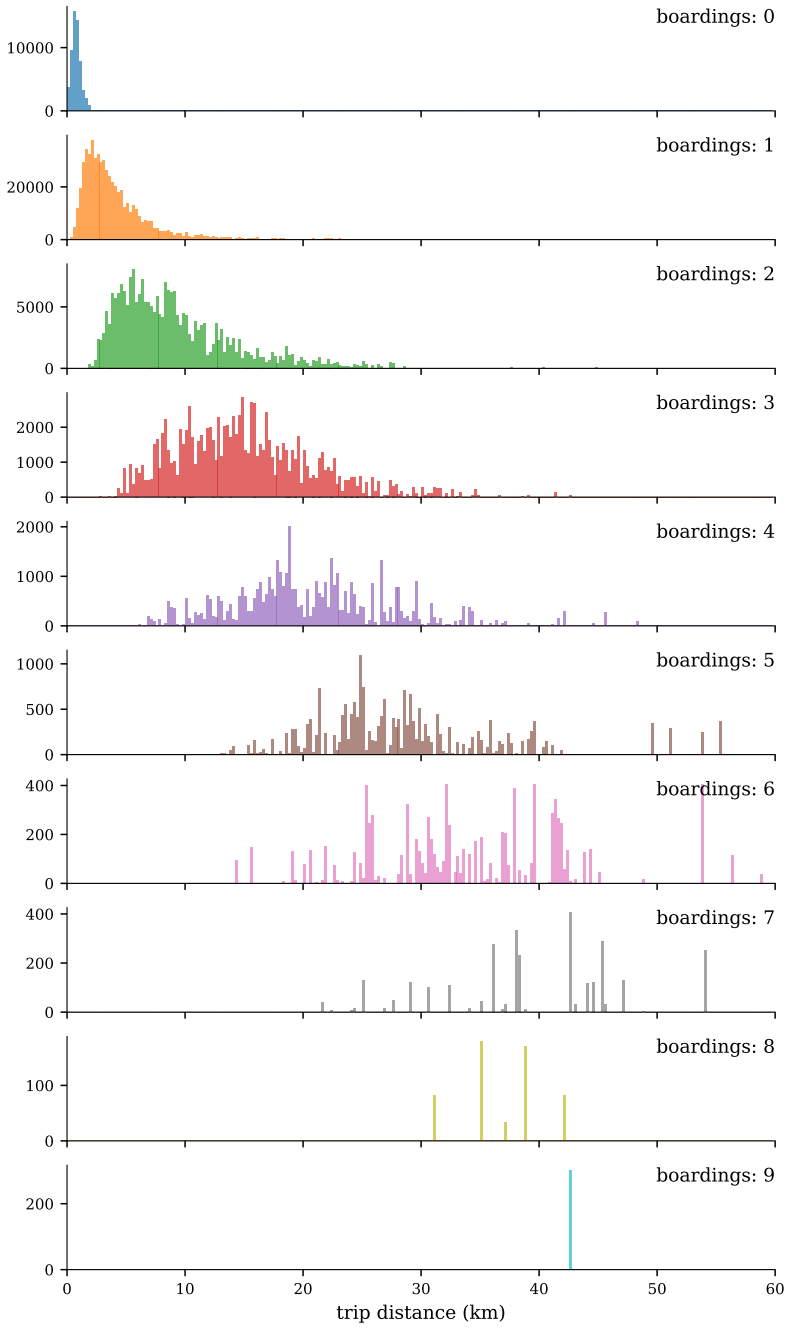

Figure 2.22: Distance distributions, broken down by the number of boardings for the trip. More boardings are more likely to be associated with longer trips.

2.4 Summary

We demonstrated the use of a gravity model to estimate the Origin–Destination trip demands in Metro Manila. This method needs three things to work: aggregated demand data at destinations, granular population residence data (census), and road/route data for calculating travel distances. While this approach fails to to capture the granularity of individual trip demands, it is able to provide a rough estimate of OD-demands aggregated at the traffic analysis zones of Metro Manila. We demonstrate a use-case example using trips tagged with the purpose of `Education`, but this could be extended to other purposes as well.

Building on the promising results of estimating ODs from destination data obtained by OD surveys, we show what the OD demands of the public education system looks like. This approach is free from the biases and constraints of the RTRS survey methodology, and represents a comprehensive picture of what the OD demands would look like for all modes of transport.

It is not our objective to fit the gravity model to the data, but rather demonstrating how the gravity model can capture the general trend of OD demands aggregated at zones. This method of estimating OD demands can be extended to any type of establishment (restaurants, malls, markets, churches) that attracts traffic, provided that the occupancy of such establishments are known, or can be estimated via other methods.

Using Origin-Destination and Boarding-Alighting datasets collected by the Road Transit Rationalization Study, we compared the difference of the zonal aggregates of both datasets. The difference of the datasets highlighted zones that are have insufficient access to public transportation, and zones that are transit hubs. Zones with insufficient access to public transportation coincide with residential areas, which rely on other forms or transport like walking, tricycles, and even private vehicles to begin their commutes. We also identified transit zones where commuters frequently change lines to continue

their travel. These zones are not necessarily associated with the presence of transit infrastructure such as terminals.

Our analysis shows that the differences in the OD and BA datasets highlight Metro Manila's dependence on paratransit modes of transportation that are not captured by the BA surveys. Even with the proposed PUV modernization program that will implement smart cards to that will give updated BA patterns in the city, it is unlikely that this will be able to cover all of the gaps in data collection that will allow us to better analyze traffic patterns in Metro Manila.

Finally, we analyzed the characteristics of Metro Manila as a driving network and public transportation network. We found that trip distances for driving are shorter than using public transportation, but not by a large amount, with most of the difference in trips longer than 5 km. Our public transportation system is efficient where almost all trips can be completed in 2-or-less rides. Public transportation looks good on paper but is unable to discourage private car use; this indicates other factors such as wait times and ride commuting comfort must also be addressed in order to shift commuters to PT, reduce the volume of traffic, and alleviate congestion.

Conclusion

Platooned flow is characterized by bunched slow-moving buses that leave gaps on the road. Having fewer riders or more buses dissolves platoons and results in an equal headway configuration. With multiple lanes, platooned dynamics spill over to the non-bus lane, and persists until the platoon dissolves.

Interactions at stops drive platoon formation. Spacing out stops speeds up traffic by minimizing platooning and queuing-induced congestion. However, spacing them too far would make walking to a midpoint destination impractical. Load-anywhere behavior, while convenient for an individual, comes at the cost of slower traffic flow and longer waiting times. To mitigate the formation of platoons, sticking to a bus schedule helps. Cities may also incentivize staggered work schedules while ensuring adequate bus supply.

The analysis of transitions in our traffic model provide transportation planners with an idea of how to best impact traffic with the least amount of change to the system. The phase transition point tells us to reduce the volume of vehicles, but this approach is impractical to attain given the high traffic volume in cities. The crossover transition involves the interactions between public transportation and commuters, and can occur at densities above the phase transition; this means that it is possible to improve the current state of traffic by managing the supply and demand side of public transportation.

Our work also uses public school enrollment to show deficiencies in the Origin–Destination surveys conducted by the RTRS. The focus of the OD demand survey on commuters that use PUBs and PUJs misses out on com-

muters that do not use public transportation. Crafting policy based on incomplete data can undermine the governments efforts to improve public transportation and alleviate the worsening traffic congestion in Metro Manila. We demonstrated a data driven method of estimating trip demands using a gravity-model, school enrollment, and barangay census data that provides a good estimate of trip demands aggregated at zones, which can be used to augment traditional OD surveys. Our method can be extended to any type of establishment, provided that establishment occupancies are known; this information can be collected and updated yearly through the different annual registration procedures businesses already go through.

Finally, we also analyzed the public transportation system of Metro Manila using a variety of data. We highlighted zones that had deficiencies in access to public transportation, which can play a role for the preference of driving cars for some commuters. The public transportation network of Metro Manila itself is efficient with regards to the directness of routes (short travel distances with the least amount of transfers), although we we excluded the start and end legs of a commute for this analysis. However, our daily experience with PT still shows problems beyond route design. Since our analysis highlights Metro Manila's dependence on paratransit modes, the availability and convenience of these modes also play an important role in the perception of PT. And for as long as riding public transportation involves long lines and uncomfortable commutes, driving and personal modes of transport will continue to attract commuters and contribute to congestion. While the LTFRB has proposed new set of routes that overhaul the PT routes in our current system, we found that this results in longer trips, additional rides, and increased transfer distances for commuters. The LTFRB must carefully weigh the additional inconvenience to commuters against improvements brought about by the overhaul of routes if the end goal is to convince more commuters to use public transportation.

Recommendation

In the course of this dissertation, we encountered several challenges in processing geospatial data. While working on census data, we encountered map alignment issues with the 2002 PSA barangay shapefiles. An updated 2015 census was available in tabular form but lacked the shapefiles necessary in our analysis. While the PSA is responsible for census data, a separate agency, NAMRIA creates map data used as the basis for barangay shapefiles. Geospatial datasets for the Philippines must have better integration with the base shapefiles provided by NAMRIA to avoid alignment and matching issues.

Part of this work was supposed to evaluate the public transportation system using traffic simulations at the scale of Metro Manila. We abandoned this approach upon realizing that unifying the route geometries with the road geometries was a non-trivial task. Our observations showed that route data stored in two separate shapefiles were incompatible — even if these supposedly referred to the same routes. Inspection of the raw data suggested that generating these shapefiles involved manually selecting points on a map to *draw* the shape of the route. Shapefiles maintained by different agencies share visual similarities but contain different and incompatible data layers. Because different agencies maintain their own siloed shapefiles with no way of (painlessly) integrating, this hinders large-scale research that involves processing and analyzing such datasets.

To avoid these issues, we recommend that agencies and data providers determine and agree on a base dataset that serves as a common reference

point. In the context of geospatial data such as roads and administrative boundaries, we propose to use OpenStreetMaps (OSM) as a reference. Route and road segments should explicitly reference OSM objects, which allows for cross-referencing across datasets. Using a base dataset will make it easier to create, distribute, and analyze geospatial data and reduce time spent cleaning and pre-processing data. Future studies, researchers, and even concerned agencies would greatly benefit as research can focus on complex analyses of transportation problems in Metro Manila.

Appendix A

Theoretical ρ_{crit}

In Section A.1, we show how to compute for the theoretical ρ_{crit} on a road with maximum speed $v_{\max}$ and vehicle lengths l.

A.1 Dependence of ρ_{crit} on $v_{\max}$ and l

One primary difference between a road with all cars ($\kappa = 1$) and all motorcycles ($\kappa = 0$) is that the critical densities ρ_{crit} for the phase transition from a free flow state to an interacting state lies at different density values. We can immediately see from **??** that the median velocity of motorcycles start to decrease at $\rho_{\text{crit}} \approx 0.1$, while for cars, it occurs at $\rho_{\text{crit}} \approx 0.16$. The difference between cars and motorcycles in our model is in the length of the two vehicle types. Because of this, there will be less cars than motorcycles, even if the overall density is the same.

To find theoretical ρ_{crit}, we want to find the range of values of ρ wherein vehicles have minimal interactions. At these low densities, the primary means of interaction is the **Deceleration Rule**. Let us first consider a road with a single lane $W = 1$, with length $L = 100$, maximum velocity $v_{\max} = 5$, with no lanechanging $p_\lambda = 0$ and no random slowdown $p_{\text{slow}} = 0$. The road is populated with vehicles of length l. In the free flow state, we expect vehicles to move at $v = v_{\max}$. Since the only way vehicles can decelerate if through

the **Deceleration Rule**, then it must follow that all vehicles must have headways $d_i >= v_{\max}$. It follows that the maximum number of vehicles N that can fit into the road satisfying this condition is

$$N_{\text{crit}} = \frac{L}{l + v_{\max}}. \tag{A.1}$$

The road density that corresponds to the number of vehicles N is given by

$$\begin{aligned} \rho_{\text{crit}} &= \frac{N_{\text{crit}} l}{L} \\ \rho_{\text{crit}} &= \frac{l}{l + v_{\max}}. \end{aligned} \tag{A.2}$$

The above equation tells us that ρ_{crit} depends on the length of the vehicle l and the maximum speed attainable $v_{\max}$. For $l = 1, v_{\max} = 1$, $\rho_{\text{crit}} = 0.5$, which is the result obtained by [78]. For the values in our model, we have $v_{\max} = 5$ and $l = 1$ for motorcycles, $l = 2$ for cars. Those values correspond to $\rho_{\text{crit}} = \frac{1}{6} \approx 0.17$ for motorcycles and $\rho_{\text{crit}} = \frac{2}{7} \approx 0.29$.

When we factor in p_λ, vehicles can use lane changing to avoid interactions, effectively raising ρ_{crit}. The random slowdown probability in turn introduces a probabilistic chance of forcing an interaction via the **Deceleration Rule**, effectively lowering ρ_{crit}. This lowering of ρ_{crit} due to non-zero values of p_{slow} is demonstrated in **??**.

Bibliography

[1] O. J. O'Loan, M. R. Evans, and M. E. Cates, "Jamming transition in a homogeneous one-dimensional system: The bus route model," Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top. **58**, 1404–1418 (1998).

[2] Y. J. Luo, B. Jia, X. G. Li, C. Wang, and Z. Y. Gao, "A realistic cellular automata model of bus route system based on open boundary," Transp. Res. Part C Emerg. Technol. **25**, 202–213 (2012).

[3] L.-M. Kieu, D. Ngoduy, N. Malleson, and E. Chung, "A stochastic schedule-following simulation model of bus routes," Transp. B Transp. Dyn. **7**, 1588–1610 (2019).

[4] D. Chowdhury and R. C. Desai, "Steady-states and kinetics of ordering in bus-route models: Connection with the Nagel-Schreckenberg model," Eur. Phys. J. B **15**, 375–384 (2000).

[5] T. Nagatani, "Bunching transition in a time-headway model of a bus route," Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top. **63**, 1–7 (2001).

[6] D. Helbing, "Empirical traffic data and their implications for traffic modeling," Phys. Rev. E **55**, 25–28 (1997).

[7] C. Gershenson and L. A. Pineda, "Why Does Public Transport Not

Arrive on Time? The Pervasiveness of Equal Headway Instability," PLoS One **4**, e7292 (2009).

[8] D. Q. Nguyen-Phuoc, G. Currie, C. De Gruyter, I. Kim, and W. Young, "Modelling the net traffic congestion impact of bus operations in Melbourne," Transp. Res. Part A Policy Pract. **117**, 1–12 (2018).

[9] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation," Phys. Rev. E **51**, 1035–1042 (1995).

[10] Y. Sugiyama, M. Fukui, M. Kikuchi, K. Hasebe, A. Nakayama, K. Nishinari, S. I. Tadaki, and S. Yukawa, "Traffic jams without bottlenecks-experimental evidence for the physical mechanism of the formation of a jam," New J. Phys. **10** (2008).

[11] R. Nagai, T. Nagatani, and N. Taniguchi, "Traffic states and jamming transitions induced by a bus in two-lane traffic flow," Physica A **350**, 548–562 (2005).

[12] R. Z. Koshy and V. T. Arasan, "Influence of Bus Stops on Flow Characteristics of Mixed Traffic," J. Transp. Eng. **131**, 640–643 (2005).

[13] Y.-M. Yuan, R. Jiang, Q.-S. Wu, and R. Wang, "Traffic Behavior in a Two-Lane System Consisting of a Mixture of Buses and Cars," Int. J. Mod. Phys. C **18**, 1925–1938 (2008).

[14] D. N. Dailisan and M. T. Lim, "Agent-based modeling of lane discipline in heterogeneous traffic," Physica A **457**, 138–147 (2016).

[15] D. N. Dailisan and M. T. Lim, "Vehicular traffic modeling with greedy lane-changing and inordinate waiting," Physica A **521**, 715–723 (2019).

[16] Metropolitan Manila Development Agency, "Metropolitan Manila Annual Average Daily Traffic," (2018).

[17] Y. Boquet, "Bus Transportation in the Philippines," 18th Hong Kong Soc. Transp. Stud. Conf. pp. 511–518 (2013).

[18] A. Rey, "EXPLAINER: What is the provincial bus ban?" Rappler (2019).

[19] JICA, "Roadmap for Transport Infrastructure Development for Metro Manila and Its Surrounding Areas (Region III and Region IV-A)," Tech. rep., Japan International Cooperation Agency (2014).

[20] B. D. Greenshields, J. T. Thompson, H. C. Dickinson, and R. S. Swinton, "The photographic method of studying traffic behavior," Highw. Res. Board Proc. **13** (1934).

[21] B. Greenshields, J. Bibbins, W. Channing, and H. Miller, "A study of traffic capacity," Highw. Res. Board Proc. **1935** (1935).

[22] K. Nagel and M. Schreckenberg, "A cellular automaton model for freeway traffic," J. Phys. I **2**, 2221–2229 (1992).

[23] M. J. Lighthill and G. B. Whitham, "On Kinematic Waves . II . A Theory of Traffic Flow on Long Crowded Roads Article," Proc. R. Soc. London A **229** (1955).

[24] P. I. Richards, "Shock Waves on the Highway," Oper. Res. **4**, 42–51 (1956).

[25] D. Helbing, *Verkehrsdynamik: neue physikalische Modellierungskonzepte* (Springer-Verlag, 2013).

[26] C. F. Daganzo, M. J. Cassidy, and R. L. Bertini, "Possible explanations of phase transitions in highway traffic," Transp. Res. A **33**, 365–379 (1999).

[27] M. Treiber and D. Helbing, "Macroscopic Simulation of Widely Scattered Synchronized Traffic States," J. Phys. A. Math. Gen. **32** (1999).

[28] D. Helbing, "Traffic and related self-driven many-particle systems," Rev. Mod. Phys. **73**, 1067–1141 (2001).

[29] D. Helbing, "Derivation and Empirical Validation of a Refined Traffic Flow Model," Science (80-.). **233**, 253–282 (1998).

[30] A. Reuschel, "Fahrzeugbewegungen in der kolonne," Osterreichisches Ingenieur Archiv **4**, 193–215 (1950).

[31] D. C. Gazis, R. Herman, and R. B. Potts, "Car-Following Theory of Steady-State Traffic Flow," Oper. Res. **7**, 499–505 (1959).

[32] M. Treiber and A. Kesting, "Traffic flow dynamics," Traffic Flow Dyn. Data, Model. Simulation, Springer-Verlag Berlin Heidelb. (2013).

[33] B. S. Kerner, S. L. Klenov, and M. Schreckenberg, "Probabilistic physical characteristics of phase transitions at highway bottlenecks: Incommensurability of three-phase and two-phase traffic-flow theories," Phys. Rev. E **89**, 1–13 (2014).

[34] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic Traffic Simulation using SUMO," in "IEEE Intell. Transp. Syst. Conf.", (IEEE, 2018).

[35] J. Casas, J. L. Ferrer, D. Garcia, J. Perarnau, and A. Torday, *Traffic Simulation with Aimsun* (Springer New York, New York, NY, 2010), pp. 173–232.

[36] M. Fellendorf, "VISSIM: A microscopic Simulation Tool to Evaluate Actuated Signal Control including Bus Priority," 64th Inst. Transp. Eng. Annu. Meet. pp. 1–9 (1994).

[37] K. Ramachandra Rao and C. Mallikarjuna, "Cellular Automata Model for Heterogeneous Traffic," J. Adv. Transp. **43**, 321–345 (2008).

[38] K. Nagel and M. Paczuski, "Emergent traffic jams," Phys. Rev. E **51**, 2909–2918 (1995).

[39] M. Schreckenberg, A. Schadschneider, K. Nagel, and N. Ito, "Discrete stochastic models for traffic flow," Phys. Rev. E **51**, 2939–2949 (1995).

[40] J. S. L. Combinido and M. T. Lim, "Crowding effects in vehicular traffic." PLoS One **7**, e48151 (2012).

[41] J. S. L. Combinido and M. T. Lim, "Modeling U-turn traffic flow," Physica A **389**, 3640–3647 (2010).

[42] D. Chowdhury, D. E. Wolf, and M. Schreckenberg, "Particle hopping models for two-lane traffic with two kinds of vehicles: Effects of lane-changing rules," Physica A **235**, 417–439 (1997).

[43] D. Chowdhury, L. Santen, and A. Schadschneider, "Statistical physics of vehicular traffic and some related systems," Phys. Reports-Review Sect. Phys. Lett. **329**, 199–329 (2000).

[44] G. Hou and S. Chen, "An improved cellular automaton model for work zone traffic simulation considering realistic driving behavior," J. Phys. Soc. Japan **88**, 1–10 (2019).

[45] X. Li, Y. Xiao, B. Jia, and X. Li, "Modeling mechanical restriction differences between car and heavy truck in two-lane cellular automata traffic flow model," Physica A **451**, 49–62 (2016).

[46] S. Lübeck, M. Schreckenberg, and K. D. Usadel, "Density fluctuations and phase transition in the Nagel-Schreckenberg traffic flow model S." Phys. Rev. E **57**, 1171–1174 (1998).

[47] K. Nagel and H. J. Herrmann, "Deterministic models for traffic jams," Physica A **199**, 254–269 (1993).

[48] G. Hou, S. Chen, and Y. Han, "Traffic Performance Assessment Methodology of Degraded Roadway Links Following Hazards," J. Aerosp. Eng. **32**, 1–12 (2019).

[49] A. T. Murray and X. Wu, "Accessibility tradeoffs in public transit planning," J. Geogr. Syst. **5**, 93–107 (2003).

[50] L. E. Olmos, S. Çolak, S. Shafiei, M. Saberi, and M. C. González, "Macroscopic dynamics and the collapse of urban traffic," Proc. Natl. Acad. Sci. **115**, 12654–12661 (2018).

[51] W. Jang, "Travel Time and Transfer Analysis Using Transit Smart Card Data," Transp. Res. Rec. J. Transp. Res. Board **2144**, 142–149 (2010).

[52] X. Ma, Y. J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders' travel patterns," Transp. Res. Part C Emerg. Technol. **36**, 1–12 (2013).

[53] E. F. Legara, C. Monterola, K. K. Lee, and G. G. Hung, "Critical capacity, travel time delays and travel time distribution of rapid mass transit systems," Physica A **406**, 100–106 (2014).

[54] A. Alsger, B. Assemi, M. Mesbah, and L. Ferreira, "Validating and improving public transport origin-destination estimation algorithm using smart card fare data," Transp. Res. Part C Emerg. Technol. **68**, 490–506 (2016).

[55] M. Munizaga, F. Devillaine, C. Navarrete, and D. Silva, "Validating travel behavior estimated from smartcard data," Transp. Res. Part C Emerg. Technol. **44**, 70–79 (2014).

[56] J. L. Toole, S. Colak, B. Sturt, L. P. Alexander, A. Evsukoff, and M. C. González, "The path most traveled: Travel demand estimation using big data resources," Transp. Res. Part C Emerg. Technol. **58**, 162–177 (2015).

[57] M. S. Iqbal, C. F. Choudhury, P. Wang, and M. C. González, "Development of origin-destination matrices using mobile phone call data," Transp. Res. Part C Emerg. Technol. **40**, 63–74 (2014).

[58] C. Wu, J. Thai, S. Yadlowsky, A. Pozdnoukhov, and A. Bayen, "Cellpath: Fusion of Cellular and Traffic Sensor Data for Route Flow Estimation via Convex Optimization," Transp. Res. Procedia **7**, 212–232 (2015).

[59] R. Jurdak, K. Zhao, J. Liu, M. AbouJaoude, M. Cameron, and D. Newth, "Understanding human mobility from Twitter," PLoS One **10**, 1–16 (2015).

[60] The World Bank, "World Development Indicators," accessed 2018-04-01 from https://datacatalog.worldbank.org/dataset/world-development-indicators.

[61] Department of Transportation and Communications, "Metro Manila Road Transit Rationalisation Study," Tech. rep., Department of Transportation and Communication (2014).

[62] Department of Transportation and Communications, "Metro Manila Road Transit Rationalisation Study: Development of Mass Transit Corridors," Tech. Rep. January, Department of Transportation and Communication (2016).

[63] D. Dailisan, *Modeling lane changing and random slowdown in vehicular traffic using a cellular automata model (MSc thesis)* (University of the Philippines, 2017).

[64] L. J. M. Rubio, D. N. Dailisan, M. J. P. Osorio, C. C. David, and M. T. Lim, "Modeling the residential distribution of enrolled students to assess boundary-induced disparities in public school access," PLoS One **14**, 1–15 (2019).

[65] Philippine Statistics Authority, "Data Kit of Official Philippine Statistics (DATOS)," Philippine Statistics Authority (2015).

[66] OpenStreetMap contributors, "Planet dump retrieved from https://planet.osm.org," https://www.openstreetmap.org (2019).

[67] G. K. Zipf, "The P_1P_2/D hypothesis: on the intercity movement of persons," American sociological review **11**, 677–686 (1946).

[68] A. M. Voorhees, "A general theory of traffic movement," Proceedings, Institute of Traffic Engineers (1955). Reprinted in Transportation **40** 1105–1116 (2013).

[69] L. Sjaastad, "The relationship between migration and income in the united states," Papers in Regional Science **6**, 37–64 (1960).

[70] V. Palchykov, M. Mitrovic, H. H. Jo, J. Saramäki, and R. K. Pan, "Inferring human mobility using communication patterns," Sci. Rep. **4**, 1–6 (2014).

[71] K. Bhattacharya, G. Mukherjee, J. Saramäki, K. Kaski, and S. S. Manna, "The international trade network: weighted network analysis and modelling," Journal of Statistical Mechanics: Theory and Experiment **2008**, P02002 (2008).

[72] D. B. West *et al.*, *Introduction to graph theory*, vol. 2 (Prentice hall Upper Saddle River, NJ, 1996).

[73] E. W. Dijkstra *et al.*, "A note on two problems in connexion with graphs," Numerische mathematik **1**, 269–271 (1959).

[74] F. Simini, M. C. González, A. Maritan, and A. L. Barabási, "A universal model for mobility and migration patterns," Nature **484**, 96–100 (2012).

[75] F. Zhao and I. Ubaka, "Transit Network Optimization - Minimizing Transfers and Optimizing Route Directness," J. Public Transp. **7**, 63–82 (2004).

[76] J. Zhou, "Sustainable commute in a car-dominant city: Factors affecting alternative mode choices among university students," Transp. Res. Part A Policy Pract. **46**, 1013–1029 (2012).

[77] L. G. Rizada, A. C. Balingit, and M. Lim, "Feasibility of a time-resolved index of commuter comfort," Proceedings of the Samahang Pisika ng Pilipinas (2018).

[78] D. A. Rosenblueth and C. Gershenson, "A Model of City Traffic Based on Elementary Cellular Automata," Complex Syst. pp. 305–322 (2011).