Downloaded from orbit.dtu.dk on: May 11, 2024

DTU Library

Mesoscopic Simulation of Multi-Modal Urban Traffic

Paulsen, Mads

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Paulsen, M. (2020). *Mesoscopic Simulation of Multi-Modal Urban Traffic*.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

DTU Management
Department of Management, Technology and Economics

Mesoscopic Simulation of Multi-Modal Urban Traffic

PhD Thesis

Mads Paulsen

June 2020

DTU Management
Department of Management, Technology and Economics

Mesoscopic Simulation of Multi-Modal Urban Traffic

PhD Thesis
June 2020

By
Mads Paulsen

Main supervisor:	Otto Anker Nielsen, Professor at Department of Management, Technology and Economics, Technical University of Denmark.
Co-supervisor:	Thomas Kjær Rasmussen, Associate Professor at Department of Management, Technology and Economics, Technical University of Denmark.
Copyright:	Reproduction of this publication in whole or in part must include the customary bibliographic citation, including author attribution, report title, etc.
Cover photo:	News Øresund – Johan Wessman, 2013 under CC BY 2.0 license. Visible firm name erased from photo.
Published by:	DTU, Department of Management, Technology and Economics, Akademivej, Building 358, DK-2800 Kongens Lyngby, Denmark. www.man.dtu.dk

Preface

This PhD thesis entitled *Mesoscopic Simulation of Multi-Modal Urban Traffic* is submitted to meet the requirements for obtaining a PhD degree at the Department of Management, Technology and Economics, DTU Management, Technical University of Denmark. The PhD project was supervised by Professor Otto Anker Nielsen and co-supervised by Associate Professor Thomas Kjær Rasmussen, both from DTU Management. The thesis is paper-based and consists of the chapters listed in the tables of content, including separate chapters for each of the following papers:

Paper 1: M. Paulsen, T. K. Rasmussen, and O. A. Nielsen (2018). "Output variability caused by random seeds in a multi-agent transport simulation model". In: *Procedia Computer Science* 130, pp. 850–857. DOI: 10.1016/j.procs.2018.04.078. URL: https://linkinghub.elsevier.com/retrieve/pii/S187705091830440X.

Paper 2: M. Paulsen, T. K. Rasmussen, and O. A. Nielsen (2020a). "Impacts of real-time information levels in public transport: A large-scale case study using an adaptive passenger path choice model". Under review at Transportation Research Part A: Policy and Practice.

Paper 3: M. Paulsen, T. K. Rasmussen, and O. A. Nielsen (2019). "Fast or forced to follow: A speed heterogeneous approach to congested multi-lane bicycle traffic simulation". In: *Transportation Research Part B: Methodological* 127, pp. 72–98. DOI: 10.1016/j.trb.2019.07.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0191261518310336.

Paper 4: M. Paulsen and K. Nagel (2019). "Large-Scale Assignment of Congested Bicycle Traffic Using Speed Heterogeneous Agents". In: *Procedia Computer Science* 151, pp. 820–825. DOI: 10.1016/j.procs.2019.04.112. URL: https://linkinghub.elsevier.com/retrieve/pii/S1877050919305769.

Paper 5: M. Paulsen, T. K. Rasmussen, and O. A. Nielsen (2020b). "Including right-of-way in a joint large-scale agent-based dynamic traffic assignment model for car and bicycle traffic". Manuscript submitted to Networks and Spatial Economics.

Acknowledgements

Although I am the sole author of this PhD thesis, the thesis and the included papers would not have been possible without the support of several persons and organisations.

Firstly, I would like to acknowledge my main supervisor Professor Otto Anker Nielsen and my co-supervisor Associate Professor Thomas Kjær Rasmussen. The fortnightly PhD meetings have helped me to stay on the right track and to guide my research in a favourable direction. Furthermore, the feedback on article drafts has been a great help. A special thanks to Thomas for being available between meetings for discussing ad-hoc issues that arose along the way.

In this regard, I will also like to thank my colleagues at DTU, especially my fellow PhD students in our division. Both for being helpful when reaching into your respective areas of expertise, but also for always providing a joyful atmosphere in the building.

I spent 4.5 months visiting Professor Kai Nagel at Technische Universität Berlin from October 2018 to March 2019. His assistance with implementing my methodology for bicycle simulation in MATSim has been crucial for my research. The evenings with double wall projection (occasionally with both feet on the table) were extremely beneficial, and have given me invaluable knowledge on MATSim. On top of this, I am also highly grateful for his hospitality during my stay. Likewise, I would also like to thank the remainder of the staff at VSP.

A special thanks is deserved for Rasmus Albrink for introducing me to and collected the video tracking data used in Paper 3. This data has severely increased the quality and validity of my research.

Penultimately, I acknowledge the financial support for my PhD provided by the IPTOP (Integrated Public Transport Optimisation and Planning) Project granted by Innovation Fund Denmark (grant 4109-00005B).

I am very grateful for my friends who have taken the initiative to meet when I was not able to. Special thanks to the board of Grundtvigselever for arranging the biannual alumni reunion weekends at Grundtvigs Højskole where I truly disconnected from my PhD.

Most importantly, however, I am especially appreciative of my family for bearing with me through the entirety of my PhD despite my endless mood swings. Although I owe Inga so many hours of walking and belly rubs, she has never guilted me – not even when forcing her to move to Berlin. What a natural born survivor she is ♡. Finally, I owe Louise so many hours of mental presence and household duties, and – especially during the COVID-19 crisis – just occasionally not being home.

Mads Paulsen, June 2020.

Summary

Urbanisation is one of the most dominant global megatrends. As many metropolitan transport networks are already under pressure, the prospect of having to fit even more people into such often car dominated urban systems is disturbing. Not only does it result in countless hours wasted in traffic, but also increasing local pollution and green house gas emissions.

One way to relieve motorised traffic in urban areas is to provide appropriate alternatives to car traffic. Public transport systems where passengers can reach their destinations reliably and dedicated bicycle infrastructure allowing people to travel smoothly by bicycle can for many be seen as sustainable and effective alternatives to car traffic in urban areas. However, even in cities with a high share of cyclists, the models for evaluating future projects and initiatives concerning bicycle traffic are of much lower quality than their counterparts for car traffic, for instance regarding their ability to estimate travel times. As such projects are regarded as socioeconomic investments with the return being received in travel time savings, there is a risk that the lack of appropriate models is setting back bicycle-friendly development.

In order to be able to evaluate transport policies with a higher level of detail, in the past decade transport models have shifted towards agent-based models. Agent-based models follow every individual of the population for the entire day, and makes it possible to evaluate very detailed effects that macroscopic models cannot. Despite agent-based models being capable of simulating very large areas in great detail, the aforementioned capability gap between models for car traffic and other modes of transport persists, at least when disregarding dedicated microscopic models limited to small geographical areas.

This PhD thesis aims at reducing this gap by modelling detailed individual behaviour in large-scale agent-based transport simulation models. The trade-off between detail level and computational performance is a vital focus point, as the developed models are intended for application on a metropolitan scale. Even while doing so, the thesis contributes diversely to the literature through five papers on agent-based modelling divided into three parts: i) Output variability in agent-based simulation of transport systems (Paper 1), ii) Agent-based passenger delay modelling with real-time information (Paper 2), and iii) Agent-based simulation of bicycle traffic and interaction with cars (Papers 3-5).

Part I deals with one of the downsides of most agent-based simulation models – the strong dependence on pseudo-random numbers. Whereas most aggregated models can get away with stating choices by their choice distribution, the individual constituents of agent-based models require explicit choices to be made, either by resolving to completely deterministic, rule-based setups or by using pseudo-random numbers to draw such choices stochastically. However, as sequences of pseudo-random numbers are entirely determined

by their random seed, changing the seed will also change modelled choices, eventually alter the final model output. As such, the output of stochastic agent-based models can be considered as samples from – extremely complex – random distributions.

The first paper of the thesis investigates the variability of such outputs by conducting an experiment in MATSim with a large-scale scenario of Santiago de Chile using 100 different runs – each using a different random seed. The corresponding link flows are analysed, and the between-seed variation is generally found to be small with tolerable coefficients of variation. However, for some links the relative error is occasionally serious when compared to other random seeds, indicating that the results based on a single seed may be deceiving. Finally, it is found that the variation between the last and penultimate iteration is almost entirely dominated by the variation across the different seeds, suggesting the need for using multiple random seeds when analysis results.

Part II of the thesis is concerned with agent-based passenger delay modelling under consideration of real-time information. Public transport systems are often uncertain, and automated vehicle location (AVL) data of their vehicles can be used to measure vehicle punctuality. Although being extremely relevant, evaluating how such delays influence passenger travel times is rarely done, presumably because doing so is inherently tricky as passengers may adapt to changes along the way, especially when real-time information is present.

The objective of *the second paper* of the thesis is to formulate a simulation model that based on recorded vehicle delays from AVL data can determine the corresponding passenger delays at a large scale under different levels of real-time information prevalence. The model lets its passengers search for new alternatives every two and a half minutes while they travel through the system based on the at any time accessible real-time information. The model is applied to the public transport system of Metropolitan Copenhagen and 812,359 daily trips are modelled for 65 days where real-life AVL data from the rail and bus network was collected. The study shows how a particularly irregular railway line causes many passengers to pursue alternative routes, and that information of better alternatives often occur at stations with many high-classed options. In line with existing literature it is found that despite some trips arriving earlier than expected, on average the passenger delays are far larger than the vehicle delays that cause them. The study further discovers that passenger delays can be reduced considerably by using real-time information obtained at the beginning of the trip and even further when acquiring information en-route.

Part III consists of three papers that successively develop a novel, tailor-made methodology for simulating bicycle traffic and interaction with car traffic at a metropolitan scale. The first paper of the part – *the third paper* of the thesis – develops a computationally fast methodology for realistically simulating congestion in separated bicycle traffic. As the heterogeneity of cyclists is profound, taking overtaking into account is essential for

being able to simulate bicycle traffic appropriately. The research uses video tracking data of low intensity bicycle traffic to investigate how desired speeds and preferred headway distances vary across cyclists, and estimates appropriate probability distributions for these cyclist characteristics used as input for the developed model. The model is based on simple assumptions, and allows overtaking by having cyclists explicitly choosing lanes when entering a link. A bottleneck network is used for testing the methodology under a wide spectrum of traffic intensities. As designed for, the model is shown to be more likely to pose excess travel time on cyclists with high desired speeds, whereas high traffic intensities are required to consistently delay cyclists with preferences for slower speeds. Still, the derived fundamental diagrams of the small-scale experiments are validated with observed video tracking data of 3,763 cyclists from a morning peak hour at Queen Louise's Bridge in Central Copenhagen. The research carried out in the paper exposes that recognising cyclist heterogeneity is essential for realistic simulation of bicycle traffic.

The fourth paper is a natural extension of the previous paper. Before this paper, no dedicated bicycle traffic assignment models with feedback between route choice and travel times were present in the literature. The paper alters this by using the methodology from the third paper to simulate bicycle traffic and integrates it in a traffic assignment model with meaningful feedback between supply and demand. The method is implemented in MATSim and applied to a large-scale case study of Metropolitan Copenhagen with 1,082,958 bicycle trips. Although the excess travel time of cyclists is low compared to other modes, through scenarios using better and worse bicycle infrastructure the study shows that this can to a large degree be contributed to the high level of bicycle infrastructure in Copenhagen. Still, it is shown that flows differ significantly between the initial and final iterations, demonstrating that feedback from the network is needed when modelling bicycle traffic in cities with a high share of cyclists – even more so if the infrastructure is insufficient.

Whereas the two preceding papers focus on simulated bicycle traffic on links, *the fifth and final paper* addresses how to model the intersections of the network. In cities with a high level of separated bicycle traffic, intersections are of particular interest as practically all interactions between bicycles and motorised traffic occur here. The paper formulates a joint car and bicycle traffic assignment model capable of modelling right-of-way at every network node, ensuring that conflicting moves do not take place simultaneously. The agent-based model is implemented in MATSim and applied to the same large-scale case study of Metropolitan Copenhagen as the fourth paper, but with the addition of 3,210,685 car trips, 299,416 truck trips, and a car network increasing the network size to 572,935 links and 144,060 nodes. Simulations are run with the right-of-way node model and the original node model of MATSim for setups with car/truck, bicycle, and all three modes, respectively. Without increasing computation times notably, the study shows that omitting to include yielding due to right-of-way at intersections underestimates travel

times and causes too much traffic to be led through the city centre. It is furthermore shown that inter-modal conflicts between bicycle traffic and motorised traffic at intersections delay traffic more than the intra-modal conflicts, emphasising the need for joint modelling of multi-modality.

In summary, this PhD thesis has contributed to the literature on large-scale agent-based simulation of urban transport systems, spanning across output variability, public transport passenger delay modelling, and dynamic traffic assignment of bicycle traffic. This includes recommendations for future practices when dealing with agent-based modelling, but also in terms of developing methodologies for new types of models for detailed large-scale modelling of passenger delays and congested bicycle traffic, allowing estimating effects that were previously ignored when planning and evaluating metropolitan transport systems. Although there are still additional steps to be taken within each of the topics covered by the five papers of the PhD thesis, their contributions constitute considerable improvements to, not only the understanding, but also the capabilities of agent-based transport simulation models for multi-modal urban traffic.

Resumé (Danish sumary)

Urbanisering er en af de mest dominerende globale megatrends. Transportnetværk er allerede under pres i mange storbyer, hvorfor udsigten til at skulle presse endnu flere mennesker ind i sådanne ofte bildominerede urbane systemer er foruroligende. Det fører ikke alene til talrige timer spildt i trafikken, men også til øget lokalforurening og udledning af drivhusgasser.

En måde at aflaste trængslen i byområder er ved at tilbyde hensigtsmæssige alternativer til biler. Kollektive transportsystemer, hvor passagerer kan nå deres destinationer pålideligt, og cykelinfrastruktur, der tillader folk at cykle gnidningsfrit, kan for mange ses som bæredygtige og effektfulde alternativer til biltrafikken i byområder. Imidlertid, selv for byer med en høj andel af cyklister er modellerne til at evaluere fremtidige projekter og initiativer målrettet cykeltrafik af langt lavere kvalitet end deres pendanter for biltrafik, fx angående deres evne til at estimere rejsetider. Da sådanne projekter betragtes som samfundsøkonomiske investeringer, hvor afkastet tilbagebetales i rejsetidsbesparelser, er der en risiko for, at manglen på egnede modeller bremser cykelvenlig udvikling.

For at være i stand til at evaluere transporttiltag med en højere detaljeringsgrad, er der i det forgangne årti sket et skifte mod agentbaserede modeller. Agentbaserede modeller følger hver enkelt individ i populationen hele døgnet, og gør det muligt at evaluere detaljerede effekter som makroskopiske modeller ikke kan. Selvom agentbaserede modeller er i stand til at simulere store områder detaljeret, eksisterer det førnævnte gab i formåen mellem modeller for biltrafik og øvrige transportformer stadig, i hvert fald når man ser bort fra mikroskopiske modeller begrænset til et lille geografisk område.

Denne ph.d.-afhandling sigter mod at reducere dette gab ved at modellere detaljeret individuel adfærd i storskalaagentbaserede transportsimuleringmodeller. Afvejningen mellem detaljeringsgrad og beregningmæssig ydeevne er et vitalt fokuspunkt, da de udviklede modeller er tiltænkt anvendelse på storbyniveau. Selv med dette in mente, bidrager afhandlingen alsidigt til litteraturen gennem fem artikler vedrørende agentbaseret modellering indelt i tre dele: i) Outputvariabilitet i agentbaseret simulering af transportsystemer, ii) Agentbaseret passagerforsinkelsesmodellering med realtidsinformation og iii) Agentbaseret simulering af cykeltrafik og interaktion med biler.

Del I omhandler en af ulemperne ved de fleste agentbaserede simuleringsmodeller – deres store afhængighed af pseudo-tilfældige tal. Hvor de fleste aggregerede modeller kan slippe afsted med at angive valg gennem deres valgsandsynligheder, kræver de individuelle bestanddele af agentbaserede modeller, at eksplicitte valg træffes enten ved at benytte fuldt ud deterministiske, regelbaserede opsætninger eller ved at anvende pseudo-tilfældige tal til at udtrække sådanne valg stokastisk. Imidlertid, idet rækker af pseudo-tilfældige tal afgøres fuldtstændigt af deres initialiseringsværdier, vil en ændring af initialiseringsværdien føre til ændringer i modellerede valg og i sidste ende også ændre det endelige modeloutput.

Således kan outputtet fra stokastiske agentbaserede modeller betragtes som stikprøver fra – meget komplekse – sandsynlighedsfordelinger.

Den første artikel i afhandlingen undersøger variabilititen af sådanne outputs ved at foretage et eksperiment i MATSim med et storskalascenarie for Santiago, Chile med 100 forskellige kørsler baseret på 100 forskellige initialiseringsværdier. De resulterende daglige trafikmængder på netværkskanter analyseres, og det fremgår at variationen mellem initialiseringsværdier generelt er lav med tolerable variationskoefficienter. Dog er den relative forskel lejlighedsvis alvorlig sammenholdt med andre initialiseringsværdier, hvilket indikerer, at resultater baseret på en enkelt initialiseringværdi kan være misvisende. Endeligt finder studiet, at variationen mellem den sidste og næstsidste iteration er næsten fuldstændigt domineret af variationen på tværs af initialiseringsværdier pegende i retning af, at der bør bruges mere end en enkelt initialiseringsværdi, når resultater skal analyseres.

Del II af afhandlingen beskæftiger sig med agentbaseret passagerforsinkelsesmodellering når realtidsinformation er til stede. Kollektive transportsystemer er ofte upålidelige, og data fra tog og bussers positioner i realtid (AVL-data) kan bruges til at måle køretøjspunktligheden. På trods af at være ekstremt relevant bliver det sjældent evalueret, hvordan sådanne forsinkelser influererer passagerrejsetider, formentligt fordi opgaven er vanskelig, da passagerer har mulighed for at adaptere deres valg hen ad vejen, i særlig grad når realtidsinformation er tilgængelig.

Formålet med afhandlings *anden artikel* er at udarbejde en simuleringmodel, som baseret på faktiske køretøjsforsinkelse fra AVL-data kan beregne dertilhørende passagerforsinkelser i stor skala under forskellige niveauer af realtidsinformationstilgængelighed. Modellen lader passagerer søge efter nye alternativer hvert 150. sekund, mens de rejser gennem systemet, baseret på den til enhver tid tilgængelige realtidsinformation. Modellen er anvendt på det kollektive transportsystem for Hovedstadsområdet, og 812,359 daglige ture er modelleret for 65 dage, hvor AVL-data blev indsamlet for tog- og busnetværket. Studiet viser, hvordan den særligt irregulære jernbanelinje Kystbanen resulterer i, at mange passagerer til forfølger alternativer ruter, og at information om bedre alternativer ofte forekommer på stationer med mange højklassede muligheder. I tråd med den eksisterende litteratur, vises det, at på trods af at nogle passagerer ankommer før tid, er passagerforsinkelser gennemsnitligt set langt større end de køretøjsforsinkelser, der skaber dem. Studiet finder yderligere, at passagerforsinkelser kan reduceres betragteligt ved at anvende realtidsinformation ved turens start og endnu mere ved løbende at anskaffe sig sådan information undervejs.

Del III består af tre artikler, der successivt udvikler en skræddersyet metodologi til at simulere cykeltrafik og interaktion med biltrafik på storbyniveu. Delens første artikel – afhandlingens *tredje artikel* – udvikler en beregningsmæssig hurtig metodolgi til realistisk at simulere trængsel på cykelstier. Da heterogenitet blandt cycklister er udtalt, er

det essentielt at tage overhalinger med i betragtning for at være i stand til at simulere cykeltrafik hensigtsmæssigt. Artiklen benytter videodata af lavintens cykeltrafik til at undersøge, hvordan ønskede hastigheder og præferencer for afstand til forankørende varierer på tværs af cyklister og estimerer passende sandsynlighedsfordelinger disse cyklistkarakteristika, der anvendes som input til modellen. Modellen er baseret på simple antagelser og tillader overhalinger ved at cyklister eksplicit vælger bane, når de ankommer til en netværkskant. Et flaskehalsnetværk er brugt til at teste metoden under et bredt spektrum at trafikintensiteter. I tråd med modellens hensigt viser det sig, at modellen er mere tilbøjelig til at påføre cyklister med høje ønskede hastigheder forøget rejsetid, hvorimod højere trafikintensiteter er nødvendige for konsekvent at forsinke cyklister med præference for lavere hastigheder. Alligevel kan fundamentaldiagrammerne udledt af småskalaeksperimenterne valideres af videodata af 3,763 cyklister fra Dronning Louises Bro i en morgenmyldretidstime. Forskningen foretaget i artiklen klarlægger, at det er essentielt at anerkende cyklistheterogenitet for at simulere cykeltrafik realistisk.

Den fjerde artikel er en naturlig udvidelse af den den foregående artikel. Før denne artikel var der ingen eksempler på dedikerede vejvalgsmodeller for cykeltrafik med feedback mellem rutevalg of rejsetider i litteraturen. Det gør artiklen op med ved at bruge metodologien fra den tredje artikel til at simulere cykeltrafik og integrere den i en vejvalgsmodel med meningsfyldt feedback mellem udbud og efterspørgsel. Metoden er implementeret i MATSim og anvendt på et storskalacasestudie for Hovedstadsområdet med 1,082,958 cykelture. Selvom den øgede rejsetid er lav for cyklister sammenlignet med andre køretøjstyper, viser studiet gennem scenarier med værre og bedre infrastruktur, at dette i høj grad kan tilskrives den gode københavnske cykelinfrastruktur. Stadigvæk vises det, at forskelle i trafikmængder afviger signifikant fra hinanden i første og sidste iteration, hvilket demonstrerer, at det er nødvendigt med feedback fra netværket, når man modellerer cykeltrafik i byer med en høj andel af cyklister – og i endnu højere grad hvis infrastrukturen er utilstrækkelig.

Hvor de to forudgående artilker fokuserer på at simulere cykeltrafik på netværkskanter, adresserer *den femte og sidste artikel*, hvordan man modellerer netværkets kryds. I byer med en høj grad af separeret cykeltrafik er krydsene af særlig interesse, idet praktisk talt alle interaktioner mellem cykeltrafikken og motoriserede køretøjer finder sted heri. Artiklen formulerer en fælles vejvalgsmodel for bil- og cykeltrafik, der evner at modellere vigepligt i alle netværksknuder, og sikrer sig at ingen konfliktende bevægelser kan finde sted samtidigt. Den agentbaserede model er implementeret i MATSim og anvendt på det samme storskalacasestudie som den fjerde artikel, men med tilføjelsen af 3,210,685 bilture, 299,416 lastbilsture og et bilnetværk, der øger netværkstørrelsen til 572,935 kanter og 144,060 knuder, Simulationskørsler med vigepligtsnodemodellen og den originale MATSim-nodemodel er foretaget med tre forskellige opsætninger med bil- og lastbilstrafik, cykeltrafik og med alle tre køretøjstyper. Uden at at øge beregningstider nævneværdigt,

viser studiet, at rejsetider undervurderes, og at for meget trafik ledes ind gennem byen, når vigepligt i kryds ikke inkluderes. Det vises ydermere, at intermodale konflikter mellem cykeltrafik og motoriseret trafik i kryds forsinker trafikken mere end intramodale konflikter, hvilket understreger nødvendigheden af at modellere multi-modalitet samlet.

Sammenfattende har denne ph.d.-afhandling bidraget til litteraturen inden for agentbaseret simulation af urbane transportsystemer, strækkende sig over outputvariabilitet, passagerforsinkelser i kollektiv transport samt dynamisk vejvalgsmodellering af cykeltrafik og interaktioner med biltrafik. Dette inkluderer anbefalinger til fremtidig praksis ved anvendelse af agentbaseret modellering, men også ved at udvikle metodologier for nye typer af modeller for storskalamodellering af passagerforsinkelser og højintens cykeltrafik, som tillader at estimere effekter, der tidligere er blevet ignoreret, når transportsystemer for storbyer skulle planlægges og evalueres. Selvom der stadig er yderligere skridt at tage inden for alle de emner, der berøres af afhandlingens fem artikler, udgør deres bidrag betydelige forbedringer af ikke bare forståelsen men også den modelmæssige formåen af agentbaserede transportsimuleringsmodeller for multimodal urban trafik.

Contents

1 Introduction

This introductory chapter begins by an overall motivation of the PhD thesis in Section 1.1. Section 1.2 divides the thesis into three parts, and presents the motivation, aim and main contributions for each of the parts and associated papers. Finally, an outline of the remainder of thesis is provided in Section 1.3.

1.1 Background

Transport systems are complex systems where choices of users and network performance are highly interdependent. Urban transport systems have additional complexity due to their traffic generally consisting of a more diverse range of transport modes and more alternatives are available.

However, even in cities like Copenhagen that are praised for theirs high market shares of cyclists, car traffic still constitutes a serious part of the overall traffic. As cars are tremendously space inefficient, even a moderate proportion of the overall number of trips is enough to cause severe congestion. Congested time in Metropolitan Copenhagen is projected to reach 18.4 millions hours in 2025 (The Danish Commission on Congestion, 2013), and with the global and national urbanisation in mind projecting two thirds of the world's population to live in cities by 2050 (United Nations, 2019) and a growth in the number citizens in The City of Copenhagen of 24.1% since 2005 (The City of Copenhagen, 2020), there is no reason to believe that these numbers will drop without serious interventions.

Car traffic does not only cause congestion, though. It also contributes to green house gas emissions and local pollution, particularly undesirable in cities where the population density is higher. Despite having a large share of cyclists, urban road traffic is still a primary source of local pollution in cities like Copenhagen, as it was exemplified at the 2011 UCI World Championships where the concentration of ultrafine particles dropped by 30% during the seven day restriction on cars in central Copenhagen (Colville-Andersen, 2011; Miljøpunkt Indre By-Christianshavn, 2011).

To circumvent these tendencies, and as a means to facilitate sustainable development and preserve efficient transport systems, many cities have adopted strategies for increasing the use of bicycles and public transport (Giles-Corti et al., 2016). As an example, The City of Copenhagen (2012) has targeted having at least 75% of trips to, from or within the municipality to be conducted by bicycle, public transport or by walking.

Whereas the travel time effects of interventions targeting car traffic are reasonably understood in the literature through decades of research in transport models centred around car

traffic, equivalent models focusing on public transport users and cyclists in particular have received much less attention, especially concerning the bicycle traffic and the movements of cyclists within such (Twaddle et al., 2014). This is problematic, as when asked why they travel by bicycle 46% of cyclists in Copenhagen mention the bicycle being the faster option, only exceeded by being the easier option (55%) (The City of Copenhagen, 2019).

Furthermore, the lack of appropriate bicycle transport models is specifically pointed out as a barrier among stakeholders according to Aldred et al. (2019), who further highlights the importance of closing the gap allowing backing cycling projects with transport model predictions. Whereas the studies of long and short term effects on public health (Pucher and Dijkstra, 2003; Oja et al., 2011; Garrard et al., 2012; Götschi et al., 2016) and equity (Goodman et al., 2013) can also be used to advocate cycling, such effects are much harder to include in cost benefit analyses, especially since existing models do not fully manage to determine the hours and kilometres bicycled.

In parallel, there has been a tendency towards developing continuously more disaggregate models, to a point where – due to methodological and computational development – entire metropolitan areas can now be modelled in high detail agent-based simulation models (Lin et al., 2008; Wegener, 2011; Bazzan and Klügl, 2014; Wise et al., 2017). Although the principles of such models have existed for more than a quarter of a century (Jones, 1979; Axhausen and Gärling, 1992), the development of such models have intensified in the past decade (Bazzan and Klügl, 2014).

The motivation for agent-based simulation solutions approaching the microscopic level has been multifaceted and includes their suitability for modelling heterogeneous systems and interactions between agents (Bazzan and Klügl, 2014). However, there are still plenty of challenges that need to be addressed within the field. According to Bazzan and Klügl (2014), the main challenge of agent-based transport models is whether such models are reliable or chaotic. Similarly, Wegener (2011) points out the stochastic nature of the outputs as being a particular challenge for agent-based models.

One of the widely used agent-based traffic simulators is the open-source model MATSim (Horni et al., 2016) which has been used in a wide range of applications and case studies around the world. All of the papers in this PhD thesis are either explicitly using MATSim (Paper 1, Paper 4, and Paper 5), utilising elements of MATSim (Paper 2) or proposing a methodology suited for implementation in MATSim (Paper 3).

Whereas MATSim originally only allowed agents to make decisions between iterations, recent applications concerning e.g. autonomous taxis (Maciejewski and Bischoff, 2018), parking search (Bischoff et al., 2019), and rerouting in car traffic (Kaddoura and Nagel, 2018) and public transport (Leng and Corman, 2020) in case of incidents, have also allowed within-day decisions to be made.

This PhD thesis contributes to the development of microscopic individual on-the-go decisions within large-scale agent-based simulation by developing models for two very different applications – passenger delay modelling and bicycle traffic simulation – and applying them to large-scale case studies of Metropolitan Copenhagen. In addition, the stochastic variation that is deemed as one of the main point of criticism for agent-based simulation models is also studied.

1.2 Aim and main contributions

The overarching aim of this PhD thesis is through development of new methodologies and case studies to test the capabilities of large-scale agent-based models when faced with complex modelling tasks concerning multi-modal urban traffic. As opposed to most existing multi-modal approaches where each mode is treated separately, the thesis aims at also modelling inter-modal interactions through joint simulation. Striving for large-scale applicability, careful selection of which aspects to model at the microscopic level during the traffic simulation – and how to model them – is a cardinal point throughout the thesis. The idea is to draw inspiration from microscopic simulation, further develop and implement methodologies computationally efficiently, and apply them in large-scale applications.

This is done across five papers categorised into one of three research areas that each constitutes a part of the thesis:

I. Output variability in agent-based simulation of transport systems.

II. Agent-based passenger delay modelling with real-time information.

III. Agent-based simulation of bicycle traffic and interaction with cars.

The motivation and aim of each part and their associated papers are presented in the following subsections.

1.2.1 Output variability in agent-based simulation of transport systems

Part I examines output uncertainty in agent-based transport simulation models caused by using pseudo-random numbers to randomly perform explicit choices. Transport models are generally prone to uncertainty as assumptions are often simplified or the required input is uncertain or downright flawed (Manzo et al., 2015). However, models explicitly utilising random numbers have additional variability caused by the choice of the *random seed* that completely determines the sequence of random numbers. As opposed to the input and model assumptions, the random seed has no unique correct value why intrinsic uncertainty is unavoidable. The aim of this part is to investigate how the output of an agent-based transport simulation model varies across different sequences of such random numbers.

The part includes a single paper, *Output variability caused by random seeds in a multi-agent transport simulation model* published in *Procedia Computer Science* 130, 2018 (Paper 1), that performs a variability analysis of outputs of a large-scale case study of MATSim (Horni et al., 2016) based on the open data scenario of Santiago de Chile (Kickhöfer et al., 2016). The downscaled scenario with more than 650,000 agents is run 100 times with different random seeds but otherwise same settings in each run. The large number of repetitions is much larger than previous studies on the subject, allowing for a detailed analysis of variation within each run but – more importantly – also between different runs. This allows to answer whether the output of a single random seed is reliable or supplementary runs with other random seeds should be used when analysing results of such models.

The study has been presented at the *7th International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications (ABMTRANS 2018) in conjunction with the 9th International Conference on Ambient Systems, Networks and Technologies (ANT 2018)* in Porto, Portugal.

1.2.2 Agent-based passenger delay modelling with real-time information

The second paper, *Impacts of real-time information levels in public transport: A large-scale case study using an adaptive passenger path choice model* under review at *Transportation Research Part A: Policy and Practice* (Paper 2), constitutes Part II of the thesis. As opposed to the easily measurable vehicle delays that have been studied intensively, passenger delays and other passenger perspectives are generally underrepresented in the literature (Parbo et al., 2016). For instance, despite an increasing presence of real-time information, its impact on passenger delays is not thoroughly understood in large-scale transport systems (Brakewood and Watkins, 2019). Remarkable, as operating public transport systems in a way that allows passengers to travel through the system fast and reliably should be a main objective for transport operators and authorities, and tools for determining passenger travel times and delays would be a helpful tool for evaluating current performances in this regard.

The aim of this part is to develop a model that based on Automated Vehicle Location (AVL) data can model corresponding real-time information and adaptive passenger responses, such that delays for all passengers in a metropolitan multi-modal public transport system can be determined while taken the level of real-time information into account.

The paper builds on top of Paulsen et al. (2018) – a conference paper presented at the *4th Conference on Advanced Systems in Public Transport (CASPT) and Transit Data 2018* in Brisbane, Australia – but corrects several shortcomings. The old, heuristic, iterative approach was rejected in favour of a new, adaptive one-shot passenger path choice model considering various levels of real-time information that is sufficiently computationally efficient to model all 801,719 daily trips of every passenger in Metropolitan Copenhagen.

Modelling adaptive passenger path choices is generally important as intended paths may become infeasible along the way, and is a downright necessity when considering real-time information to which passengers may react en-route. The size of the case study in terms of network and trips is much larger than any existing multi-modal studies, and the inclusion of the entire public transport system provide passengers with a variety of possible alternatives through transfers. The paper also contributes to literature by modelling passenger delays of actual days in a large-scale multi-modal network by including 65 days of observed AVL data of trains and buses in the area, whereas earlier studies on the subject applying real-life AVL data only dealt with railways and few transfers (Nielsen et al., 2009; Lijesen, 2014). The study models passenger responses to five levels of real-time information and analyses the overall and marginal effects on the corresponding passenger delays.

An extended abstract using the improved methodology has been accepted for presentation at the *9th Symposium of the European Association for Research in Transportation (hEART 2020)* in Lyon, France.

1.2.3 Agent-based simulation of bicycle traffic and interaction with cars

The final part of the thesis, Part III, is concerned with large-scale simulation of bicycle traffic and consists of Papers 3, 4, and 5. Even though some cities such as Copenhagen have high shares of cyclists, bicycle traffic is still modelled way too simplistically in existing models. Demand unresponsive route choice effects such as gradient (Menghini et al., 2010), land-use (Prato et al., 2018), and surface (Prato et al., 2018), are decently covered in the literature, and can be estimated independent on the size of bicycle flows. Estimating demand sensitive modelling of travel times, on the other hand, requires data from the entirety of a bicycle traffic stream, and has received much less attention in the literature (Twaddle et al., 2014). Consequently, projects designed for improving bicycle traffic have a disadvantage when being evaluated due to important effects such as travel time savings being neglected.

In three successive stages, Part III of the thesis develops a novel large-scale traffic assignment model capable of simulating bicycle traffic realistically and computationally efficiently (Paper 3) and applies it in two large-scale case studies (Papers 4, and 5), one of which also includes joint modelling of car traffic and inter-modal interactions at intersections (Paper 5).

The first of these papers, *Fast or forced to follow: A speed heterogeneous approach to congested multi-lane bicycle traffic simulation* published in *Transportation Research Part B: Methodological* 127, 2019 (Paper 3), takes on the task of developing a methodology specially designed for simulating bicycle traffic on dedicated infrastructure. The absence of existing transport simulators capable of modelling congestion for such bicycle traffic means that traffic assignment models including cyclists have predominantly omitted

congestion effects altogether, assigning constant, demand independent travel times for cyclists. Flow-based methods designed for car traffic are not a viable solution either as they do not accommodate the high heterogeneity across cyclists and the good opportunities to overtake in bicycle traffic. Instead, this paper develops a tailor-made agent-based simulation model for bicycle traffic on dedicated infrastructure based on assumptions and characteristics more appropriate for bicycle traffic, and that can be estimated on low intensity bicycle traffic data. As opposed to very detailed microscopic models, the model should still be sufficiently computationally efficient to be used within a traffic assignment model for an entire metropolitan area.

An initial version of the methodology was presented at the *7th Symposium of the European Association for Research in Transportation (hEART 2018)* in Athens, Greece.

The second paper of this part, *Large-scale assignment of congested bicycle traffic using speed heterogeneous agents* published in *Procedia Computer Science* 151, 2019 (Paper 4), builds on top of the methodology from Paper 3. The bicycle traffic simulation model from Paper 3 is integrated with an individualised route choice model and implemented in MATSim (Horni et al., 2016). The resulting model is the first ever dedicated bicycle traffic assignment model with proper feedback capabilities between supply and demand. Three illustrative scenarios are designed for testing the model on a large-scale case study of Metropolitan Copenhagen with 1,082,958 trips assigned to a detailed bicycle network. The formulated model is intended for evaluating travel time effects of large interventions targeting bicycle traffic.

The study was presented at the *8th International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications (ABMTRANS 2019) in conjunction with the 10th International Conference on Ambient Systems, Networks and Technologies (ANT 2019)* in Leuven, Belgium,

Paper 5 entitled *Including right-of-way in a joint large-scale agent-based dynamic traffic assignment model for car and bicycle traffic* submitted to *Networks and Spatial Economics* further develops the research by extending the model to simulate interactions at intersections in a joint large-scale simulation model for car and bicycle traffic. Although some cities such as Copenhagen to a large degree separate bicycle traffic from the remaining traffic, interactions are bound to happen at intersections. Despite right-of-way being designed to favour cyclists, delays can still occur when cyclists have to turn left or cross larger roads, whereas turning car traffic has to yield for dense streams of bicycle traffic. The study contributes to literature by formulating and implementing in MATSim (Horni et al., 2016) an efficient method for simulating intra- and extra-modal interactions at intersections in a joint agent-based traffic assignment model for car and bicycle traffic, as well as analysing the impacts of multi-modal right-of-way in a large scale case study of Metropolitan Copenhagen with 4,593,059 daily trips on a network of 572,935 links and

144,060 nodes.

An extended abstract on the subject has been accepted for presentation at the now postponed *8th International Symposium on Dynamic Traffic Assignment (DTA 2020)* in Seattle, WA, USA.

1.3 Outline

The remainder of the PhD thesis consists of the three introduced parts and their papers as well as final concluding chapter. Part I on output variability of agent-based models contains Paper 1 in Chapter 2. Paper 2 found in Chapter 3 constitutes Part II concerning passenger delay modelling. Papers regarding agent-based simulation of congested bicycle traffic are found in Chapters 4-6 of Part III. Finally, Chapter 7 sums up the individual and mutual contributions and major takeaways from the five papers and points towards relevant research directions and policy implications outlined by the thesis.

References

Aldred, R., T. Watson, R. Lovelace, and J. Woodcock (2019). "Barriers to investing in cycling: Stakeholder views from England". In: *Transportation Research Part A: Policy and Practice* 128, pp. 149–159. DOI: 10.1016/j.tra.2017.11.003. URL: https://linkinghub.elsevier.com/retrieve/pii/S096585641730410X.

Axhausen, K. W. and T. Gärling (1992). "Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems". In: *Transport Reviews* 12.4, pp. 323–341. DOI: 10.1080/01441649208716826. URL: http://www.tandfonline.com/doi/abs/10.1080/01441649208716826.

Bazzan, A. L. C. and F. Klügl (2014). "A review on agent-based technology for traffic and transportation". In: *The Knowledge Engineering Review* 29.3, pp. 375–403. DOI: 10.1017/S0269888913000118. URL: https://www.cambridge.org/core/product/identifier/S0269888913000118/type/journal_article.

Bischoff, J., M. Maciejewski, T. Schlenther, and K. Nagel (2019). "Autonomous Vehicles and their Impact on Parking Search". In: *IEEE Intelligent Transportation Systems Magazine* 11.4, pp. 19–27. DOI: 10.1109/MITS.2018.2876566. URL: https://ieeexplore.ieee.org/document/8510918/.

Brakewood, C. and K. Watkins (2019). "A literature review of the passenger benefits of real-time transit information". In: *Transport Reviews* 39.3, pp. 327–356. DOI: 10.1080/01441647.2018.1472147.

Colville-Andersen, M. (2011). *Massive Fall in Air Pollution During World Championships*. URL: http://www.copenhagenize.com/2011/12/massive-fall-in-air-pollution-during.html.

Garrard, J., C. Rissel, and A. Bauman (2012). "Health Benefits of Cycling". In: *City Cycling*. Ed. by J. Pucher and R. Buehler. MIT Press. Chap. 3, pp. 31–55. URL: https://books.google.dk/books?id=226mCyz9JaEC&pg=PA31&hl=da&source=gbs_toc_r&cad=2#v=onepage&q&f=false.

Giles-Corti, B., A. Vernez-Moudon, R. Reis, G. Turrell, A. L. Dannenberg, H. Badland, S. Foster, M. Lowe, J. F. Sallis, M. Stevenson, and N. Owen (2016). "City planning and population health: a global challenge". In: *The Lancet* 388.10062, pp. 2912–2924. DOI: 10.1016/S0140-6736(16)30066-6. URL: https://linkinghub.elsevier.com/retrieve/pii/S0140673616300666.

Goodman, A., J. Panter, S. J. Sharp, and D. Ogilvie (2013). "Effectiveness and equity impacts of town-wide cycling initiatives in England: A longitudinal, controlled natural experimental study". In: *Social Science & Medicine* 97, pp. 228–237. DOI: 10.1016/j.socscimed.2013.08.030. URL: https://linkinghub.elsevier.com/retrieve/pii/S0277953613004826.

Götschi, T., J. Garrard, and B. Giles-Corti (2016). "Cycling as a Part of Daily Life: A Review of Health Perspectives". In: *Transport Reviews* 36.1, pp. 45–71. DOI: 10.1080/01441647.2015.1057877. URL: http://www.tandfonline.com/doi/full/10.1080/01441647.2015.1057877.

Horni, A., K. Nagel, and K. W. Axhausen, eds. (2016). *The Multi-Agent Transport Simulation MATSim*. London: Ubiquity Press. DOI: 10.5334/baw. URL: https://www.ubiquitypress.com/site/books/10.5334/baw/.

Jones, P. (1979). "New approaches to understanding travel behaviour: The human activity approach". In: *Behavioral Travel Modeling*. Ed. by D. Hensher and P. Stopher. London: Redwood Burn Ltd. Chap. 2, pp. 55–80.

Kaddoura, I. and K. Nagel (2018). "Using real-world traffic incident data in transport modeling". In: *Procedia Computer Science* 130, pp. 880–885. DOI: 10.1016/J.PROCS.2018.04.084. URL: https://www.sciencedirect.com/science/article/pii/S1877050918304460.

Kickhöfer, B., D. Hosse, K. Turner, and A. Tirachini (2016). *Creating an open MATSim scenario from open data: The case of Santiago de Chile*. VSP Working Paper 16-02. TU Berlin. URL: https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2016/16-02/.

Leng, N. and F. Corman (2020). "The role of information availability to passengers in public transport disruptions: An agent-based simulation approach". In: *Transportation Research Part A: Policy and Practice* 133, pp. 214–236. DOI: 10.1016/j.tra.2020.01.007. URL: https://linkinghub.elsevier.com/retrieve/pii/S0965856419305075.

Lijesen, M. G. (2014). "Optimal Traveler Responses to Stochastic Delays in Public Transport". In: *Transportation Science* 48.2, pp. 256–264. DOI: 10.1287/trsc.2013.0465. URL: http://pubsonline.informs.org/doi/abs/10.1287/trsc.2013.0465.

Lin, D.-Y., N. Eluru, S. T. Waller, and C. R. Bhat (2008). "Integration of Activity-Based Modeling and Dynamic Traffic Assignment". In: *Transportation Research Record: Journal of the Transportation Research Board* 2076.1, pp. 52–61. DOI: 10.3141/2076-06. URL: http://journals.sagepub.com/doi/10.3141/2076-06.

Maciejewski, M. and J. Bischoff (2018). "CONGESTION EFFECTS OF AUTONOMOUS TAXI FLEETS". In: *Transport* 33.4, pp. 971–980. DOI: 10.3846/16484142.2017.1347827. URL: https://journals.vgtu.lt/index.php/Transport/article/view/212.

Manzo, S., O. A. Nielsen, and C. G. Prato (2015). "How uncertainty in input and parameters influences transport model :output A four-stage model case-study". In: *Transport Policy* 38, pp. 64–72. DOI: 10.1016/j.tranpol.2014.12.004. URL: http://linkinghub.elsevier.com/retrieve/pii/S0967070X14002522.

Menghini, G., N. Carrasco, N. Schüssler, and K. Axhausen (2010). "Route choice of cyclists in Zurich". In: *Transportation Research Part A: Policy and Practice* 44.9, pp. 754–765. DOI: 10.1016/j.tra.2010.07.008. URL: http://linkinghub.elsevier.com/retrieve/pii/S0965856410001187.

Miljøpunkt Indre By-Christianshavn (2011). *VM i partikelmålinger*. Technical report. URL: http://www.a21.dk/pdf/aarsplaner/2011/partikelrapport_111208_b.pdf.

Nielsen, O. A., O. Landex, and R. D. Frederiksen (2009). "Passenger delay models for rail networks". In: *Schedule-Based Modeling of Transportation Networks*. Boston, MA: Springer US. Chap. 2, pp. 27–49. DOI: 10.1007/978-0-387-84812-9_2. URL: http://link.springer.com/10.1007/978-0-387-84812-9_2.

Oja, P., S. Titze, A. Bauman, B. de Geus, P. Krenn, B. Reger-Nash, and T. Kohlberger (2011). "Health benefits of cycling: A systematic review". In: *Scandinavian Journal of Medicine and Science in Sports* 21.4, pp. 496–509. DOI: 10.1111/j.1600-0838.2011.01299.x.

Parbo, J., O. A. Nielsen, and C. G. Prato (2016). "Passenger Perspectives in Railway Timetabling: A Literature Review". In: *Transport Reviews* 36.4, pp. 500–526. DOI: 10.1080/01441647.2015.1113574. URL: https://www.tandfonline.com/doi/full/10.1080/01441647.2015.1113574.

Paulsen, M., T. K. Rasmussen, and O. A. Nielsen (2018). "Modelling Railway-Induced Passenger Delays in Multi-Modal Public Transport Networks". In: *Proceedings for the 14th Conference on Advanced Systems in Public Transport and Transit Data 2018*. Ed. by M. D. Hickman. Brisbane, Australia.

Prato, C. G., K. Halldórsdóttir, and O. A. Nielsen (2018). "Evaluation of land-use and transport network effects on cyclists' route choices in the Copenhagen Region in value-of-distance space". In: *International Journal of Sustainable Transportation*, pp. 1–12. DOI: 10.1080/15568318.2018.1437236. URL: https://www.tandfonline.com/doi/full/10.1080/15568318.2018.1437236.

Pucher, J. and L. Dijkstra (2003). "Promoting Safe Walking and Cycling to Improve Public Health: Lessons From The Netherlands and Germany". In: *American Journal of Public Health* 93.9, pp. 1509–1516. DOI: 10.1016/j.ypmed.2009.07.028.

The City of Copenhagen (2012). *CPH 2025 Climate Plan*. The Technical and Environmental Administration. URL: https://kk.sites.itera.dk/apps/kk_pub2/pdf/983_jkP0ekKMyD.pdf.

The City of Copenhagen (2019). *Copenhagen - City of Cyclists*. The Bicycle Account 2018. The Technical and Environmental Administration. URL: https://kk.sites.itera.dk/apps/kk_pub2/pdf/1962_fe6a68275526.pdf.

The City of Copenhagen (2020). *Status på København 2019*. Den Tværgående Analyseenhed. URL: https://www.kk.dk/sites/default/files/status_paa_koebenhavn_2019.pdf.

The Danish Commission on Congestion (2013). *Mobilitet og fremkommelighed i hovedstaden*. Betænkning 1539. URL: https://www.trm.dk/media/2593/betaenkning-1539-mobilitet-og-fremkommelighed-i-hovedstaden-hovedrapport-netversion.pdf.

Twaddle, H., T. Schendzielorz, and O. Fakler (2014). "Bicycles in Urban Areas". In: *Transportation Research Record: Journal of the Transportation Research Board* 2434.1, pp. 140–146. DOI: 10.3141/2434-17. URL: http://journals.sagepub.com/doi/10.3141/2434-17.

United Nations (2019). *World Urbanization Prospects: The 2018 Revision*. Department of Economic and Social Affairs, Population Division. ST/ESA/SER.A/420. New York. DOI: 10.4054/demres.2005.12.9. URL: https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf.

Wegener, M. (2011). "From macro to micro - how much micro is too much?" In: *Transport Reviews* 31.2, pp. 161–177. DOI: 10.1080/01441647.2010.532883.

Wise, S., A. Crooks, and M. Batty (2017). "Transportation in Agent-Based Urban Modelling". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 10051 LNAI. Springer Verlag, pp. 129–148. DOI: 10.1007/978-3-319-51957-9_8. URL: http://link.springer.com/10.1007/978-3-319-51957-9_8.

I

Output variability in agent-based simulation of transport systems

2 Paper 1: Output variability caused by random seeds in a multi-agent transport simulation model

The following pages contain the article:

M. Paulsen, T. K. Rasmussen, and O. A. Nielsen (2018). "Output variability caused by random seeds in a multi-agent transport simulation model". In: *Procedia Computer Science* 130, pp. 850–857. DOI: 10.1016/j.procs.2018.04.078. URL: https://linkinghub.elsevier.com/retrieve/pii/S187705091830440X.

Please cite accordingly.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 130 (2018) 850–857

Procedia Computer Science

www.elsevier.com/locate/procedia

The 7th International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications (ABMTRANS 2018)

Output variability caused by random seeds in a multi-agent transport simulation model

Mads Paulsen[a,*], Thomas Kjær Rasmussen[a], Otto Anker Nielsen[a]

[a]*DTU Management Engineering, Technical University of Denmark, Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark*

Abstract

Dynamic transport simulators are intended to support decision makers in transport-related issues, and as such it is valuable that the random variability of their outputs is as small as possible. In this study we analyse the output variability caused by random seeds of a multi-agent transport simulator (MATSim) when applied to a case study of Santiago de Chile. Results based on 100 different random seeds shows that the relative accuracies of estimated link loads tend to increase with link load, but that relative errors of up to 10 % do occur even for links with large volumes. Although the proportion of links having large relative errors is roughly the same for all of the investigated seeds, it is shown the variations of individual link loads between seeds largely dominate variations between the two last iterations within a seed.

© 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the Conference Program Chairs.

Keywords: Transport Simulation, Stochasticity, Simulation Variability, MATSim

1. Introduction

Uncertainty in transport models is a well-known problem that needs to be addressed when using their outputs[1]. Whereas the uncertainty of traditional analytical transport models is caused by uncertainty of input parameters, models that rely on simulation have additional variability in their output due to the stochasticity of the series of pseudo-random numbers that are drawn. These series are determined by the so-called random seed.

Although some types of transport models use stochasticity/simulation for choice set generation, it is most widely used in dedicated transport simulators. The open-source software MATSim[2] is one example of such large-scale transport simulator used for analysing transport scenarios all across the world. As transport simulators are generally intended to facilitate policy support, it is of great interest to examine the output variability caused by random seeds in MATSim in order to determine whether this can potentially overshadow the impacts of suggested infrastructure investments under consideration.

This paper investigates the variability in output caused by random seeds for an open-data scenario of Santiago de Chile[3,4,5]. This is done by running the model repeatedly with different seeds using otherwise the same default

*Corresponding author. Tel.: +45 45 25 65 98.
E-mail address: madsp@dtu.dk

1877-0509 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the Conference Program Chairs.
10.1016/j.procs.2018.04.078

configurations and inputs. In the remainder of the paper, some additional background is initially given in section 2, followed by an introduction to the applied methodology in section 3. Section 4 presents how random numbers are generated in MATSim, while the case study and results are presented in sections 5 and 6, respectively. Section 7 discusses the findings and gives suggestions for further work.

2. Background

While a lot of studies have dealt with uncertainties in transport models in general — in particular uncertainties due to uncertainties of input – only a few studies have previously focused on the pure stochastic variability caused by the selection of random seeds in activity-based models.

Veldhuisen et al.[6] concluded that Monte Carlo errors are negligible when considering aggregate results of the microsimulation model RAMBLAS.

Castiglione et al.[7] used an approach where the partial mean of the first n seeds is compared to the final mean ($n = 100$) for the so-called San Francisco Model. Although this is also interesting, the partial mean is dependent on the order in which the seeds are used. Also it does not generally answer how large errors are when using a single seed, as only one of such cases is investigated (the first partial mean).

Lawe et al.[8] investigated the sensitivity to variations in random seed based on five different random seeds in TRANSIMS. Although based on a small sample, the coefficient of variation of link loads were small on average for all investigated links.

Also with TRANSIMS, Ziems et al.[9] follows the same approach as in Castiglione et al.[7] using 20 seeds, but extends the analysis by changing and keeping the random seeds of different parts of the simulation. Coefficients of variations of link loads for two investigated links were 1.7 % when changing the random seed of the entire model.

Cools et al.[10] investigated the variability of demand for each mode when using FEATHERS. Based on 200 different seeds the coefficient of variation was generally below 1 %.

Bekhor et al.[11] investigates the variability that arises from random components of the model specifications of the Tel Aviv Transportation Model. It also studies variability solely caused by random seeds by running various settings with three different seeds. The number of different seeds used is too low to quantify the findings, though.

Several studies by Nagel (et al.)[12,13,14] touches upon the issue of simulation errors caused by selection of random seeds. It is noted and visualised that such variability exists based on two different random seeds, but an in-depth investigation of the variation has not been published, although unpublished work by Raney et al. is mentioned[13].

One of the studies by Nagel et al.[14] mentions the need to further investigate the variability caused by random seeds in MATSim specifically. This is supported by a chapter[15] of the MATSim book[2], stating that quantities of interest from the output of the model can be found by averaging over a series of runs with different seeds. This is encouraged due to the *broken ergodicity*[16] of the model, meaning that once a random seed is selected some till then possible outcomes of the model are no longer reachable[15].

3. Methodology

In order to evaluate the variations caused by random seeds in a multi-agent transport simulator, a number of measures will be applied in this study. The first is the coefficient of variation, c_v, defined as the sample standard deviation divided by the sample mean, both considered across all seeds s in the set of seeds S.

We also introduce two additional measures. With x_s^l denoting the link load of link l within the set of links L when using seed $s \in S$, we introduce the following notation with Iverson brackets to denote the empirical probability of the link, $l \in L$, having a link load that is more than q times off its sample mean across all seeds, $\bar{x}^l$,

$$r_q^l = \frac{\sum_{s \in S} \left[\frac{|x_l^s - \bar{x}_l|}{\bar{x}_l} > q \right]}{|S|}, \qquad l \in L,\ q \geq 0. \tag{1}$$

Likewise, we denote the proportion of links having link loads further than q times away from their corresponding empirical mean values when the seed, $s \in S$, was used by,

$$r_q^s = \frac{\sum_{l \in L} \left[\frac{|x_l^s - \bar{x}_l|}{\bar{x}_l} > q \right]}{|L|}, \qquad s \in S,\ q > 0. \tag{2}$$

By considering r_q^l and r_q^s across all $l \in L$ and $s \in S$, respectively, a series of values, $\mathbf{r}_q^L$ and $\mathbf{r}_q^S$, is obtained.

Furthermore, we introduce the terms within-seed variation, W_l, and the between-seed variation, B_l, for a link $l \in L$,

$$W_l = \frac{1}{|S|} \sum_{s \in S} \left(x_l^{s,|I|-1} - x_l^{s,|I|} \right)^2, \qquad l \in L,$$

$$B_l = \frac{1}{|S| - 1} \sum_{s \in S} \left(x_l^{s,|I|} - \bar{x}_l^{|I|} \right)^2, \qquad l \in L.$$

The calculation of the within-seed variation utilises that the (ex post) expected value of iteration $|I| - 1$ is the value obtained in the last iteration, $|I|$. For the between seed variation, we consider that each seed might converge to a different solution, why the expected value is found as the mean across all seeds $s \in S$ in iteration $|I|$, i.e. $\bar{x}_l^{|I|} = \frac{1}{|S|} \sum_{s \in S} x_l^{s,|I|}$. We can compare B_l and W_l by calculating $\tilde{R}_l$,

$$\tilde{R}_l = \sqrt{\frac{B_l + W_l}{W_l}}, \qquad l \in L.$$

If this is close to 1, then the between-seed variation is negligible, whereas the opposite is true for large values of $\tilde{R}_l$.

4. MATSim: A Multi-Agent Transport Simulator

MATSim[2] is an activity- and individual-based simulation model, where each agent aims at obtaining the highest possible value of their scoring function. It uses a co-evolutionary algorithm (Algorithm 1[14]) to reach its final output.

Algorithm 1 Co-evolutionary, population based search

1. **Initiation:** Generate at least one plan for every agent.
2. **Iterations:** Repeat the following many times.
   (a) **Plan Selection:** Select one plan for every agent.
   (b) **Plan Scoring:** Obtain a score for every agent's selected plan by executing all selected plans simultaneously in a mobility simulator and attach a performance measure (score) to each executed plan.
   (c) **Plan Innovation:** Generate new plans for some of the agents by mutating existing plans or re-routing.

The algorithm utilises on random numbers on several occasions. Firstly, they are used to determine the order in which to handle in-going links of each node in every time step of the mobility simulator. Here any link is chosen with a probability proportional to its capacity.

Secondly, they are used to determine which agents use which plan mutation strategy and plan selection strategy in each iteration. In each iteration 30 % of agents are selected to generate a new plan for their choice set – half of these by re-routing (new shortest path searches), the other half by substituting the mode of a discrete-uniformly randomly selected sub-tour by another mode from the pool of possible modes. If the generation of the plan forces the choice set to exceed the maximum number of plans, the plan with the worst score is removed from the choice set.

Thirdly, the 70 % of agents who do not generate a new plan, draw a plan, P^*, from their choice set with discrete uniform probability, and selects this plan with probability $\min\left\{1, 0.01 \cdot e^{\frac{U_{P^*} - U_{P^C}}{2}}\right\}$, where U_{P^*} and U_{P^C} are the latest

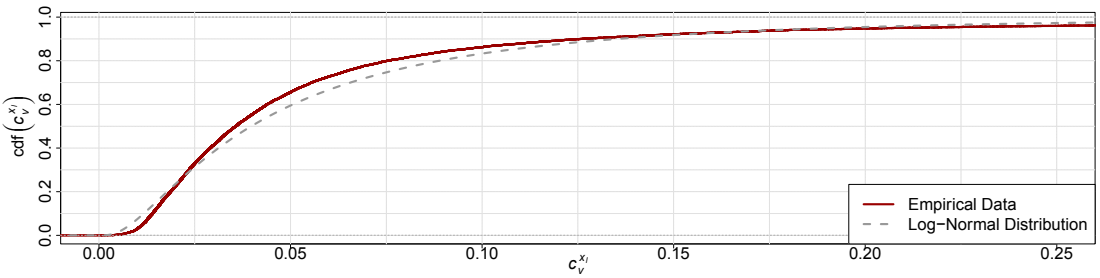


Fig. 1. Empirical cumulative distribution functios of the coefficient of variation of link loads across all links $l \in L$.

scores of the drawn plan, P^*, and the currently selected plan, P^C, respectively. If P^* is not chosen, the agent will execute the current plan again in the next iteration.

Finally, in order to have a larger diversity on the routes suggested when re-routing, the utility of money is drawn from a log-normal distribution when initiating a router.[17]

The random numbers are thus predominantly used for choosing between reasonable discrete alternatives. Therefore, MATSim does not at seem to be particularly vulnerable to variability arising from random seeds a first glance.

With the role of random numbers in MATSim established, we now turn to discuss how pseudo-random numbers are created in MATSim. They are to a large extent based on Java's built-in Linear Congruential Generator (LCG). The traditional LCG[18] is of the form,

$$X_{n+1} = aX_n + c \mod m,$$

where X_n is the latest draw, X_{n+1} is the next draw, and a, c and m are parameters having the values $m = 2^{48}$, $a = 25,214,903,917$, and $c = 11$ in Java. In order to get a uniformly distributed number X_{n+1} is divided by m.

MATSim offers a few extensions to Java's built-in random number generation. First of all, when a random number generator is instantiated, it immediately draws and throws away the first 100 draws, as these are found not to be sufficiently random.[15] Secondly, whenever a new random number generator is created from an existing one, the value of an internal counter is multiplied by 23 and added to the latest used seed before discarding the next 100 draws of the newly created instance. The initial random number generator is established with the globally assigned random seed as X_0, meaning that this number completely determines the sequence of all other random numbers used in the run.

5. Case Study

The model has been run for 100 different initial global random seeds on the Santiago de Chile open data scenario[3,4,5]. The scenario includes a toll system on selected roads, and schedule-based public transport running on dedicated public transport links. The default "out-of-the-box" configurations set by the developers of the scenario have been used, including a population sample of 10 % corresponding to 665,201 agents, a network consisting of 22,981 unidirectional links for car traffic, and 100 iterations per seed. Choice sets for every agent are locked after 80 iterations, at which point MSA is enabled for the (at most) five plans per agent. Considered modes for substitution of a sub-tour are walk, public transport, and – if available to the agent – car.

6. Results

The results show that coefficients of variation of link loads (dedicated public transport links excluded), $c_v^{x_l}$, are generally quite low with 65.7 % of the links not exceeding 5 % (see Figure 1). The distribution has a skewness of 16.7 and has some resemblance to a log-normal distribution, which is also shown in the figure.

The correlations between the mean ($\bar{x}_l$), the standard deviation ($\hat{\sigma}^{x_l}$), and the coefficient of variation ($c_v^{x_l}$) across all links $l \in L$ are shown in Table 1. It is seen that there is a strong correlation between $\bar{x}_l$ and $\hat{\sigma}^{x_l}$ (0.849) indicating

Table 1. Sample Pearson correlation coefficient between the mean, the standard deviation, and the coefficient of variation of link loads.

	Mean ($\bar{x}_l$)	Standard Deviation ($\hat{\sigma}^{x_l}$)	Coefficient of Variation $\left(c_v^{x_l}\right)$
Mean ($\bar{x}_l$)	1	0.849	-0.113
Standard Deviation ($\hat{\sigma}^{x_l}$)	0.849	1	-0.115
Coefficient of Variation $\left(c_v^{x_l}\right)$	-0.113	-0.115	1

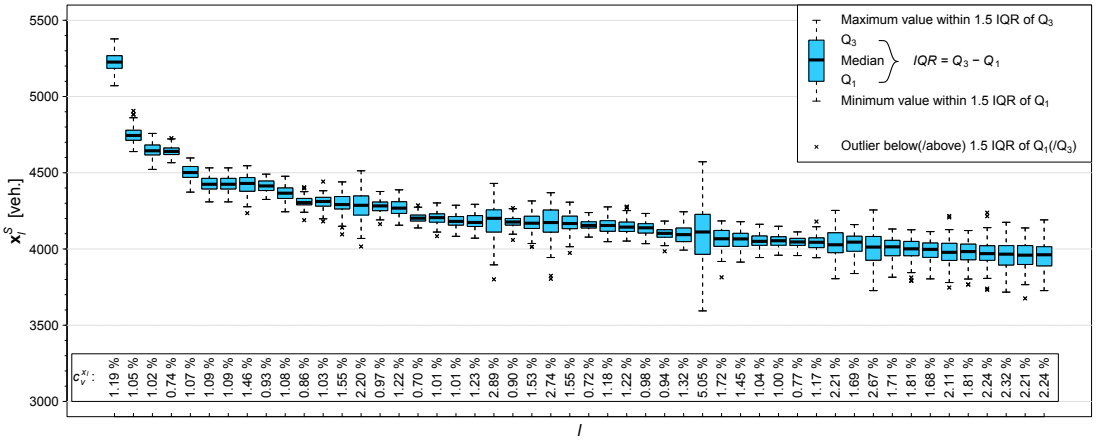


Fig. 2. Boxplots of link loads $\left(\mathbf{x}_l^S\right)$ for the 50 busiest links and their corresponding coefficient of variation $\left(c_v^{x_l}\right)$.

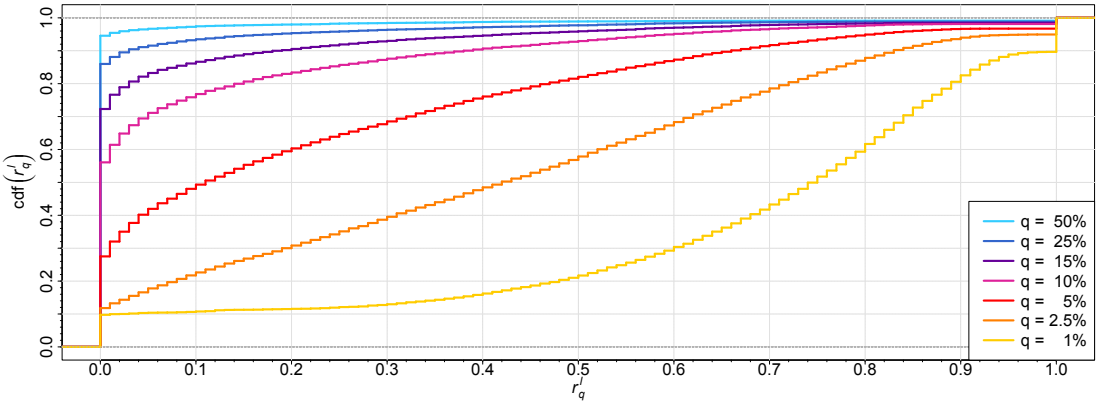


Fig. 3. Empirical cumulative distribution function of $\mathbf{r}_q^L$. for various values of q.

that using $c_v^{x_l}$ as a measure seems valid. However, since the correlation between $(\bar{x}_l)$ and $c_v^{x_l}$ is negative (-0.113), $c_v^{x_l}$ is generally expected to be a little smaller for links with high flows.

In Figure 2 it is seen that the coefficient of variation generally is about 1 % for the 50 busiest links, but that some links have a significantly higher $c_v^{x_l}$. The figure also shows that the ranges for individual links seem to be large, indicating that for some seeds the error can be notably higher than a few percent.

As mentioned in section 3, $\mathbf{r}_\mathbf{q}^\mathbf{L}$ holds the empirical probabilities of each link having a relative error larger than q when using a single seed. The corresponding cumulative distribution function is found in Figure 3 for various values

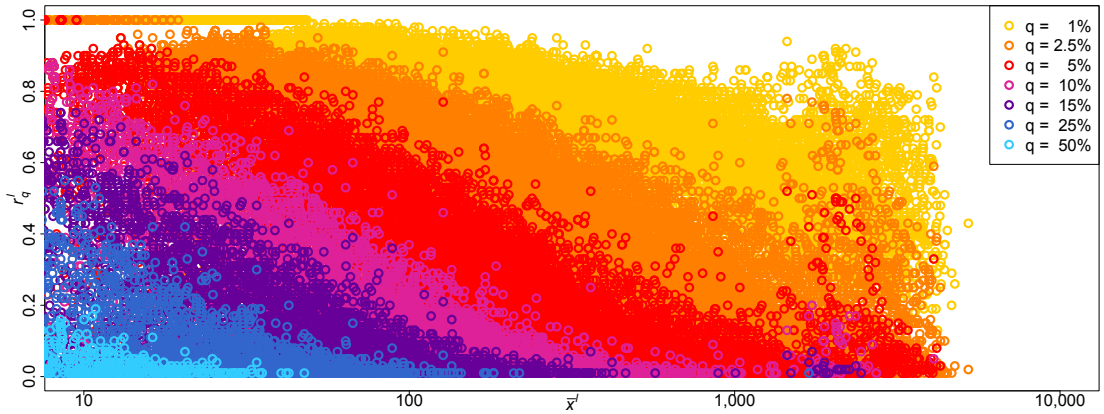


Fig. 4. Scatter plot of r_q^l against the average link load $\left(\bar{x}^l\right)$ on a logarithmic axis for all links $l \in L$ grouped by q for $r_q^l > 0$, plotted in the same drawing order as the legend.

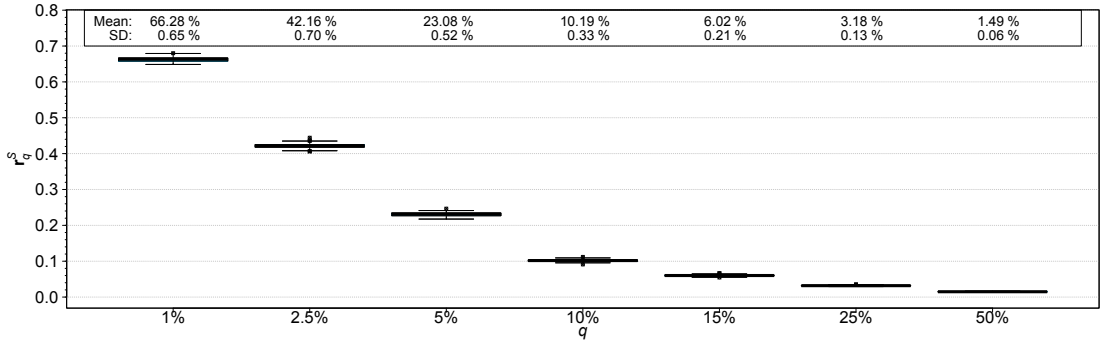


Fig. 5. Boxplots of $\mathbf{r}_q^S$ grouped by q.

of q. Whereas very low errors [1 %, 2.5 %[are frequent, relative errors of more than 50 % never occurred for 94.6 % of the links. Still, it is seen that 17.9 % of the links had a relative error exceeding 5 % in at least half of the runs.

Figure 4, however, shows that the extreme errors are primarily seen for less busy links. An interesting phenomenon occurs for links with a link load between 1,000 and 10,000, though, as the probabilities of 15 %, 10 %, and especially 5 % and 2.5 % errors seem to be larger than for the links with a mean link load just below 1,000.

In the above we have found that at the link level, there is a risk of getting relatively high deviations when only using a single seed value. In Figure 5 we consider each seed as a whole across all links, and find the corresponding proportion of links having a relative error larger than q, $\mathbf{r}_\mathbf{q}^\mathbf{S}$. It is seen that the seeds generally perform equally well, but that about 24 % of links are expected to have a link load deviating at least 5 % from their empirical mean.

So far we have established that between-seed variation is in fact considerable, however, without comparing it to other types of known variations. This is done by calculating $\tilde{R}_l$ as described in section 3. The empirical cumulative distribution function of $\tilde{R}_l$ across all $l \in L$ (Figure 6) shows that between seed variation greatly dominates the within seed variation for almost every link. This trend is even more evident for larger links, as $\tilde{R}_l$ is at least 4 for 90.2 % of links with a link load above 10 (see Figure 7).

Variability of mode markets shares has also been investigated, but was found to be negligible with the largest coefficient of variation being $8 \cdot 10^{-4}$. Variability of travel speeds, on the other hand, showed a pattern that was nearly indistinguishable from that of link loads.

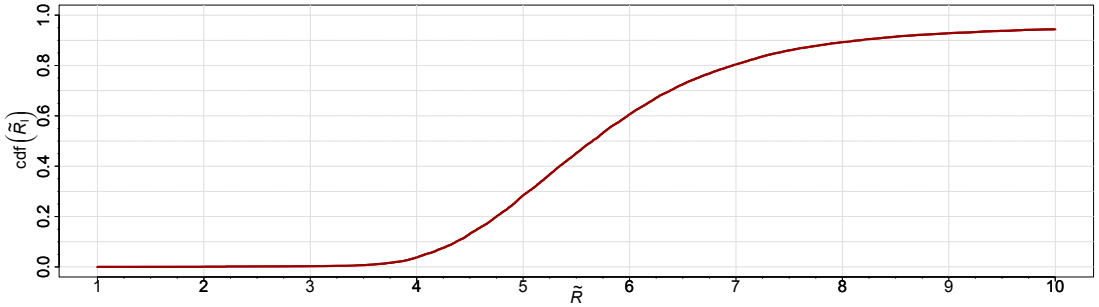


Fig. 6. Empirical cimulative distribution function of $\tilde{R}$.

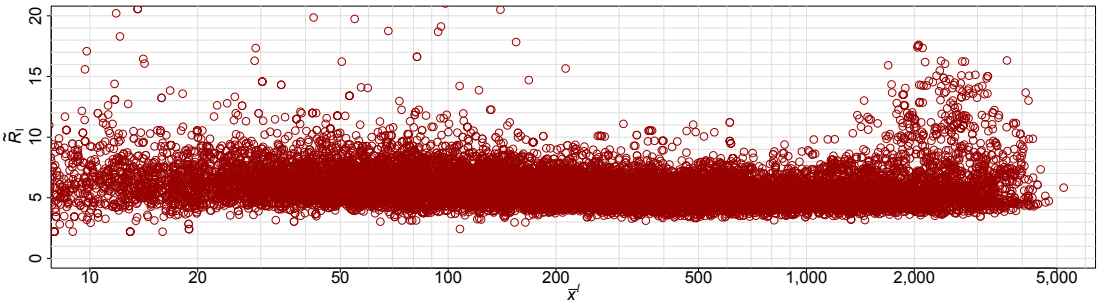


Fig. 7. Scatter plot of $\tilde{R}_l$ against $\bar{x}_l$ in a logarithmic axis.

7. Discussion and Future Work

The results of this study indicate that the potential relative error when using a single seed is generally a decreasing function of the amount of traffic on the link. However, among the largest links there seems to be a few links that are particularly error-prone. This should give rise to some concern as in many cases such busy links are the most interesting for policy support purposes. A geographical analysis of the links with a high coefficient of variation would therefore be a natural extension of the current work.

The finding that the between-seed variations largely dominate the within-seed variations suggests that broken ergodicity might in fact be a problem that needs to be addressed in transport simulators dealing with discrete choices of agents. Especially, since the tests in this paper were conducted using a low number of iterations (it is not uncommon for MATSim applications to use as much as 1,000 iterations) and thus were expected to have a relatively high level of within-seed variation.

One way to overcome this problem could be to average the results across multiple runs with different seeds, however at the theoretical risk that the resulting (averaged) solution is not an equilibrium solution. Alternatively, it may be favourable to present the results as a distribution by running the model multiple times with different seeds as proposed in the literature [15]. Especially since the results of this paper suggest that the information gained by such additional runs exceeds that of running supplementary iterations of a single seed.

Representing results as distributions instead of point estimates would be of particular interest for project appraisal purposes. Here the additional information would contribute to a further understanding of the uncertainty of the responses to a proposed project. Potentially this could answer whether the effects of infrastructural changes risk being overshadowed by between-seed variations. Running tests under such project appraisal-like circumstances with a base network and an alternative network is thus recommended for future work.

Another method previously proposed in the literature [19,20] aims at producing a more stable model outcome in the first place. This is, however, beyond the scope of this paper.

8. Acknowledgements

A special thanks to the developers of the Santiago de Chile scenario Benjamin Kickhöfer and Alejandro Tarachini, as well as the creators of MATSim, Kai Nagel and Kay Axhausen. Without the open access of the scenario and model this study would not have been possible.

References

1. Rasouli, S., Timmermans, H.. Uncertainty in travel demand forecasting models: literature review and research agenda. *Transportation Letters* 2012;**4**(1):55–73. URL: http://www.tandfonline.com/doi/full/10.3328/TL.2012.04.01.55-73. doi:10.3328/TL.2012.04.01.55-73.
2. Horni, A., Nagel, K., Axhausen, K.W., editors. *The Multi-Agent Transport Simulation MATSim*. London: Ubiquity Press; 2016. URL: https://www.ubiquitypress.com/site/books/10.5334/baw/. doi:10.5334/baw.
3. Kickhöfer, B., Hosse, D., Turner, K., Tirachini, A.. Creating an open MATSim scenario from open data: The case of Santiago de Chile. *VSP Working Paper 16-02* 2016;URL: https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2016/16-02/.
4. Kickhöfer, B., Tirachini, A.. Santiago de Chile. In: Horni, A., Nagel, K., Axhausen, K.W., editors. *The Multi-Agent Transport Simulation MATSim*; chap. 84. London: Ubiquity Press; 2016, p. 547–550.
5. Camus, L., Tirachini, A., Kickhöfer, B.. The MATSim Santiago open data model: development and first applications. *Presentation at the MATSim User Meeting - Haifa, Israel* 2017;.
6. Veldhuisen, J., Timmermans, H., Kapoen, L.. Microsimulation Model of Activity-Travel Patterns and Traffic Flows: Specification, Validation Tests, and Monte Carlo Error. *Transportation Research Record: Journal of the Transportation Research Board* 2000;**1706**:126–135. URL: http://trrjournalonline.trb.org/doi/10.3141/1706-15. doi:10.3141/1706-15.
7. Castiglione, J., Freedman, J., Bradley, M.. Systematic Investigation of Variability due to Random Simulation Error in an Activity-Based Microsimulation Forecasting Model. *Transportation Research Record: Journal of the Transportation Research Board* 2003;**1831**:76–88. URL: http://trrjournalonline.trb.org/doi/10.3141/1831-09. doi:10.3141/1831-09.
8. Lawe, S., Lobb, J., Sadek, A., Huang, S., Xie, C.. TRANSIMS Implementation in Chittenden County, Vermont. *Transportation Research Record: Journal of the Transportation Research Board* 2009;**2132**:113–121. URL: http://trrjournalonline.trb.org/doi/10.3141/2132-13. doi:10.3141/2132-13.
9. Ziems, S.E., Sana, B., Plotz, J., Pendyala, R.M.. Stochastic Variability in Microsimulation modeling Results and Convergence og Corridor-Level Characteristics. In: *Transportation Research Board 90th Annual Meeting*; vol. 11-3560. 2011, URL: http://rampendyala.weebly.com/uploads/5/0/5/4/5054275/stochasticvariation_ziemssanaplotzpendyala_trb2011.pdf.
10. Cools, M., Kochan, B., Bellemans, T., Janssens, D., Wets, G.. Assessment of the effect of micro-simulation error on key travel indices: Evidence from the activity-based model Feathers. In: *90th Annual Meeting of the Transportation Research Board*. Washington D.C.; 2011, URL: http://orbi.ulg.ac.be/bitstream/2268/134326/1/CETAL-MICROSIML-NOV2010.pdf.
11. Bekhor, S., Kheifits, L., Sorani, M.. Stability analysis of activity-based models: case study of the Tel Aviv transportation model. *European Journal of Transport and Infrastructure Research* 2014;**14**(4):311–331.
12. Nagel, K.. Experiences with iterated traffic microsimulations in Dallas. In: Schreckenberg, M., Wolf, D.E., editors. *Traffic and granular flow '97*. Heidelberg: Springer; 1997, p. 199–214. URL: https://arxiv.org/pdf/adap-org/9712001.pdf. arXiv:9712001v1.
13. Nagel, K., Rickert, M., Simon, P.M., Pieck, M.. The dynamics of iterated transportation simulations. *eprint arXiv:nlin/0002040 - Earlier version presented at the TRIannual Symposium on Transportation ANalysis (TRISTAN-III) in San Juan, Puerto Rico, 1998* 2000; URL: https://arxiv.org/pdf/nlin/0002040.pdf. arXiv:0002040.
14. Nagel, K., Flötteröd, G.. Agent-based traffic assignment: going from trips to behavioral travelers. In: Bhat, C.R., Pendyala, R.M., editors. *Travel Behaviour Research in an Evolving World–Selected papers from the 12th international conference on travel behaviour research*. Jaipur, India: International Association for Travel Behaviour Research; 2012, p. 261–294.
15. Flötteröd, G.. MATSim as a Monte-Carlo Engine. In: Horni, A., Nagel, K., Axhausen, K.W., editors. *The Multi-Agent Transport Simulation MATSim*; chap. 48. London: Ubiquity Press; 2016, p. 361–370.
16. Palmer, R.. Broken ergodicity. *Advances in Physics* 1982;**31**(6):669–735.
17. Nagel, K., Kickhöfer, B., Joubert, J.W.. Heterogeneous Tolls and Values of Time in Multi-agent Transport Simulation. *Procedia Computer Science* 2014;**32**(d):762–768.
18. Knuth, D.E.. *The Art of Computer Programming: Seminumerical Algorithms*; vol. 2. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.; 3rd ed.; 1997. ISBN 0-201-89684-2.
19. Bowman, J.L., Bradley, M.A., Gibb, J.. The Sacramento Activity-Based Travel Demand Model: Estimation and Validation Results. In: *Proceedings of the European Transport Conference (ETC) 2006*. Strasbourg, France: Association for European Transport; 2006, URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1017.6668&rep=rep1&type=pdf. doi:10.1.1.1017.6668.
20. Gibb, J., Bowman, J.L.. Convergence of an activity-based travel model system to equilibrium: Experimental design and findings. In: *Proceedings of the 11th National Transportation Planning and Application Conference of the Transportation Research Board*. Daytona Beach, FL; 2007, URL: http://www.trbappcon.org/2007conf/files/049Gibbfinal.pdf.

II

Agent-based passenger delay modelling with real-time information

3 Paper 2: Evaluation of real-time passenger information: A large-scale case study using an adaptive passenger path choice model

The following pages contain the article:

M. Paulsen, T. K. Rasmussen, and O. A. Nielsen (2020). "Impacts of real-time information levels in public transport: A large-scale case study using an adaptive passenger path choice model". Under review at Transportation Research Part A: Policy and Practice.

Please cite accordingly.

Impacts of real-time information levels in public transport: A large-scale case study using an adaptive passenger path choice model

Mads Paulsen[a,*], Thomas Kjær Rasmussen[a], Otto Anker Nielsen[a]

[a]*Transport Division, Department of Technology, Management and Economics, Technical University of Denmark, Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark*

Abstract

Public transport services are often uncertain, causing passengers' travel times and routes to vary from day to day. However, since door-to-door passenger delays depends on both intended and realised routes, they are difficult to calculate, as opposed to vehicle delays which can be derived directly from the widely available Automated Vehicle Location (AVL) data of the public transport system. In this study we use three months of such historical AVL data to calculate corresponding realised routes and passengers delays for more than 800,000 daily trips in a large-scale, multi-modal transport network of the Metropolitan Area of Copenhagen by applying an adaptive passenger path choice model. The proposed model allows analysing five different levels of real-time information provision, ranging from no information at all to global real-time information being available everywhere. The results show that variability of passengers' travel time is considerable and much larger than that of the public transport vehicles. It is also shown that obtaining global real-time information at the beginning of the trip reduces passengers delay dramatically, although still being inferior to receiving such along the trip. Additionally, being able to automatically obtain real-time passenger information while walking and being on-board public transport services was found not to lead to considerable improvements compared to acquiring such information manually while waiting at stops.

Keywords: Passenger Delays, Passenger Path Choice Model, Public Transport, Real-Time Information, AVL Data, Agent-Based Simulation

*Corresponding author
Email address: madsp@dtu.dk (Mads Paulsen)

Manuscript submitted to Transportation Research Part A: Policy and Practice *May 27, 2020*

1. Introduction

Public transport vehicles being delayed is a more or less an inherent part of operating a public transport system. Obviously, it would make sense for public transport planners and operators to focus on reducing the door-to-door passenger delays that passengers incur when travelling through the public transport system. However, despite operating vehicles in the public transport system is only a means to fulfil the overall task of allowing passengers to travel through the system, the punctuality and reliability of vehicle delays is often the overshadowing focus for public transport planners and operators when evaluating the performance of the system (Parbo et al., 2016). Vehicle delays can be extracted directly from Automated Vehicle Location (AVL) data making them easily accessible and unambiguous.

Passenger delays, on the other hand, are inherently difficult to calculate as vehicle delays often affects passenger delays in a non-linear fashion when passengers transfer between different services. In such cases, delays can even lead to a travel time gain if an earlier departure is sufficiently delayed to depart after the passenger arrives at a stop (Nielsen et al., 2009). They will, however, more often lead to increased travel times, naturally when the last leg of the path is delayed, but also when delays cause passengers to arrive at a stop later than expected and thereby miss the connection to an intended departure. This will cause the passenger to wait for the next departure or to search for an alternative route. As such, it is very complicated to determine the travel time of public transport passengers, as not only the public travel services but also the routes of passengers may deviate from what was initially intended – and may be influenced by the information that is offered to them.

This means that in order to calculate passenger delays, an adaptive passenger path choice model that can determine passengers' intended paths as well as their path choices when faced with real-time information of current delays in the system is required. Hickman (1993) was the first to introduce an adaptive passenger path choice model that supported real-time passenger information. Although more advanced methodologies have since been developed, and larger case studies have been carried out, no studies have calculated passenger delays of adaptive passengers for an entire metropolitan area without the use of additional data sources such as Automated Fare Collection (AFC) (Antos and Eichler, 2016; Sun et al., 2016b,a) or smartphone data (Carrel et al., 2015). Neither has the effects of real-time information on passengers' choices and travel times been evaluated on the basis of actual delays of an entire metropolitan area in the 27 years that have passed.

This study introduces an adaptive passenger path choice model for a stochastic schedule-based public transport network with no a-priori knowledge of the distributions of travel time uncertainties. The model is capable of handling an entire metropolitan area in manageable time by exploiting the computational gains of restricting all passengers to only be able to reconsider their options at the same moments in time (every 150 seconds). This allows the system to be solved as a series of successive dynamic passenger path choice models where locations of passengers and public transport vehicles are updated accordingly in between assignments.

The study will model five real-time information levels, and analyse the marginal effect of increasing the real-time information offered to passengers from one level to the next.

R0 *No information*: Agents will pursue their intended path, only allowing temporal adaptations.

R1 *Pre-trip information*: Agents search for the shortest path at the beginning of their trip using the real-time information available at that point.

R2 *Information at stops*: In addition to potentially updating their path at the beginning of their trip (R1), agents can also adapt to real-time information while waiting at stops/stations (but not on-board a service).

R3 *Information everywhere*: No restriction on where agents can adapt, meaning that they also search for better alternatives using real-time information while walking and while on-board public transport services.

R∞ *Perfect information*: As opposed to the other real-time information levels, the passengers know all past, current, and future delays in advance, allowing always choosing the optimal path a priori without en-route adaptation.

It may seem tedious to acquire the amount of information required by R3, but – as (Nökel and Wekeck, 2009; Estrada et al., 2015; Zargayouna et al., 2018) suggest – is achievable if thought of as an app continuously performing searches and providing notifications when better alternatives emerge. In this regard, R1 (Pre-trip) and R2 (At stops) can be thought of as using a similar app, but where users manually have to search for new paths at the origin (R1,R2) and at stops (R2). R∞ is obviously unrealistic, and would require the ability to predict all future delays accurately at the beginning of the day. It is solely included in order to establish a hypothetical lower bound for the passenger delays of each trip.

For each of the five real-time information levels we will use realised timetables constructed from 65 weekdays of actual AVL data for trains and buses in the metropolitan area of Copenhagen, and combine them with an adaptive passenger path choice model in order to determine *realised* routes and travel times for 801,719 trips daily trips in the area. By comparing these to the corresponding *intended* routes and travel times based on the planned timetable, the door-to-door *passenger delays* can be determined, see Figure 1.

Our contribution is threefold. Firstly, we model real-time information and door-to-door passenger delays for an entire metropolitan area. Secondly, we simulate passenger behaviour of actual days by applying observed AVL data of trains and buses. Finally, we model the impact of specific levels of real-time information that no prior studies have been able to isolate.

The remainder of the paper is structured as follows. Section 2 contextualises our contributions by reviewing the literature on adaptive passenger choice models with real-time information. Section 3 provides a detailed description of our proposed methodology for adaptive passenger assignment model considering real-time information. Section 4 tests the methodology on a large-scale case study of the Copenhagen metropolitan area, focusing on the effect of providing real-time information to passengers. Section 5 concludes and outlines directions for future work.

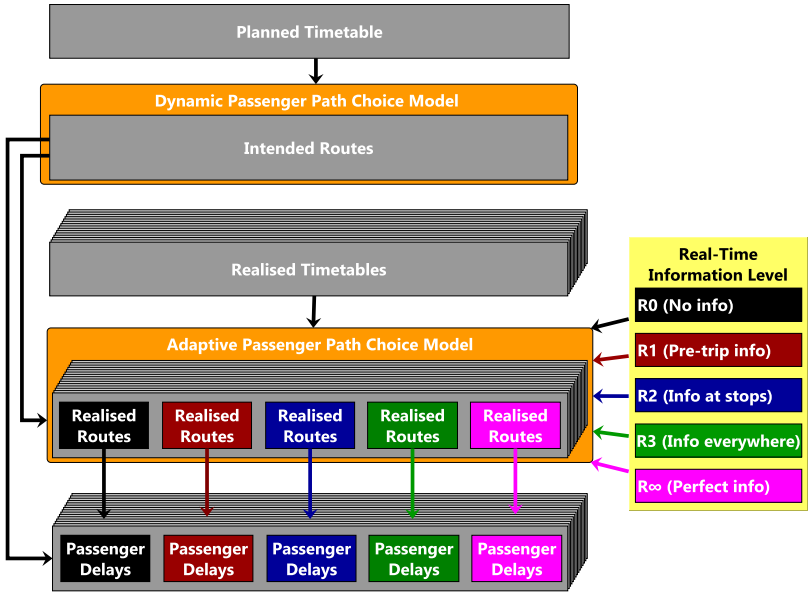


Figure 1: The overall framework. Adapted from Paulsen et al. (2018).

2. Literature Review

In this section we provide a review of relevant studies using real-time information in a schedule-based path choice context. Several studies deal with real-time information in frequency-based systems, e.g. stop-level real-time information (Gentile et al., 2005; Shimamoto et al., 2005; Trozzi et al., 2013), real-time information comprehensiveness (Nökel and Wekeck, 2009; Chen and Nie, 2015; Oliker and Bekhor, 2018), and crowding information (Fonzone and Schmöcker, 2014). However, as frequency-based models are inappropriate in low-frequency areas – which our case study will also cover – such models are not suitable for our problem at hand, and hence omitted from our review.

For studies on other matters of real-time information than its effect on passenger delays and/or path choice, we refer to the reviews by Ben-Elia and Avineri (2015); Nuzzolo et al. (2016); Brakewood and Watkins (2019).

2.1. Real-time information and passenger path choice in schedule-based systems

The concept of a "clever" passenger that actively reconsiders which route to board while waiting at a stop dates all the way back to Chriqui and Robillard (1975). The first adaptive passenger path choice model supporting such behaviour was proposed by Hall (1983) with uncertain public transport services with known distributions. Passengers could adapt en-route by utilising their (certain) time of arrival at a stop as information.

It took another ten years for real-time information about the public transport system to be considered in connection with passenger path choice and passenger travel times in Hickman (1993), who allowed passengers to receive and

adapt to real-time information of varying precision during their trip. The same methodology was applied in Hickman and Wilson (1995), but with a case study where real-time information was only relevant at the origin stop.

Nuzzolo et al. (2001) allowed passengers to obtain stop-level real-time information when deciding which service to board. The doubly stochastic model was applied to a morning peak hour of the 58 line schedule-based bus network of the Italian town of Salerno.

Landex and Nielsen (2006) calculated station-to-station passenger delays for adaptive train passengers in the suburban railway network of Copenhagen. Vehicle delays were micro-simulated, and passengers searched in a realised timetable if their delay at the station was exceeded by a predefined threshold. The methodology was later applied in Nielsen et al. (2009) using actual AVL data from the suburban railway network. Less than 10% of trips involved a transfer, and perfect information was acquired if passengers had their threshold exceeded at transfer stations. Using also real-life AVL data, Lijesen (2014) modelled station-to-station passenger delays in a subset of the Dutch railway network with departure time adaptation. Passenger delays of 16 station pairs with direct connections were analysed.

Wahba (2008) evaluated four different levels of real-time provision using MILATRAS (Wahba and Shalaby, 2009). Global pre-trip information based on the previous day, as well as various implementations of local real-time information were evaluated in a case study of the 22 routes and 284 stops network of the city of Brampton, Ontario, Canada. Although MILATRAS models real-time information internally and has since been applied in large scale, e.g. for Toronto, Canada (Wahba and Shalaby, 2011), the effect of real-time information has not been studies since.

Coppola and Rosati (2009) applied an Advanced Public Transport Information System (APTIS) with information on waiting times and vehicle bus occupancies on a nine line subnetwork of Naples, Italy. Comi et al. (2016); Nuzzolo et al. (2016) extended this to include real-time predictive information of on-board crowding and travel time applied to a single OD pair of a small network with less than 10 lines and stops.

Cats et al. (2011) proposed a dynamic agent-based station-to-station public transport simulator, BusMezzo, to evaluate the impact of different levels of comprehensiveness of real-time information for the metro system of Stockholm. Providing global information about all trains of the network at platforms was found to reduce passenger travel times compared to only doing so for trains using that particular platform or station (local information), especially when reducing frequencies. Vehicle delays were drawn from statistical distributions. The same topic was studied for similar sized Stockholm case studies using BusMezzo in Cats (2014); Cats and Jenelius (2014).

Estrada et al. (2015) simulated five levels of real-time information (none, at origin, at stops, always, and for a single line) for a 13 line network of Rivera, Uruguay. However, transfers were not allowed. Zargayouna et al. (2018) also dealt with omnipresent real-time information, but allowed transfers. The effect of having different proportions of passengers connected to such information was evaluated in a case study of Toulouse, France.

Nuzzolo and Comi (2016) showed in two cases studies of at most six lines how an individualised travel planner with real-time information improved the likelihood of a traveller choosing his/her preferred door-to-door route.

Rambha et al. (2016) proposed a Markov Decision Process supporting real-time information and applied it on four lines of the Austin transit network. Nuzzolo and Comi (2019) applied the same methodology to a single line example

from Rome, Italy whilst reducing dimensionality.

Yao et al. (2017) modelled passenger responses to real-time information for the metro system of Beijing for more than a million station-to-station trips on a 344 stations and 18 lines network with 7,539 daily train runs. The analysis did not involve comparing with a scenario without real-time information, though.

Paulsen et al. (2018) modelled door-to-door passenger delays for 1% of the population in the metropolitan area of Copenhagen using real-life AVL-data. Based on an iterative approach implemented in MATSim (Horni et al., 2016), agents could only react to information between iterations, and perfect information was assumed at some point during their trips. Leng and Corman (2020) also modelled door-to-door passenger delays using MATSim, but with pre-trip real-time information. The disruption model was applied to a large scenario of Zurich with a single disruption.

Zhu and Goverde (2019) also addressed disruptions, and proposed a station-to-station dynamic passenger assignment model for major railway disruptions incorporating passenger responses to real-time information at stations or both at stations and in trains. The model was applied for three hours to a small train network with 17 stations, and was exposed to a one hour disruption, for which the passengers knew the end time once it occured.

2.2. Research gap

The literature review has identified several research gaps that will be addressed in this study. Firstly, no studies have previously applied a multi-modal door-to-door passenger path choice model with real-time information on a large scale. The study in Yao et al. (2017) was geographically large-scale, but modelled station-to-station paths in uni-modal network. Paulsen et al. (2018); Leng and Corman (2020) modelled door-to-door passenger delays in large model areas, but only considered 1% of trips in the areas. Furthermore, only perfect information (Paulsen et al., 2018) or pre-trip information (Leng and Corman, 2020) were used.

Secondly, no studies have used real-life AVL data of both buses and trains to model passenger path choices and corresponding delays. Nielsen and Frederiksen (2009) and Lijesen (2014) used AVL data of railway networks with no or very few transfers, and Paulsen et al. (2018) did not include AVL data of buses.

Finally, none of the studies including transfers have been able to isolate the marginal effect of being able to use global real-time information to search for alternatives while walking and on-board public transport services (R3), nor being able to also adapt to global real-time information while waiting at stops (R2) as opposed to only doing so pre-trip (R1). Estrada et al. (2015) did so, but disregarded transfers. Wahba (2008); Cats and Jenelius (2014); Oliker and Bekhor (2018); Zargayouna et al. (2018) all included transfers and scenarios equivalent to R1 or R3, but the scenarios most similar to R2 were based on local real-time information, i.e. passenger could only obtain arrival time predictions concerning the stops they were currently at – in Cats and Jenelius (2014) all stops within 500m. Finally, Cats et al. (2011) studied a scenario equivalent to R2 but only compared it to local information scenarios. Thus, there still is a need to isolate the marginal effects of R2 and R3 in larger systems where transfers cannot be ignored.

3. Methodology

3.1. Adaptive passenger path choice model

3.1.1. Simulation framework

This study proposes an agent-based adaptive passenger path choice model that allows en-route decisions of its agents based on global real-time information. Figure 2 provides a graphical representation of the model, and shows how our adaptive passenger assignment model takes agents adaptively through the public transport system. For every day we are simulating we start by importing the realised timetable of that day. This is used to keep track of the actual arrival and departure times of every service in the public transport system.

We then begin our simulation at $t = t_s$, and continues in timesteps of size τ until we reach the the end-time of our simulation, t_e. In each of such timesteps we import the real-time timetable for time t, which is a timetable enriched with current delays at time t and estimations of future delays, and constructs the corresponding RAPTOR graph (Delling et al., 2015). These RAPTOR graphs are used for the shortest path searches that the agents perform during the simulation (see Section 3.1.3), but are always overruled by the fully realised timetables (actual operation) in terms of when vehicles arrives and departs at/from stops. Although this study estimate future delays very simplistically (see Section 4.1.3), the framework does allow for more advanced vehicle delay prediction methods.

In the following paragraphs the adaptive passenger path choice model will be introduced one module at the time, beginning with the activity module.

Activity module. Each agent has a current time and a current location, which for any value of the current time gives the current location of the agent. When the agent ends an activity, i.e. start a trip, the current location is set equal to the activity location, the current time is updated to be the end time of the activity, and the destination of the agent is determined. If the real-time information is at least R1 (Pre-trip info), the agent will use such real-time information to search for the shortest path to the destination, or otherwise stick to the intended route based on the planned timetable. The agent is subsequently passed on to the walking algorithm.

Walk module. Walking agents starts by finding out whether they should search for a new shortest path. They only do so, if they have not previously searched for a shortest path in the timestep and the information level is R3 (Everywhere). If so, they perform a shortest path from their current time and location to their destination. Once a shortest path has been established, assuming walks being performed with constant speeds of 1m/s in straight lines in the plane, they calculate whether they can reach their next location (activity or stop) before the next timestep. If they cannot, they walk as far as possible towards their next location, and update their current location, before setting their current time to $t + \tau$ and waiting for the next timestep. If they reach a location within the timestep, their current location is updated accordingly and the elapsed time is added to the current time. If the location matches their destination, they are passed on to the finished module, whereas the waiting module receive them if their location is now a stop/station.

Finished module. If an agent at some point manages to reach its destination its status is set to FINISHED. The realised route of the agent is exported, and the agent will no longer be processed. On the other hand, if an agent has not reached its destination when the simulation clock runs out, the trip is considered as an incomplete trip and is left out of the analysis of that particular day and real-time information level.

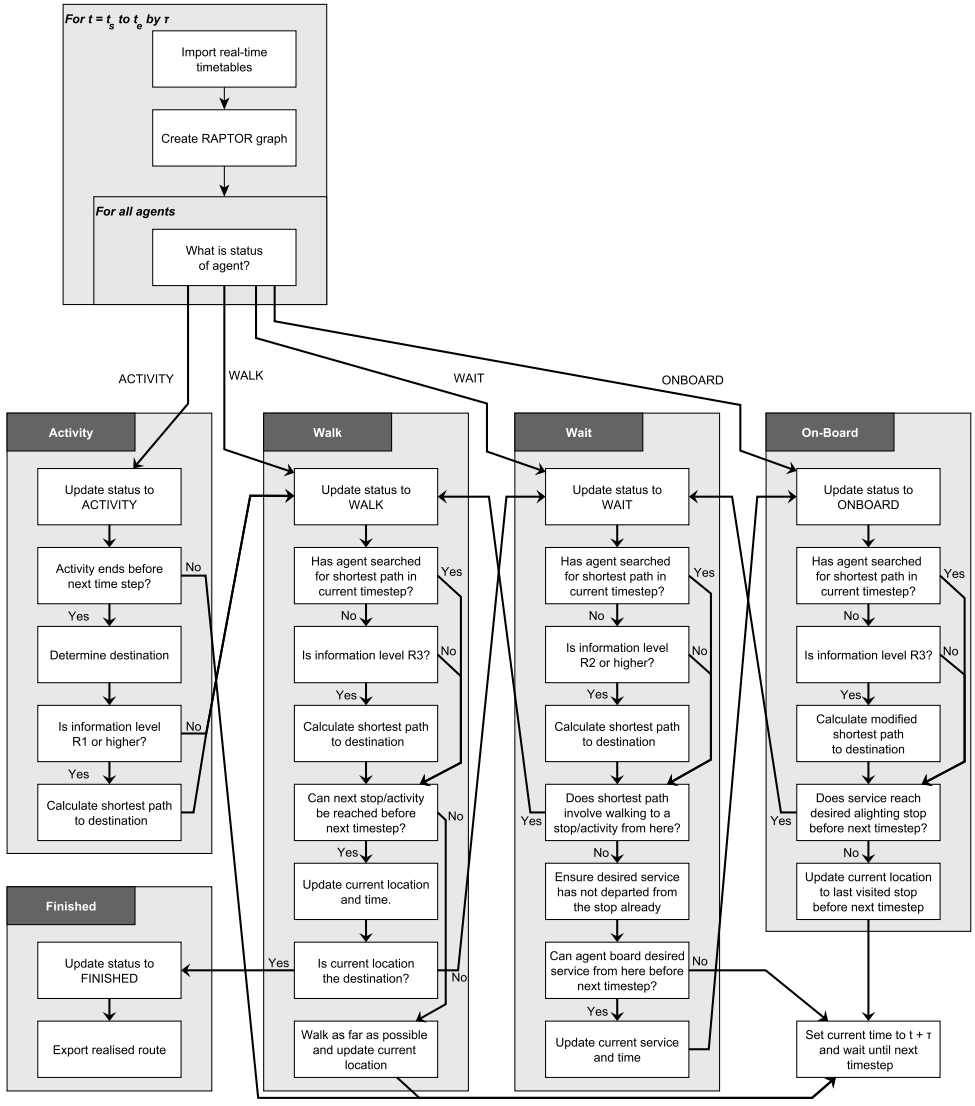


Figure 2: A graphical representation of the proposed adaptive passenger assignment algorithm.

Wait module. Waiting agents start by searching for a new path if they have not done so in the timestep and the information level is R2 (At stops) or R3 (Everywhere). It may happen that the new shortest path suggests walking to a different stop/station or all the way to the destination. If this is the case, the agent enters the walking module.

If the agent remains at the stop/station the realised timetable is consulted in order to ensure that the desired service has not already departed from the station, which may happen if the agent has arrived to the stop/station after time t (but before $t + \tau$). If it has already departed, the desired service of the agent is updated to be the next service serving the previously desired alighting station. If no such service exists, i.e. the previously desired service was the last one connecting the two stops/stations, the agent will be unable to complete the trip and will no longer be processed. In reality this would correspond to taking a taxi, calling for a ride, etc, but modelling such behaviour, which is normally more relevant in disruption models, have been deemed beyond the scope of this model.

Once the departure of the desired service is ensured to happen in the future (or now), it is determined whether it does so before the next timestep. If this is the case the agent will board the service and enter the on-board module. If it departs in a subsequent timestep, the agent waits until the next timestep.

We propose the model without capacity constraints, as denied boarding situations are extremely rare in the model area in which the model will be applied the model. If needed in other case studies, the model can be extended to comply to capacity constraints by adding passengers to a first-in-first-out queue instead of letting them board their vehicle directly. Such queues would then be processed at the end of each timestep, and agents denied boarding would then reconsider their options in the following timestep.

On-board module. The on-board module also starts by determining whether a new shortest path search is needed. A modified shortest path search will be carried out if the information level is R3 (Everywhere) and it has not previously been done in the timestep. The modified shortest path search differs slightly from the ordinary one by starting the search from the next station of the current service at the time of arrival. Furthermore, whereas a boarding penalty is normally incurred when boarding a service, no penalty is given when boarding the particular vehicle the agent is already in. Once the shortest path has been determined, the agent will alight at his desired alighting stop/station if the service reaches this stop/station before the end of the timestep. If so, the agent is moved onto the waiting module, whereas the agent will simply wait until the following timestep otherwise.

Note that when using R2 (Info at stops), an alighting passenger will at some point be passed onto the wait module, where it is possible for the agent to search for a route. However, there are still differences between R2 (At stops) and R3 (Everywhere). Firstly, in this situation, under R2 the passenger has already alighted the service, causing a boarding penalty to be incurred if the agent is to board the same train again. This means that even though the agent might follow the same route in both cases, one of them will have a higher generalised cost. Secondly, in R3 it is possible for the agent to receive a notification about an earlier alighting station being optimal. This can not happen in R2 as the search is only carried out at the desired alighting stop of the previously performed shortest path search.

3.1.2. Base and perfect information scenarios

For the base scenario where only the deterministic planned timetable is used, the algorithm would still work, but it would be horribly inefficient. Instead, a full-day RAPTOR graph (Delling et al., 2015) is created using the planned (for base) or fully realised timetable (for R∞) in which all agents search for the shortest path. In both cases the routes and travel times of these shortest path searches can be extracted directly, as no deviations would happened during the simulation.

3.1.3. Shortest path searches

The shortest path searches used in this study build on top of the RAPTOR graph (Delling et al., 2015) and shortest path search implemented for MATSim (Horni et al., 2016) presented in Rieser et al. (2018) which allows applying different utility to different submodes of the transport network, see Table 4. However, as that methodology does not ensure finding the shortest path, several extensions have been added to the model to ensure correctness as well as for speeding up computation time.

First of all, instead of only keeping the minimum cost at each node, all non-dominated alternatives are preserved at each node. With c_a denoting the cost of alternative a when departing from a node, alternative a dominates another alternative, b, if and only if it leaves the node earlier, i.e. $t_a \leq t_b$, and has a lower potential defined as,

$$\begin{aligned} c_a - t_a w &< c_b - t_b w \\ (t_b - t_a) w &< c_b - c_a, \end{aligned} \tag{1}$$

where w is the highest possible marginal disutility (in our study the marginal disutility of waiting). The potential can be seen as the worst-case time difference to cost conversion. If two alternatives leave the node at the same time, i.e. $t_a = t_b$, then only the minimal cost alternative – which is also the minimal potential alternative – is kept.

All trips have fixed coordinates for origin and destinations, but the choice of the first and final stop of the passenger path is included as part of the path choice model. We consider every stops within 3.6km of the origin as potential initial access stops in the shortest path search, but have no restrictions regarding the distance between the final egress stop and the destination. Passengers are allowed to walk directly to the destination from any stop, and are furthermore allowed to walk directly from the origin to the destination if that is favourable. From any stop in the graph agents are allowed to transfer to stops that are within 600 metres – corresponding to 10 minutes. However, in order to speed up the searches, several transfers are eliminated in the pre-processing as explained in Rieser et al. (2018). In addition to Rieser et al. (2018) we eliminate transfers that are locally dominated by transferring from the next stop of the current service or to the previous stop of the receiving service, using the aforementioned rules of dominance. Hence, we also accept travelling further than necessary in the opposite direction along the receiving line before making the transfer, which is not allowed in Rieser et al. (2018), as this in some instances can provide higher utility than transferring at the first intersection point between the two lines.

When a shortest path search is initiated, the so far optimal cost, C^*, is initialised with the value of the cost of walking directly between the origin (O) and the destination (D), denoted by $c(O, D)$. With $C(S)$ denoting the cumulative cost when reaching stop S, every time a stop is visited, $C(S) + c(S, D)$ is calculated, and C^* is updated if $C(S) + c(S, D) < C^*$.

In order to eliminate paths as early as possible, we apply a goal-directed search with pre-processed minimum travel costs between stops. These minimum but not necessarily obtainable minimum costs are found between all stops S and T in the pre-processing stage by static searches in a condensed graph as proposed in Wagner and Willhalm (2007). We denote such minimum cost by $m(S, T)$. The first time a stop, S, is visited in the shortest path search, the minimum distance to the destination, D, is determined by looping through all stops, T, within 3.6 km of the destination. With $d(T, D)$ denoting the distance between T and D, the minimum cost between S and D, denoted by $h(S, D)$, is then found as,

$$h(S, D) = \min_{T : d(T,D) \leq 3.6km} m(S, T) + c(T, D). \tag{2}$$

Paths currently at stop, S, can be eliminated if $C(S) + h(S, D) \geq C^*$.

3.2. Inverse Squared Distance Weighting

In the results section we will often have observations associated with geographical locations. In order to discover patterns in such observations, we will make use of Inverse Squared Distance Weighting (ISDW) (Shepard, 1968) for drawing surfaces. Let $k_{p,i}$ denote the rank of observation $i \in \mathcal{I}$ with respect to point p, such that i is the $k_{p,i}$'th nearest neighbour of p, let $d(p, i)$ be the distance (measured in meters) from p to i, and v_i be some value associated with observation i. Then the Inverse Squared Distance Weighting is as,

$$v_p = \begin{cases} \frac{1}{\sum_{j \in \mathcal{I}} \mathbb{1}[d(p,j)=0)]} \sum_{j \in \mathcal{I}} \mathbb{1}[d(p, j) = 0]\, v_j, & p \in \mathcal{I} \\ \frac{1}{\sum_{i \in \mathcal{I}_p} \frac{1}{d(p,i)^2}} \sum_{i \in \mathcal{I}_p} \frac{1}{d(p,i)^2} v_i, & p \notin \mathcal{I} \wedge \mathcal{I}_p \neq \emptyset\,, \\ \emptyset, & p \notin \mathcal{I} \wedge \mathcal{I}_p = \emptyset \end{cases} \tag{3}$$

with $\mathbb{1}$ denoting the indicator function, and $\mathcal{I}_p$ being the set containing the 1,000 nearest neighbours within 10km of p, i.e.

$$\mathcal{I}_p = \{\, i \in \mathcal{I} \mid d(p, i) \leq 10{,}000 \ \wedge\ k_{p,i} \leq 1{,}000 \,\}, \qquad p \notin \mathcal{I}. \tag{4}$$

Note that $\mathcal{I}_p = \emptyset$ if no observations is within 10km of p.

4. Case Study

In this section we present our case study of the metropolitan area of Copenhagen. Section 4.1 presents the data used for the case study, Section 4.2 presents the model configuration and computation times, while results are presented and analysed in Section 4.3.

4.1. Data

4.1.1. Public Transport Network

Our case study covers the entire public transport system within the metropolitan area of Copenhagen except for ferries. It contains 433 lines which are distributed across submodes according to Table 1. The public transport system consists of approximately two million daily events, i.e. occurrences of a public transport vehicle arriving at or departing from a stop in the system. The majority of the events happens in the bus system, but is supported by high-classed public transport lines which offers a large variety in options when travelling through the network.

The total number of stop facilities is 7,625 with 7,447 of them being bus stop groups, and 278 of them being railway or metro stations.

	Re/IC Trains	S-Trains	Local Trains	Metro	Bus	Total
Lines	6	7	9	2	409	433
Runs	685	1,299	777	976	28,795	32,532
Stop visits	8,286	27,847	8,734	15,128	956,542	1,016,537

Table 1: Total number of lines as well as daily runs and stop visits for each submode of the public transport network in the case study.

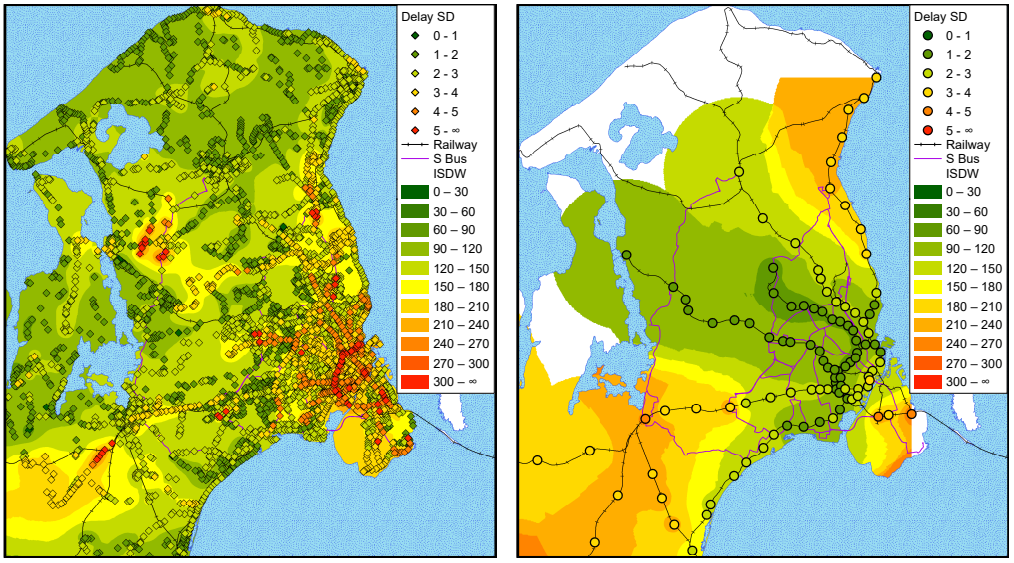
4.1.2. Automated Vehicle Location (AVL) Data

The study uses AVL data from the Danish national railway manager, Rail Net Denmark, and the regional bus company, Movia, for every weekday of September, October, and November 2014 – a total of 65 days.

The data was initially cleaned in order to secure consistency in the AVL data. The criteria used for consistency were that realised times for each service were non-decreasing from stop to stop, and that the time difference between two consecutive elements (i.e. a stop arrival or a stop departure) could not exceed two hours. For each departure, if at least one inconsistency was found, this, the previous, and the next observation were updated using linear interpolation of the delays based on their corresponding planned times. When extrapolation was needed, constant extrapolation of delays was used. This eventually caused no inconsistencies or all of the delays to be undefined. In the latter case, the observations were deemed too noisy to fix, and in lack of further information, the departure was assumed to follow the published timetable for all stops.

A total of 4,414,658 observations of a train arriving or departing at a station were deemed consistent, corresponding to approximately 67,918 observations per day, providing a coverage of almost 94%[1] on the suburban and regional train network, see Table 2.

[1] 56.60% if including local trains (8.734 daily arrivals) and metro (15,128 daily arrivals) in the calculation of coverage



(a) Buses (b) Trains

Figure 3: Map of standard deviation of vehicle arrivals at each stop/station for buses and trains. Interpolated surface estimates based on inverse squared distance weighting.

Delay data was not available for local train and metro lines, see Figure 3. Luckily, the effect of lack of data from local train and metro lines will most likely be limited as the local trains have a low demand (only 4.37% of trips use it, see Table 3), and the metro has a high frequency (down to two minutes headway in peak periods), and both of the services run on tracks with no or very few train conflicts.

For buses, delay information is spread across the entire model area but is less complete than the train data, as the bus company only collected AVL data for some of their vehicles. The vehicles capable of collecting AVL data were circulated between the different lines throughout the period. A total of 59,933,164 consistent observations were present in the dataset, corresponding to roughly 929,049 observations per day.

	Trains		Buses	
	Arrivals	Departures	Arrivals	Departures
Observations	2,207,329	2,207,329	29,966,582	29,966,582
Mean [min.]	0.18	0.31	1.98	2.13
Std. dev. [min.]	2.59	2.65	3.30	3.30
Mode [min.]	-0.67	0.5	0.70	0.70
2.5% quantile [min.]	-1.65	-1.52	-1.78	-1.57
Median [min.]	-0.28	-0.20	1.33	1.45
97.5% quantile [min.]	5.15	5.50	9.83	10.02
Coverage	93.98%		48.20%	

Table 2: Key statistical measures of delays for vehicle arrivals and departures for trains buses.

Some aggregate statistics of vehicle arrivals and departures for train and buses can be found in Table 2. It is worth noticing that the delays are generally lower for trains than for buses (which can also be seen in Figure 5). Furthermore, the standard deviation is somewhat smaller too. On Figure 3 it is clear, that the S-train lines are generally reliable, whereas the delays of the regional lines (found in the orange and yellow areas in south-western, south-eastern, and north-eastern part of the figure) are more uncertain. For buses, the delays vary most in the City of Copenhagen, but there also some small towns in the western part of the map where delays have high standard deviations.

4.1.3. Real-time timetables

Our methodology presented in Section 3.1 requires real-time timetables for every timestep. As the focus of this study is not on the real-time predictions themselves we have opted for a very simple approach. Readers are pointed towards Corman and Kecman (2018) and Petersen et al. (2019) for recent real-time prediction applications for trains and buses, respectively. The real-time timetable for time t is constructed according to Algorithm 1 for any public transport run with planned departure and arrival times $\mathbf{t}$ and recorded delays $\mathbf{d}$.

Algorithm 1 `CreateRealTimeTimetable`$(t, \mathbf{t}, \mathbf{d})$

$n \leftarrow |\mathbf{t}|$
if $t_1 + d_1 > t + t_b \ \vee \ t_n + d_n < t$ **then**
  return $\emptyset$
else if $t \leq t_1 + \min\{d_1, 0\}$ **then**
  return $\mathbf{t}$
else
  $i^* \leftarrow \min\{\, i \in \{1, ..., n\} \mid t_i + d_i \geq t \,\}$
  $d^* \leftarrow d_{i^*\text{-}1} + \dfrac{(t_{i^*} + d_{i^*}) - t}{(t_{i^*} + d_{i^*}) - (t_{i^*\text{-}1} + d_{i^*\text{-}1})} (d_{i^*} - d_{i^*\text{-}1})$
  for $j \leftarrow i^*$ **to** n **do**
    $d_j \leftarrow d^*$
  return $\mathbf{t} + \mathbf{d}$

This means that any run that has already finished is completely omitted from the timetable. For any run that was not planned to start nor has started yet, the real-time timetable is just the planned timetable. For all other runs, the estimated delay for all future stops is found by interpolating between the latest recorded delay and the next delay.

4.1.4. Demand

The demand is based on the Copenhagen Model for Person Activity Scheduling (COMPAS) (Prato et al., 2013) and consists of 801,719 trips. The usage of the different submodes of the public transport system in the basis scenario for these trips can be found in Table 3. It is seen that bus and S-train services are the most frequently used services, whereas the metro, long distance trains, and especially the local trains are less frequently used.

	Re/IC-Train	S-Train	Local Train	Metro	Bus
Use [%]	12.74	44.88	4.37	18.70	74.73

Table 3: Percentage of public transport trips using various submodes [2]

The time use of walking, waiting, and for each of the submodes as well as number of boardings are visualised in Figure 4. It is noticed that the walking time distributions is fairly wide with about a third of the trips walking for more than 20 minutes. The number of transfers are generally higher than for many previous studies on passenger delay modelling such as Landex and Nielsen (2006); Nielsen et al. (2009); Lijesen (2014), with more than 60% trips having at least one transfer (two boardings), 24% having at least two, and about 6% having more than three transfers.

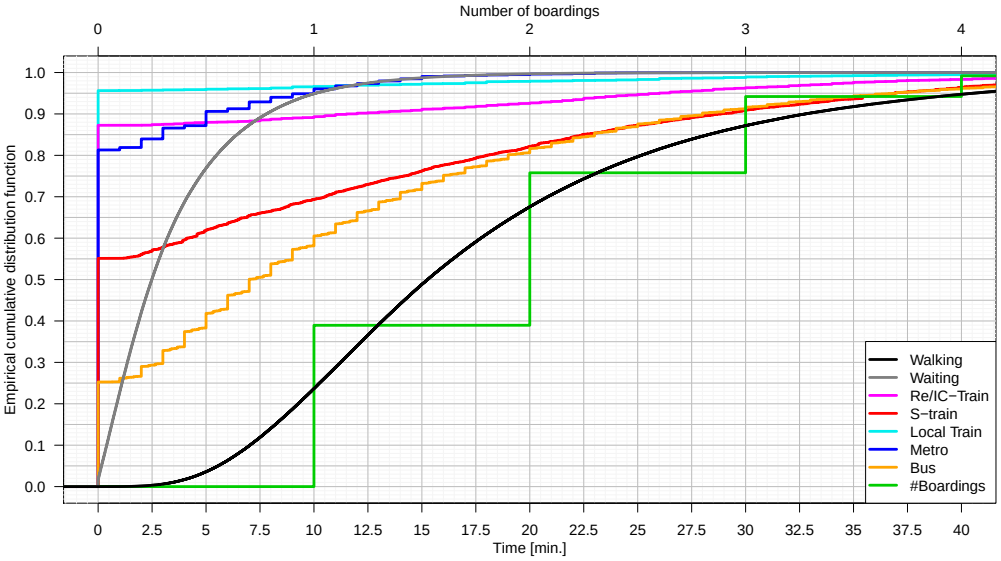


Figure 4: Empirical cumulative distribution function of number of boardings and time use in different submodes as well for waiting and walking.

4.2. *Configuration and computation times*

The model presented in Section 3.1 has been implemented in Java, and was built on top of the RAPTOR (Delling et al., 2015) algorithm developed by the Swiss Federal Railway for MATSim (Horni et al., 2016) presented in Rieser et al. (2018). We test the model using the data presented in Section 4.1 for all of the five information levels as well as a base scenario. Our simulation days ran from 3am to 3am the following day using a timestep of $\tau = 150$ seconds.

Table 4: Utilities for various modes, wait, and transfers.

	Re/IC Train	S-Train	Local Train	Metro	Bus	Wait	Walk	Boardings
Utility per minute	-1.1	-0.9	-0.9	-0.85	-1.0	-1.3	-1.6	-
Utility per event	-	-	-	-	-	-	-	-4.0

For shortest path searches the utility parameters shown in Table 4 are based on previous studies in the area, e.g. Eltved et al. (2018). When summing the generalised costs for a path, the resulting unit can be interpreted as bus

[2]Based on shortest paths in the published timetables using the parameters of Table 4.

in-vehicle minute equivalents. As this is quite verbose we will instead use the term *generalised cost units* abbreviated by *gcu*.

Although we only apply the model with a fixed parameter set in this case study, the model may relatively straightforwardly be extended to cover some taste heterogeneity, as previously pinpointed as being relevant by Nuzzolo and Comi (2016). Since the model is run in parallel on ten separate threads, each keeping the same subpopulation in every timestep and building identical raptor graphs, each thread could as well build separate graphs that corresponds to separate utility parameter sets – one for each subpopulation. As such, it would be possible to account for some heterogeneity by dividing the population into ten (or more) different categories, and running each of such categories in a separate thread, even without increasing the computation time.

Average computation times across the 65 simulation days are found in Table 5. Obviously the model is fastest when adaptivity can be disregarded completely (Base and R∞) or when paths only have to be updated if a connection is missed (R0). In both cases only a single RAPTOR graph has to be created. Computation time increases when RAPTOR graphs have to be created every τ seconds, and increases further as the number of shortest path searches increases. Nevertheless, for all levels of information the model finishes in less than five hours on average showing that the proposed methodology definitely is large-scale applicable.

	Base	R0	R1	R2	R3	R∞
Computation time [min.]	61.0	60.6	72.5	129.6	295.8	65.5

Table 5: Average computation time per day for various real-time information levels.

4.3. Results

4.3.1. Aggregate statistics

We begin the presentation of the results by comparing the similarity of trips in each of the information levels to corresponding intended trips of passengers, see Table 6. Incomplete trips are trips that have to be abandoned prematurely, because the agent can no longer find a valid path to the destination given the level of real-time information, or because the simulation time runs out. Even without real-time information this only happens for 0.89% of trips. The number of incomplete trips can be heavily reduced by providing pre-trip information, and almost entirely eliminated when applying even higher levels of real-time information.

For the no information configuration (R0) all complete trips have to use use the same stops, as the agent is not allowed to spatially deviate from the intended path. Still, 40% of such trips deviate in the temporal dimension where at least one different departure is used. This high proportion may be due to the fairly low penalties on transfers and high disutility on waiting time, which to a large degree incentivise agents to convert waiting time – that also acts as buffer time – to some sort of in-vehicle time or walking time.

When providing addtional information to agents, they are clearly seen to deviate further from their intended path, both regarding, stops, lines, and departures. In R1 more than 40% of trips choose at least one different stop, showing

that better alternative options may appear arly on. It seems as if the difference between R2 and R3 is mostly regarding the stops being used, although agents also tend to deviate slightly more regarding lines and departures.

	R0	R1	R2	R3	R∞
Complete trips	51,147,865	51,562,914	51,619,813	51,621,104	52,105,781
Incomplete trips [%]	0.93	0.13	0.02	0.01	0.01
Same stops [%]	100.00	56.04	43.43	36.31	39.59
Same lines [%]	89.48	61.87	53.84	49.94	51.13
Same departures [%]	60.18	44.70	40.24	37.05	38.66

Table 6: Similarities between intended route and modelled routes under different level of information. Incomplete trips are reported as percentage of all attempted trips, remaining figures based on complete trips only.

We now turn to look at the passenger delays, for which key figures are given in Table 7 for each of the five information levels. Key figures regarding vehicle arrivals are included in the table too for comparison.

First and foremost, it is seen that overall the passenger has a higher volatility than vehicle delays, which supports the findings of several previous studies (Landex and Nielsen, 2006; Nielsen et al., 2009; Paulsen et al., 2018). Acquiring pre-trip info reduces delays and their standard deviation considerably, even though such agents cannot spatially update their path en-route. However, it is still inferior to also applying real-time information while waiting at stops (R2) or everywhere (R3). Having the ability to apply real-time information everywhere (R3) reduces the mean delay with a negligible amount and also slightly increases the standard deviation compared to R2. Perfect information (R∞) almost reduces the average delay to zero, and also reduces the standard deviation by 50%, showing that both R2 and R3 are quite far from this theoretical lower bound.

	Vehicle arrivals		Passengers delays with real-time information level				
	Trains	Buses	R0	R1	R2	R3	R∞
N	2,207,393	29,718,431	51,147,865	51,562,914	51,619,813	51,621,104	52,105,781
Mean	0.18	1.99	10.41	5.33	3.49	3.45	0.23
Std. dev.	2.59	3.30	27.98	18.75	12.88	13.06	6.13
2.5% quantile	-1.65	-1.78	-1.49	-8.33	-9.33	-9.88	-13.96
Median	-0.28	1.33	2.75	1.12	0.87	0.91	0.16
97.5% quantile	5.13	9.85	62.80	47.89	30.44	30.26	11.14

Table 7: Key statistical measures of delays for vehicle arrivals [minutes] and passengers [gcu].

The cumulative distribution functions of passenger delays for each of our five real-time information levels, as well as the empirical distribution functions of train and bus arrivals can be found in Figure 5. It is quite clear, that without acquiring any information agents are very unlikely to save lots of time, whereas many of such trips are considerably delayed. The proportion of large delays can be reduced by acquiring pre-trip information which also facilitates decent probabilities of achieving considerable travel time savings. Savings occur even more often for R2 and R3, while also reducing the risk of large delays. Interestingly, R2 and R3 are almost indistinguishable.

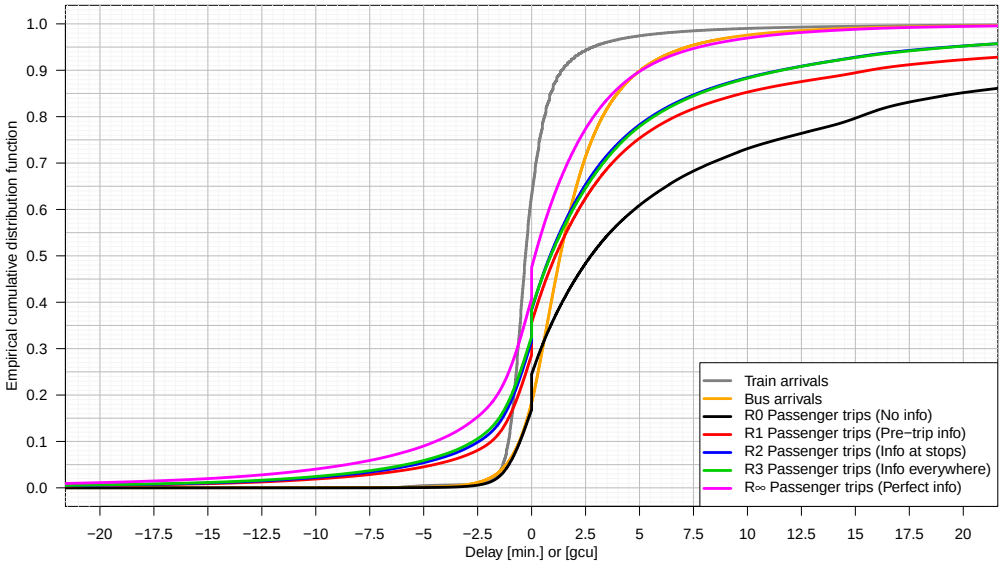


Figure 5: Empirical cumulative distribution function for bus and train arrivals [min.] and passenger delays [gcu] for the four information levels. Negative values imply having lower generalised cost than predicted by the planned timetable for passengers, and arriving early for trains/buses.

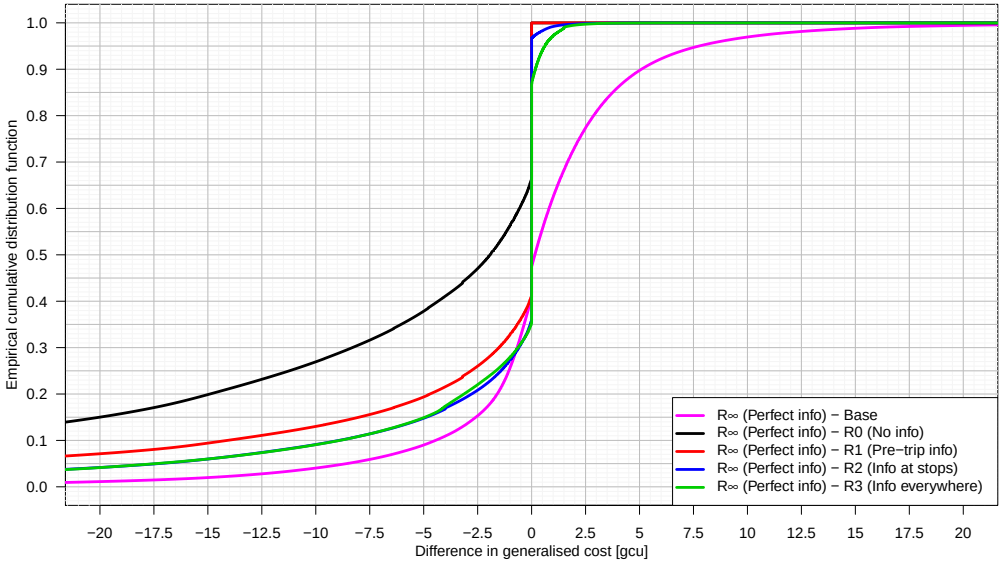


Figure 6: Cumulative distribution function for the generalised cost difference between R∞ and all other real-time information levels. Negative values imply the generalised cost of R∞ being lower.

By comparing to $R\infty$ (perfect information) instead of the intended route, see Figure 6, we can evaluate the remaining potential of the information levels. Firstly, it is seen in the upper right part of the plot that the perfect information (R∞) does not always perform better than R2 and R3. This happens because of the setup for shortest path searches, and the fact that R2 and R3 can perform successive searches that artificially increase their maximal allowed walk distance between transfers if they use an intermediate stop, Section 3.1.3. R3 also has the potential to artificially prolong walking legs by zig-zagging between optimal transfer stations, and in this way convert waiting time to walking time which has a lower disutility. R∞ always dominate R0 and R1, though.

It is seen that about a third of trips choose the optimal route even if no real-time information is available. By allowing the pre-trip information this number increases dramatically to almost 60%. Additional 5% of trips can be performed optimally (or "superoptimally") by increasing the level to R2 or R3. The remaining third of trips (those on the the left-hand side of the figure) would need real-time information with better vehicle delay predictions in order to be optimal. We return to the effect of R2 and R3 and how they differ in Section 4.3.4.

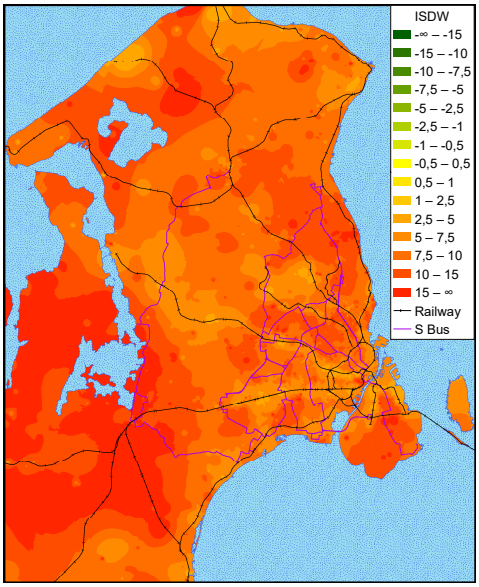
4.3.2. Geographical analysis

One of the features of the proposed door to door passenger assignment model is that we are able to analyse variation in passenger delays across origins and destinations. This can also be done in stop to stops models, however whilst ignoring the possibility of other first and last stops being optimal. In Figure 7 a inverse squared distance weighting of trip delays departing between 6am and 9am towards Central Copenhagen can be found for each of the four information levels. For each trip – which is repeated in at most 65 days – the mean of the delay of that trip has been calculated and used as value in the inverse squared distance weighting, see Section 3.2.

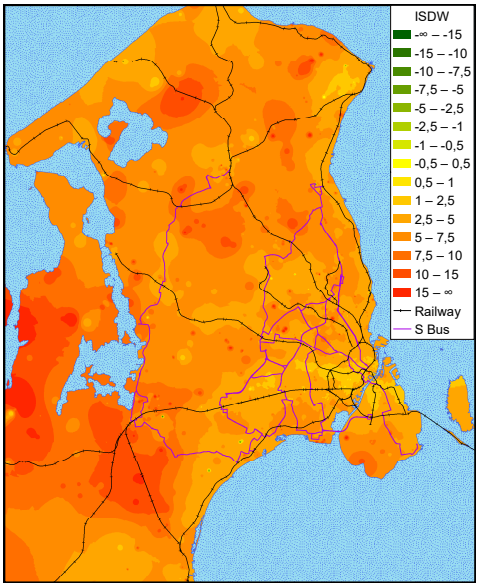
A dramatic decrease in estimated mean delay is seen when applying information when waiting or everywhere. The entire area seems to improve by receiving real-time information, even the northern part of the map that contain local lines where no AVL data was available. This shows that real-time information is not only an advantage for the urban core, but also for the less populated parts of the metropolitan area. In the Greater Copenhagen the real-time information is even more effective, causing almost all trips to have a mean delay within half a minute.

We also analyse the variability of trips by using the standard deviation of trip delays in stead of mean trip delays. The resulting maps can be found in Figure 8. We once again see that trips closer to central Copenhagen are generally better off than trips starting further away, and that they improve a lot with increased information level. As the information level increases a larger area can keep a low travel time variability. Notice, however, that the north western part of the map (areas near Hundested and Tisvilde) have their travel time variability increased. As Figure 7 shows a drop in mean passenger delay, the increase in variability stems from some trips with considerably lower travel time than what is normally the case.

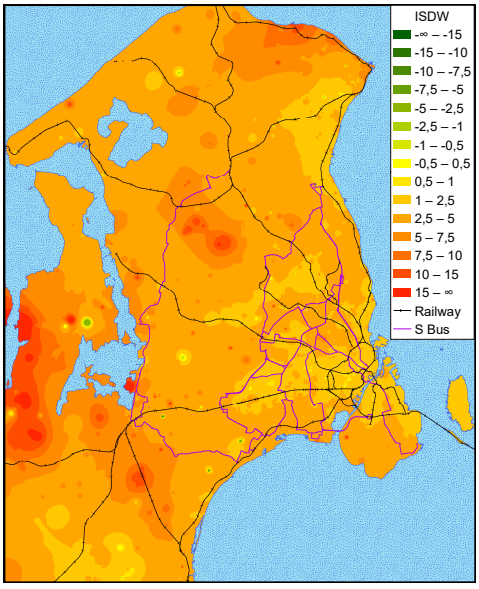
Plots equivalent to Figure 7 and 8, but for the afternoon rush hour can be found in the appendix, Figures A.1-A.2.



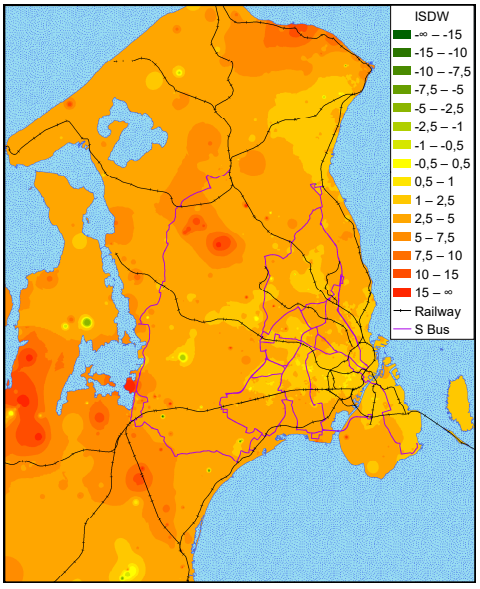
(a) R0 (No info)



(b) R1 (Pre-trip info)

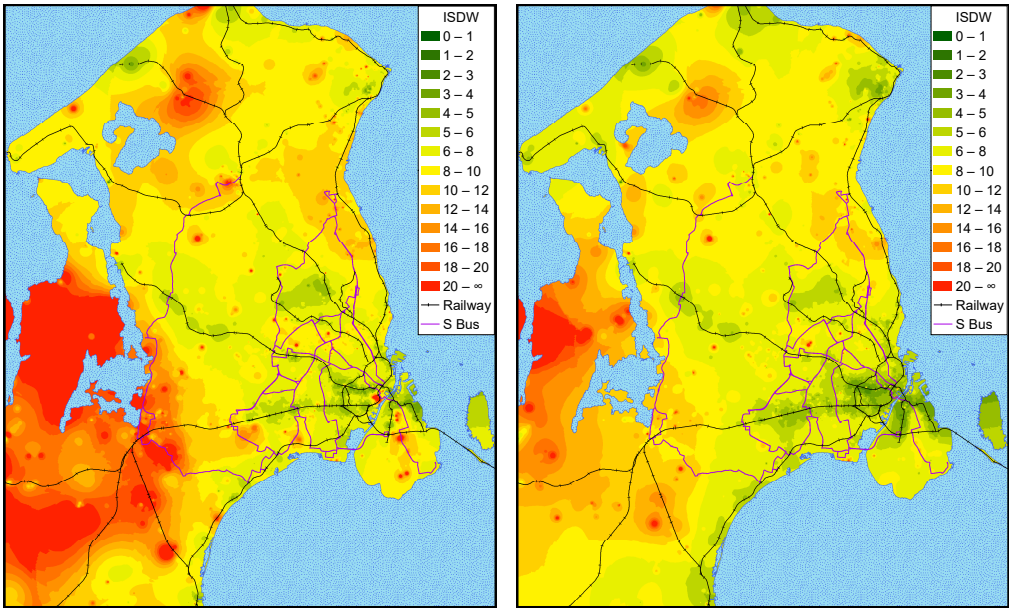


(c) R2 (Info at stops)



(d) R3 (Info everywhere)

Figure 7: Inverse squared distance weighting of the mean of delays of trips departing between 6am and 9am towards Central Copenhagen.



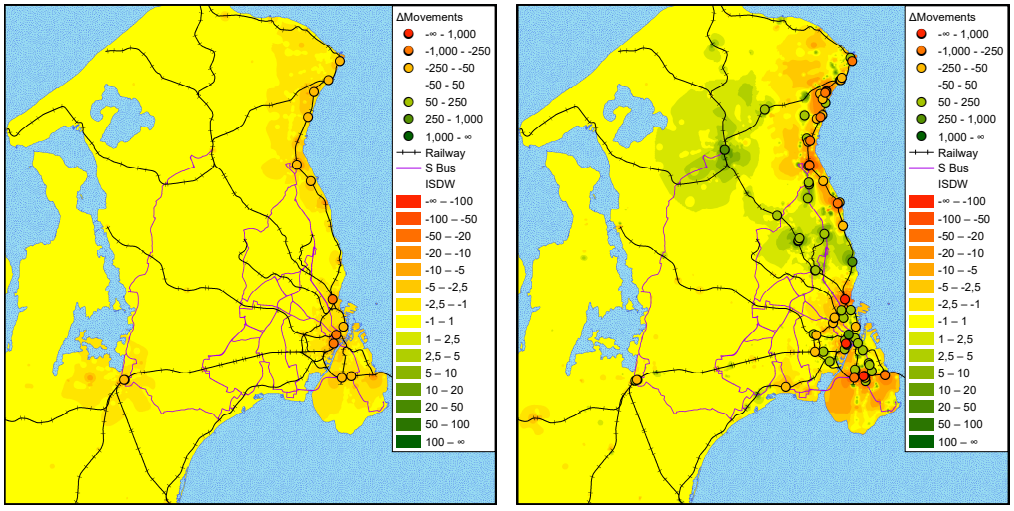
(a) R0 (No info)

(b) R1 (Pre-trip info)

(c) R2 (Info at stops)

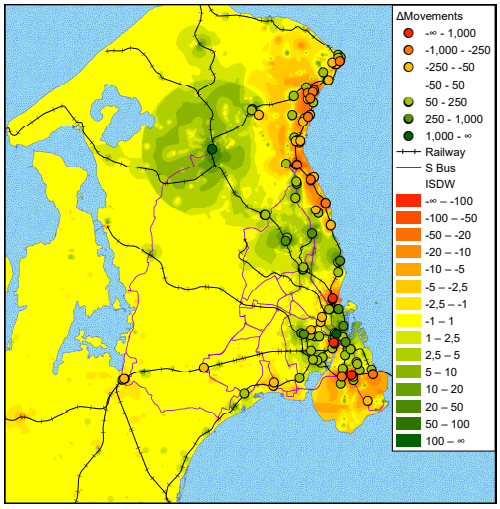
(d) R3 (Info everywhere)

Figure 8: Inverse squared distance weighting of the standard deviation of delays of trips departing between 6am and 9am towards Central Copenhagen.

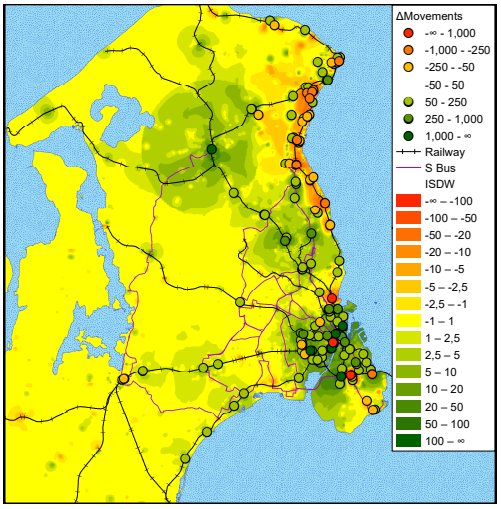


(a) R0 (No info)

(b) R1 (Pre-trip info)



(c) R2 (Info at stops)



(d) R3 (Info everywhere)

Figure 9: Difference in average daily movements (boardings and alightings) at stops w.r.t. base level. Interpolated surface estimates based on inverse squared distance weighting.

4.3.3. Passenger flow redistribution: Kystbanen

Passenger flow redistribution is often a core aspect in disruption models, e.g. Li and Xu (2011); Hong et al. (2012); De-Los-Santos et al. (2012); Rodríguez-Núñez and García-Palomares (2014); Adnan et al. (2017); Leng et al. (2018). However, lines with poor punctuality may witness a similar – but less extreme – avoidance too.

Kystbanen, running along the eastern coast of northern Zealand from Elsinore trough Copenhagen towards Sweden, had a poor punctuality in the period where AVL data was available (see Figure 3). In the following we will investigate how real-time information of such delays influence the number of passenger movements (boardings and alightings) and stations. The difference in the number of movements relative to the paths based on the planned timetable are visualised and estimated through inverse squared distance weighting in Figure 9. Only trips that intended to use Kystbanen when searching in the planned timetable are considered in this analysis. Thus, trips initially intending a different path and rerouting onto Kystbanen are not included in the analysis.

As expected, no stops gain movements when no information is offered, as all passengers stick to their intended stations - if they manage to complete their trip, that is. Some stations are seen to lose movements due to such incomplete trips. Already when offering pre-trip information, we begin to see some alternatives in the bus, metro, local railway, and suburban railway system being used more frequently, and that passengers avoid stations on the regional coast line, with some of the southern stations losing more than 1,000 movements. For R2 and R3, we see larger spatial deviation from Kystbanen, for instance with the suburban railway line from Hillerød (large dark green spot in the northern part of the map) being used very frequently, as well as local train lines and bus lines leading to it. Birkerød between Hillerød and Copenhagen shows a similar tendency. Stops served by bus 150S, which runs parallel to Kystbanen but often a few kilometres to the west, also gets an increased number of passengers. Overall, the flow redistributions seen in the maps seem behaviourally realistic.

4.3.4. Effect of en-route real-time passenger information

One of the contributions of this study is the ability to isolate the effect of acquiring global real-time information while waiting at stops (R2) as opposed to only having it available at the origin (R1), and the additional effect of being able to obtain it while walking and being onboard public transport services (R3). The remainder of this section is dedicated to an aggregate as well as disaggregate geographical analysis of the effect of such real-time information.

Figure 10 contains the the empirical cumulative distribution functions of generalised cost differences between R3 and all other scenarios. Notice, therefore that the green and magenta lines are the reversed lines of those found in Figure 5 and Figure 6, respectively. However, this plot also allows investigating in great detail the differences between R2 and R3. It is seen that R3 is in fact beaten by the less informed R2 for 15% of trips. Similar proportions of more costly trips are found when comparing to R1 and – perhaps more surprisingly – when receiving no information (R0). This indicate that the R3 information level may be a little too aggresive in its guidance of passengers.

There are at least two likely scenarios where the automatic real-time will increase the passenger delay. Assume that a passenger is on-board service A going towards a station T with a transfer to service B. Then the passenger may get a notification at a station prior to T because the real-time prediction indicates that the passenger will miss the transfer to B at T, making it more favourable to transfer to an alternative line on an earlier station instead. However, if the passenger would have been able to make the transfer to B at T, then the passenger has possibly adapted to a worse alternative. The converse can also be the case, in the sense that the passenger might be told that it is better to transfer

at a different station than T. However, there is a risk that the passenger will end up not being able to make the transfer at that station, yet again incurring a travel time loss (assuming the transfer at T was possible).

Looking on the left-hand side of Figure 10, on the other hand, it is immediately clear that the automatic real-time information can also be very beneficial. Lastly, for some trips it does not matter how much information is available. Either because there are no obvious alternatives, or because the planned shortest path remain the shortest path.

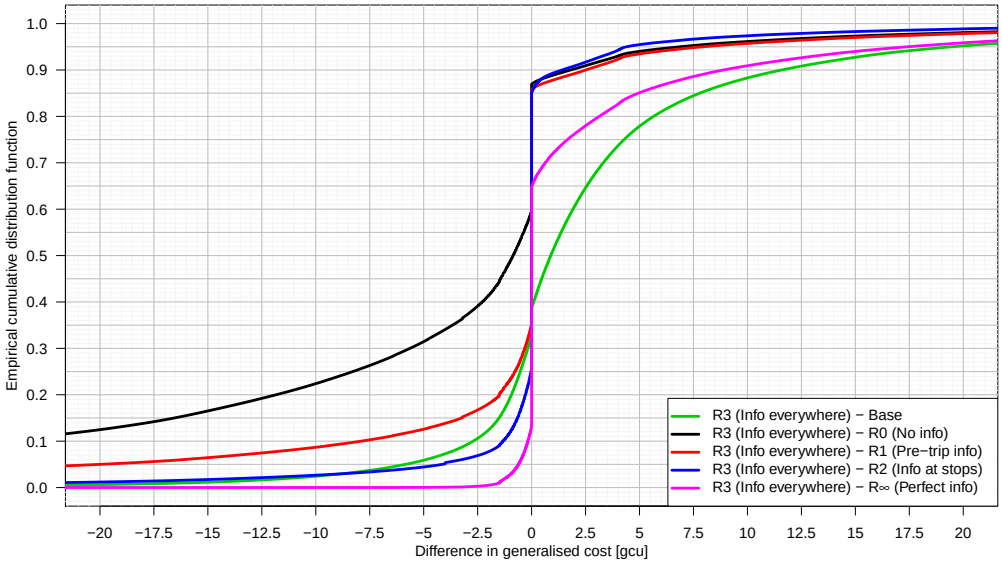


Figure 10: Distribution of differences in generalised cost when offering trip specific real-time passenger information everywhere.

Finally, we analyse the geographical distribution of locations where passengers receive notifications about better alternatives while being on-board a public transport service in R3, see Figure 11. The top 25 stations with more than 2,000 daily notifications are also listed in Table 8.

It is evident from Table 8 that the notifications are predominantly received at stations with a wide range of high classed alternatives. Surprisingly, Dybbølsbro (Dbt) – a S-train only station – has the most daily notifications. The reason for this is most likely that when arriving by S-train from the west or south, Dybbølsbro is the first available transfer station, and generally serves as a transfer station for six S-train lines.

It is also seen that train stations with a connection to regional or InterCity trains, metro, or S buses are popular locations for receiving notifications of better alternatives. This is even the case for some of the towns further away from the city center such as the ones situated along bus 400S, i.e. Hundige (Und), Ishøj (Ih), Høje Taastrup (Htå), Ballerup (Ba), and Lyngby (Ly), as well as for Ølby (Ølb) that has both a regional and a S-train line, see Figure 11.

The bus stop group with the highest number of daily notifications is located next to Copenhagen Central (Kh) and is ranked as number 42 with 584.8 daily notifications.

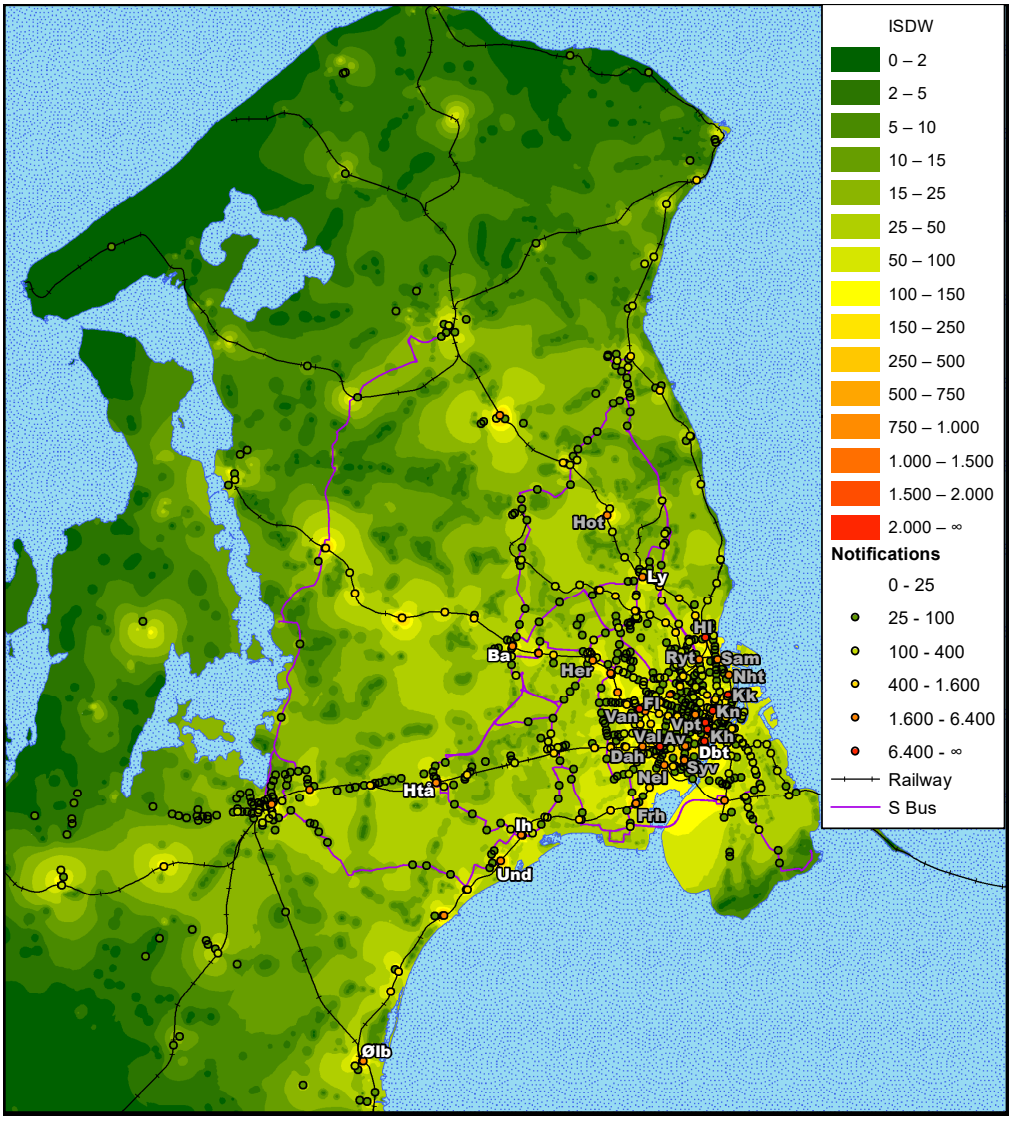


Figure 11: Geographical distribution of locations where on-board notifications of better alternatives are received in the full-adaptive setup. Interpolated surface estimates based on inverse squared distance weighting. Stations written in white are explicitly mentioned in the text.

Station	Daily notifications	High classed lines serving station			
		Re/IC Train	S Train	Metro	S Bus
Dybbølsbro (Dbt)	16,472.2	-	A,B,Bx,C,E,H	-	-
Copenhagen Central (Kh)	15,266.3	40,40a,50,54,56	A,B,Bx,C,E,H	-	250S,350S
Valby (Val)	9,963.3	40a,50,54,56	B,Bx,C,H	-	-
Nørreport (Kn)	9,076.0	40,40a,50,54,56	A,B,Bx,C,E,H	M1,M2	150S,350S
Østerport (Kk)	8,920.8	40,40a,50,54,56	A,B,Bx,C,E,H	-	-
Vesterport (Vpt)	8,075.4	-	A,B,Bx,C,E,H	-	-
Vanløse (Van)	7,148.7	-	C,H	M1,M2	-
Hellerup (Hl)	6,901.3	40,40a	A,B,Bx,C,E,H	-	-
Svanemøllen (Sam)	6,304.1	-	A,B,C,E	-	-
Høje Taastrup (Htå)	4,516.3	40a,50,54,56	B,Bx	-	400S
Flintholm (Fl)	4,248.2	-	C,H	M1,M2	-
Ny Ellebjerg (Nel)	3,257.4	50,54[3]	A,E,F	-	-
Sydhavn (Syv)	3,037.3	-	A,E	-	-
Herlev (Her)	2,979.9	-	C,H	-	300S,350S
Friheden (Frh)	2,923.2	-	A,E	-	200S
Enghave (Av)	2,721.6	-	B,Bx,C,H	-	-
Nordhavn (Nht)	2,681.0	-	A,B,C,D,E	-	-
Holte (Hot)	2,643.8	-	B,E	-	-
Ryparken (Ryt)	2,630.9	-	A,F	-	-
Lyngby (Ly)	2,465.3	-	B,E	-	200S,300S,400S
Hundige (Und)	2,325.8	-	A,E	-	400S,600S
Danshøj (Dah)	2,315.6	-	B,Bx, F	-	-
Ishøj (Ih)	2,193.7	-	A, E	-	300S,400S
Ølby (Ølb)	2,080.9	51	E	-	-
Ballerup (Ba)	2,065.9	-	C,H	-	350S,400S,500S

Table 8: An overview of the 25 stations with the highest number of daily on-board notifications and their supply of high classed public transport.

5. Conclusions and Future Work

This paper has proposed an adaptive passenger path choice model and a framework that allows large-scale evaluation of passenger delays while considering different levels of real-time information availability. The model was applied to 65 realised days considering the entire public transport system of the metropolitan area of Copenhagen and using real-life AVL data of trains and buses in the area. In each of the days the model was run with all five introduced information levels for 801,719 daily public transport trips in with average computation times ranging between one and five hours depending on the real-time information level. The computation times and size of the case study proves that the proposed model is indeed large-scale applicable.

Passenger delays were found to be considerably more volatile than the vehicles delays recorded in the area. By adapting to real-time information pre-trip passengers were shown to reduce their passenger delays noticeably, and even more so when provided with real-time information en-route. Large deviations were seen compared to the intended

[3]The regional train lines 50 and 54 only serve Ny Ellebjerg once per day per direction.

paths based on the planned timetable, with the configurations providing real-time information everywhere forcing nearly 63% of trips to differ by at least one departure. The flow redistribution was illustrated for passengers intending using a particularly unreliably line of the transport system, and the substitution patterns of passengers that explored other parts of the network were found to be realistic. A decent correlation was found between the origin and the corresponding passenger delays, with passengers beginning their trip in less serviced areas on average having larger passenger delays and travel time variability.

The effect of acquiring global real-time information and rerouting based on such information at stops (R2) was found to have a large effect compared to only being able to do so at the origin (R1). However, the net effect of being able to also be able to do so while walking and while onboard public transport services (R3) was found to be virtually zero. This was due to the fact that even though R3 outperformed the other real-time information scenarios for some trips, R3 also caused a lot of passengers to opt for alternatives that ended up performing worse than all other scenarios. This shows that R3 has potential to be beneficial, but will have to reduce the number of trips where passengers would be better off not receiving information. Developing a better strategy for the agents or improving the vehicle delay prediction methodology to reduce this number, would be an interesting direction for future research.

Another way of addressing the issue of receiving bad advice would be to adapt some day-to-day learning process to the model, as done in models like MILATRAS (Wahba and Shalaby, 2009, 2011). Previous bad experiences with a particular public transport lines or transfers could then be reflected in the utility functions, or by decreasing the trust the agent has in the travel planner.

Modelling and investigating non-obeyance of suggestions, either by having some individuals not connected as in Zargayouna et al. (2018) or by giving agents a non-zero probability of ignoring suggestions, would be interesting future research. This could also overcome some of the herd behaviour that is bound to lead to denied boardings when applying such type of model in more crowded case studies (Zargayouna et al., 2018). The model could also be adapted to directly consider effects of crowding and taste heterogeneity (as discussed in Section 3.1.1 and Section 4.2).

The computation times of the model are low enough to run in real-time or to process the passenger delays of the previous day overnight. In this way the model could be a part of an evaluation system, allowing the operators to gather insights about how yesterday's operation actually influenced the door-to-door trips of passengers. Furthermore, by using the framework with artificial delays as in e.g. Landex and Nielsen (2006); Corman and Kecman (2018), rather than historical delays, the model can be used to evaluate timetables that have not even been put into operation yet. When used in this regard, the proposed model could be a relevant tool for public transport planners.

Acknowledgements

This research was partly funded by Innovation Fund Denmark through the Integrated Public Transport Optimisation and Planning (IPTOP) project (grant 4109-00005B). We thank our IPTOP partners Rail Net Denmark and Movia for making the AVL data used in this study available as part of the project.

References

Adnan, M., Pereira, F.C., Azevedo, C.L., Basak, K., Koh, K., Loganathan, H., Peng, Z.H., Ben-Akiva, M., 2017. Evaluating Disruption Management Strategies in Rail Transit. Transportation Research Board 96th annual meeting .

Antos, J., Eichler, M.D., 2016. Tapping into Delay: Assessing Rail Transit Passenger Delay With Data from a Tap-In, Tap-Out Fare System. Transportation Research Record: Journal of the Transportation Research Board 2540, 76–83. URL: http://trrjournalonline.trb.org/doi/10.3141/2540-09, doi:10.3141/2540-09.

Ben-Elia, E., Avineri, E., 2015. Response to Travel Information: A Behavioural Review. Transport Reviews 35, 352–377. URL: http://www.tandfonline.com/doi/full/10.1080/01441647.2015.1015471, doi:10.1080/01441647.2015.1015471.

Brakewood, C., Watkins, K., 2019. A literature review of the passenger benefits of real-time transit information. Transport Reviews 39, 327–356. doi:10.1080/01441647.2018.1472147.

Carrel, A., Lau, P.S., Mishalani, R.G., Sengupta, R., Walker, J.L., 2015. Quantifying transit travel experiences from the users' perspective with high-resolution smartphone and vehicle location data: Methodologies, validation, and example analyses. Transportation Research Part C: Emerging Technologies 58, 224–239. URL: http://linkinghub.elsevier.com/retrieve/pii/S0968090X15001060, doi:10.1016/j.trc.2015.03.021.

Cats, O., 2014. An Agent-based Approach for Modeling Real-time Travel Information in Transit Systems. Procedia Computer Science 32, 744–749. URL: http://dx.doi.org/10.1016/j.procs.2014.05.485https://linkinghub.elsevier.com/retrieve/pii/S1877050914006851, doi:10.1016/j.procs.2014.05.485.

Cats, O., Jenelius, E., 2014. Dynamic Vulnerability Analysis of Public Transport Networks: Mitigation Effects of Real-Time Information. Networks and Spatial Economics 14, 435–463. URL: http://link.springer.com/10.1007/s11067-014-9237-7, doi:10.1007/s11067-014-9237-7.

Cats, O., Koutsopoulos, H.N., Burghout, W., Toledo, T., 2011. Effect of Real-Time Transit Information on Dynamic Path Choice of Passengers. Transportation Research Record: Journal of the Transportation Research Board 2217, 46–54. URL: http://journals.sagepub.com/doi/10.3141/2217-06, doi:10.3141/2217-06.

Chen, P., Nie, Y., 2015. Optimal transit routing with partial online information. Transportation Research Part B: Methodological 72, 40–58. URL: http://dx.doi.org/10.1016/j.trb.2014.11.007, doi:10.1016/j.trb.2014.11.007.

Chriqui, C., Robillard, P., 1975. Common Bus Lines. Transportation Science 9, 115–121.

Comi, A., Nuzzolo, A., Crisalli, U., Rosati, L., 2016. A New Generation of Individual Real-time Transit Information Systems, in: Nuzzolo, A., Lam, W.H.K. (Eds.), Modelling Intelligent Multi-Modal Transit Systems. CRC Press, Boca Raton, FL, USA. chapter 3, pp. 80–107. URL: http://www.crcnetbase.com/doi/10.1201/9781315368986-4, doi:10.1201/9781315368986-4.

Coppola, P., Rosati, L., 2009. Simulation-based evaluation of Advanced Public Transportation Information Systems (APTIS), in: Schedule-Based Modeling of Transportation Networks. Springer US, Boston, MA. volume 46, pp. 1–21. URL: http://link.springer.com/10.1007/978-0-387-84812-9_10, doi:10.1007/978-0-387-84812-9_10.

Corman, F., Kecman, P., 2018. Stochastic prediction of train delays in real-time using Bayesian networks. Transportation Research Part C: Emerging Technologies 95, 599–615. URL: https://linkinghub.elsevier.com/retrieve/pii/S0968090X18311021, doi:10.1016/j.trc.2018.08.003.

De-Los-Santos, A., Laporte, G., Mesa, J.A., Perea, F., 2012. Evaluating passenger robustness in a rail transit network. Transportation Research Part C: Emerging Technologies 20, 34–46. URL: https://linkinghub.elsevier.com/retrieve/pii/S0968090X10001385, doi:10.1016/j.trc.2010.09.002.

Delling, D., Pajor, T., Werneck, R.F., 2015. Round-Based Public Transit Routing. Transportation Science 49, 591–604. URL: http://pubsonline.informs.org/doi/10.1287/trsc.2014.0534, doi:10.1287/trsc.2014.0534.

Eltved, M., Nielsen, O.A., Rasmussen, T.K., 2018. The influence of frequency on route choice in mixed schedule and frequency-based public transport systems – The case of the Greater Copenhagen Area. Proceedings for the 14th Conference on Advanced Systems in Public Transport and Transit Data 2018 .

Estrada, M., Giesen, R., Mauttone, A., Nacelle, E., Segura, L., 2015. Experimental evaluation of real-time information services in transit systems from the perspective of users. Proceedings of the Conference on Advanced Systems in Public Transport (CAPST) , 1–20.

Fonzone, A., Schmöcker, J.D., 2014. Effects of transit real-time information usage strategies. Transportation Research Record 2417, 121–129. doi:10.3141/2417-13.

Gentile, G., Nguyen, S., Pallottino, S., 2005. Route Choice on Transit Networks with Online Information at Stops. Transportation Science 39, 289–297. URL: http://pubsonline.informs.org/doi/abs/10.1287/trsc.1040.0109, doi:10.1287/trsc.1040.0109.

Hall, R.W., 1983. Traveler route choice: Travel time implications of improved information and adaptive decisions. Transportation Research Part A: General 17, 201–214. doi:10.1016/0191-2607(83)90042-0.

Hickman, M.D., 1993. Assessing the impact of real-time information on transit passenger behavior. Ph.D. thesis. Massachussetts Institute of Technology. doi:10.1016/0965-8564(96)81134-7.

Hickman, M.D., Wilson, N.H., 1995. Passenger travel time and path choice implications of real-time transit information. Transportation Research Part C 3, 211–226. doi:10.1016/0968-090X(95)00007-6.

Hong, L., Gao, J., Xu, R.H., 2012. Influence of Emergency Passenger Flow Distribution in Urban Rail Network. Advanced Materials Research 450-451, 295–301. URL: https://www.scientific.net/AMR.450-451.295, doi:10.4028/www.scientific.net/AMR.450-451.295.

Horni, A., Nagel, K., Axhausen, K.W. (Eds.), 2016. The Multi-Agent Transport Simulation MATSim. Ubiquity Press, London. URL: https://www.ubiquitypress.com/site/books/10.5334/baw/, doi:10.5334/baw.

Landex, A., Nielsen, O.A., 2006. Simulation of disturbances and modelling of expected train passenger delays, in: Computers in Railways X, WIT Press, Southampton, UK. pp. 521–529. URL: http://library.witpress.com/viewpaper.asp?pcode=CR06-052-1, doi:10.2495/CR060521.

Leng, N., Corman, F., 2020. The role of information availability to passengers in public transport disruptions: An agent-based simulation approach. Transportation Research Part A: Policy and Practice 133, 214–236. URL: https://linkinghub.elsevier.com/retrieve/pii/S0965856419305075, doi:10.1016/j.tra.2020.01.007.

Leng, N., De Martinis, V., Corman, F., 2018. Agent-based simulation approach for disruption management in rail schedule. Proceedings for the 14th Conference on Advanced Systems in Public Transport and Transit Data 2018 URL: https://www.research-collection.ethz.ch:443/handle/20.500.11850/281348.

Li, F., Xu, R., 2011. Simulation of passenger distribution in urban rail transit in case of emergencies, in: Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE), IEEE. pp. 1933–1936. URL: http://ieeexplore.ieee.org/document/6199592/, doi:10.1109/TMEE.2011.6199592.

Lijesen, M.G., 2014. Optimal Traveler Responses to Stochastic Delays in Public Transport. Transportation Science 48, 256–264. URL: http://pubsonline.informs.org/doi/abs/10.1287/trsc.2013.0465, doi:10.1287/trsc.2013.0465.

Nielsen, O.A., Frederiksen, R.D., 2009. Large-scale schedule-based transit assignment - further optimization of the solution algorithms. Operations Research/ Computer Science Interfaces Series 46, 119–144. URL: http://www.scopus.com/inward/record.url?eid=2-s2.0-84867783635&partnerID=tZOtx3y1.

Nielsen, O.A., Landex, O., Frederiksen, R.D., 2009. Passenger delay models for rail networks, in: Schedule-Based Modeling of Transportation Networks. Springer US, Boston, MA. chapter 2, pp. 27–49. URL: http://link.springer.com/10.1007/978-0-387-84812-9_2, doi:10.1007/978-0-387-84812-9_2.

Nökel, K., Wekeck, S., 2009. Boarding and Alighting in Frequency-Based Transit Assignment. Transportation Research Record: Journal of the Transportation Research Board 2111, 60–67. URL: http://journals.sagepub.com/doi/10.3141/2111-08, doi:10.3141/2111-08.

Nuzzolo, A., Comi, A., 2016. Individual utility-based path suggestions in transit trip planners. IET Intelligent Transport Systems 10, 219–226. URL: https://digital-library.theiet.org/content/journals/10.1049/iet-its.2015.0138, doi:10.1049/iet-its.2015.0138.

Nuzzolo, A., Comi, A., 2019. A run-based optimal strategy search method in intelligent stochastic transit networks, in: 2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019, Institute of Electrical and Electronics Engineers Inc.. pp. 2633–2638. doi:10.1109/ITSC.2019.8917476.

Nuzzolo, A., Crisalli, U., Comi, A., Rosati, L., 2016. A mesoscopic transit assignment model including real-time predictive information on crowding. Journal of Intelligent Transportation Systems 20, 316–333. URL: https://www.tandfonline.com/doi/full/10.1080/15472450.2016.1164047, doi:10.1080/15472450.2016.1164047.

Nuzzolo, A., Russo, F., Crisalli, U., 2001. A Doubly Dynamic Schedule-based Assignment Model for Transit Networks. Transportation Science 35, 268–285. URL: http://pubsonline.informs.org/doi/abs/10.1287/trsc.35.3.268.10149, doi:10.1287/trsc.35.3.268.10149.

Oliker, N., Bekhor, S., 2018. A frequency based transit assignment model that considers online information. Transportation Research Part C: Emerging Technologies 88, 17–30. URL: https://linkinghub.elsevier.com/retrieve/pii/S0968090X18300044, doi:10.1016/j.trc.2018.01.004.

Parbo, J., Nielsen, O.A., Prato, C.G., 2016. Passenger Perspectives in Railway Timetabling: A Literature Review. Transport Reviews 36, 500–526. URL: https://www.tandfonline.com/doi/full/10.1080/01441647.2015.1113574, doi:10.1080/01441647.2015.1113574.

Paulsen, M., Rasmussen, T.K., Nielsen, O.A., 2018. Modelling Railway-Induced Passenger Delays in Multi-Modal Public Transport Networks, in: Hickman, M.D. (Ed.), Proceedings for the 14th Conference on Advanced Systems in Public Transport and Transit Data 2018, Brisbane, Australia.

Petersen, N.C., Rodrigues, F., Pereira, F.C., 2019. Multi-output bus travel time prediction with convolutional LSTM neural network. Expert Systems with Applications 120, 426–435. URL: https://linkinghub.elsevier.com/retrieve/pii/S0957417418307486, doi:10.1016/j.eswa.2018.11.028.

Prato, C.G., Rasmussen, T.K., Nielsen, O.A., Watling, D.P., 2013. A disaggregate pseudo-dynamic assignment for the activity-based model of the Greater Copenhagen Area, in: Joao, V. (Ed.), 13th World Conference on Transport Research (WCTR), Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. pp. 1–19.

Rambha, T., Boyles, S.D., Waller, S.T., 2016. Adaptive Transit Routing in Stochastic Time-Dependent Networks. Transportation Science 50, 1043–1059. URL: http://pubsonline.informs.org/doi/10.1287/trsc.2015.0613, doi:10.1287/trsc.2015.0613.

Rieser, M., Métrailler, D., Lieberherr, J., 2018. Adding Realism and Efficiency to Public Transportation in MATSim, in: 18th Swiss Transport Research Conference (STRC), Monte Verità / Ascona, Switzerland.

Rodríguez-Núñez, E., García-Palomares, J.C., 2014. Measuring the vulnerability of public transport networks. Journal of Transport Geography 35, 50–63. URL: https://linkinghub.elsevier.com/retrieve/pii/S0966692314000180, doi:10.1016/j.jtrangeo.2014.01.008.

Shepard, D., 1968. Two- dimensional interpolation function for irregularly- spaced data, in: Proceedings of the 1968 23rd Acm National Conference, ACM Press. pp. 517–524. doi:10.1145/800186.810616.

Shimamoto, H., Kurauchi, F., Yasunori, I., 2005. Evaluation on Effect of Arrival Time Information Provision Using Transit Assignment Model. International Journal of ITS Research 3, 11–18.

Sun, H., Wu, J., Wu, L., Yan, X., Gao, Z., 2016a. Estimating the influence of common disruptions on urban rail transit networks. Transportation Research Part A: Policy and Practice 94, 62–75. URL: http://linkinghub.elsevier.com/retrieve/pii/S0965856416303913, doi:10.1016/j.tra.2016.09.006.

Sun, Y., Shi, J., Schonfeld, P.M., 2016b. Identifying passenger flow characteristics and evaluating travel time reliability by visualizing AFC data: a case study of Shanghai Metro. Public Transport 8, 341–363. URL: http://link.springer.com/10.1007/s12469-016-0137-8, doi:10.1007/s12469-016-0137-8.

Trozzi, V., Gentile, G., Kaparias, I., H Bell, M.G., 2013. Route choice model and algorithm for dynamic assignment in overcrowded bus network with real-time information at bus stops. 92nd Annual Meeting of the Transportation Research Board .

Wagner, D., Willhalm, T., 2007. Speed-Up Techniques for Shortest-Path Computations, in: STACS 2007. Springer Berlin Heidelberg, Berlin, Heidelberg. volume 4393 LNCS, pp. 23–36. URL: http://link.springer.com/10.1007/978-3-540-70918-3_3, doi:10.1007/978-3-540-70918-3_3.

Wahba, M., Shalaby, A., 2009. MILATRAS: a new modeling framework for the transit assignment problem, in: Wilson, N.H., Nuzzolo, A. (Eds.), Schedule-Based Modeling of Transportation Networks. Springer US, Boston, MA, pp. 171–194. URL: http://link.springer.com/10.1007/978-0-387-84812-9_9, doi:10.1007/978-0-387-84812-9_9.

Wahba, M., Shalaby, A., 2011. Large-scale application of MILATRAS: case study of the Toronto transit network. Transportation 38, 889–908. URL: http://link.springer.com/10.1007/s11116-011-9358-5, doi:10.1007/s11116-011-9358-5.

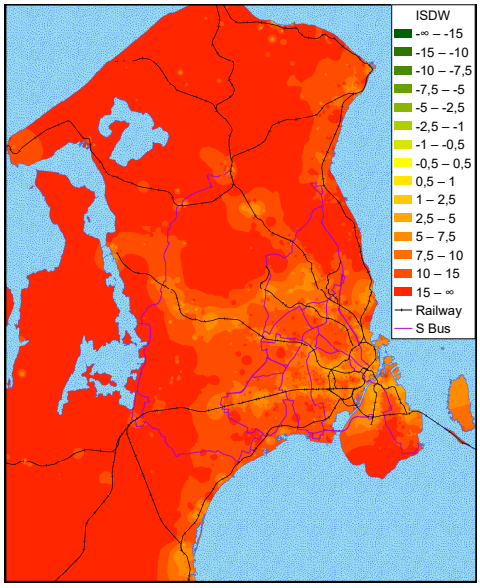
Wahba, M.M., 2008. MILATRAS: MIcrosimulation Learning-based Approach to TRansit ASsignment. Ph.D. thesis. University of Toronto, Department of Civil Engineering. URL: https://tspace.library.utoronto.ca/handle/1807/17251.

Yao, X., Han, B., Yu, D., Ren, H., 2017. Simulation-Based Dynamic Passenger Flow Assignment Modelling for a Schedule-Based Transit Network. Discrete Dynamics in Nature and Society 2017. doi:10.1155/2017/2890814.

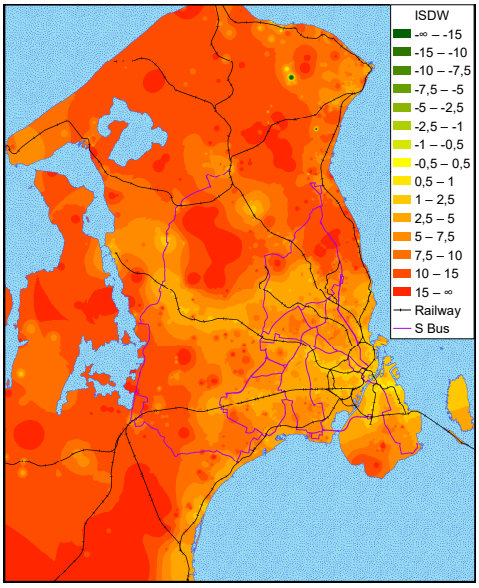
Zargayouna, M., Othman, A., Scemama, G., Zeddini, B., 2018. Multiagent Simulation of Real-Time Passenger Information on Transit Networks. IEEE Intelligent Transportation Systems Magazine 12, 50–63. URL: https://ieeexplore.ieee.org/document/8532100/, doi:10.1109/MITS.2018.2879166.

Zhu, Y., Goverde, R.M.P., 2019. Dynamic Passenger Assignment for Major Railway Disruptions Considering Information Interventions. Networks and Spatial Economics 19, 1249–1279. URL: http://link.springer.com/10.1007/s11067-019-09467-4, doi:10.1007/s11067-019-09467-4.

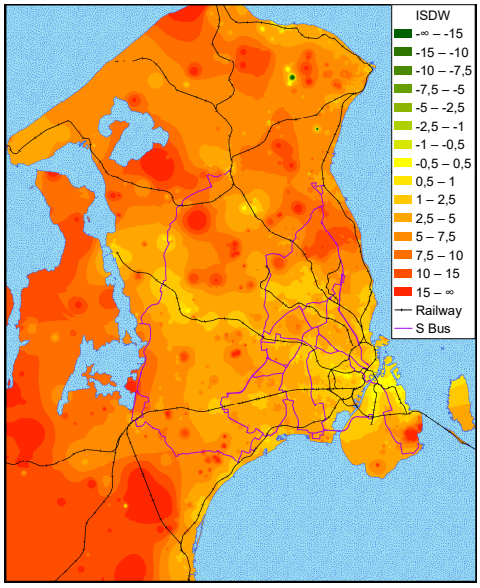
Appendix A. Supplementary Figures



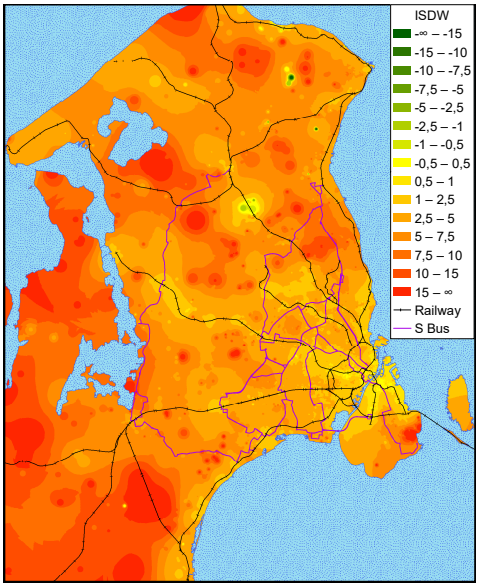
(a) R0 (No info)



(b) R1 (Pre-trip info)

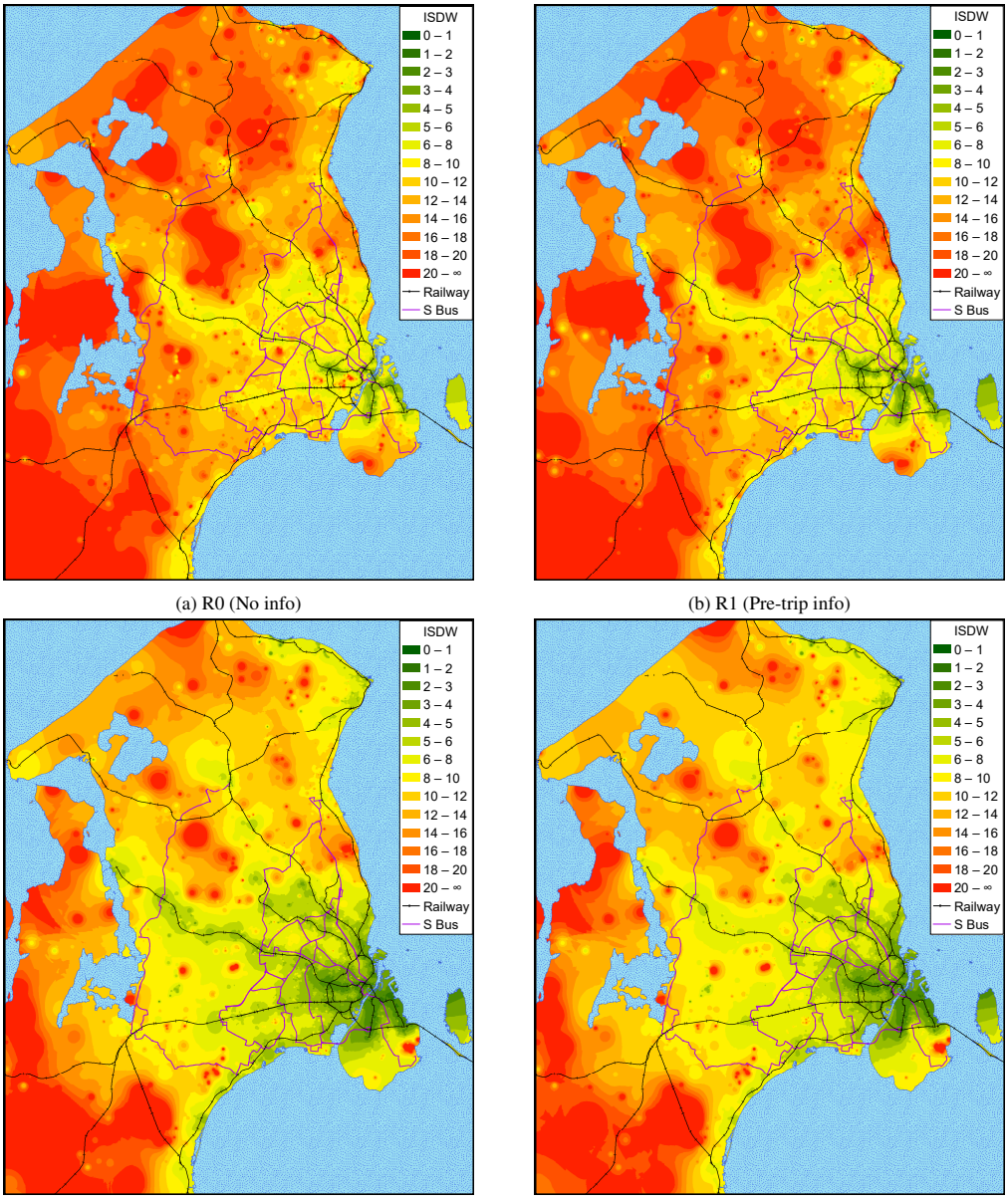


(c) R2 (Info at stops)



(d) R3 (Info everywhere)

Figure A.1: Inverse squared distance weighting of the mean of delays of trips departing between 3pm and 6pm from Central Copenhagen.



(a) R0 (No info)

(b) R1 (Pre-trip info)

(c) R2 (Info at stops)

(d) R3 (Info everywhere)

Figure A.2: Inverse squared distance weighting of the standard deviation of delays of trips departing between 3pm and 6pm from Central Copenhagen.

III

Agent-based simulation of bicycle traffic and interaction with cars

4 Paper 3: Fast or forced to follow: A speed heterogeneous approach to congested multi-lane bicycle traffic simulation

The following pages contain the article:

M. Paulsen, T. K. Rasmussen, and O. A. Nielsen (2019). "Fast or forced to follow: A speed heterogeneous approach to congested multi-lane bicycle traffic simulation". In: *Transportation Research Part B: Methodological* 127, pp. 72–98. DOI: 10.1016/j.trb.2019.07.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0191261518310336.

Please cite accordingly.

Transportation Research Part B 127 (2019) 72–98

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Fast or forced to follow: A speed heterogeneous approach to congested multi-lane bicycle traffic simulation

Mads Paulsen*, Thomas Kjær Rasmussen, Otto Anker Nielsen

Department of Technology, Management and Economics, Technical University of Denmark, Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 5 November 2018
Revised 5 July 2019
Accepted 8 July 2019

Keywords:
Bicycle traffic
Bicycle congestion modeling
Speed heterogeneity
Multi-lane traffic
Simulation
Fundamental diagram

ABSTRACT

Copenhagen is world-known for its large proportion of cyclists, forming a diverse group with a large variation of equipment and physical abilities. This leads to a considerable speed heterogeneity which needs to be taken into account when modelling the traffic on dedicated bicycle paths. Nevertheless, existing studies on bicycle traffic simulation have either neglected such speed heterogeneity altogether or modelled it by dividing cyclists into a few discrete classes ignoring the entirety of the speed distribution. This paper proposes an efficient bicycle traffic simulation model with continuously speed heterogeneous cyclists and corresponding congestion effects. Based on individual-specific desired speeds and headway distance preferences, the model shows realistic speed-flow relationships validated with on-site observations while being capable of delaying rapid cyclists more often than slower ones in moderate traffic flows. The scalability of the model allows it to be large-scale applicable for network loading purposes, and thus suitable for evaluating impacts of cycling related infrastructure investments.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

The bicycle is a very space-efficient mode of transport, why major cities across the world are starting to promote bicycling as a means of relieving congestion in urban areas. However, in places where bicycles already hold a high market share of travellers – such as Copenhagen and the Netherlands – the number of bicycles have reached a level where congestion on dedicated bicycle paths can no longer be ignored.

When modelling car traffic it is customary to consider all cars as a homogeneous mass where all cars traversing a link follow the same speed within a certain period of time. This assumption is more or less valid for car traffic as car drivers generally desire to go by the maximum allowed speed, alternatively by the highest speed possible given the congestion level. This is, however, not the case for cyclists for which higher speeds require a considerable increase in the wattage produced by the cyclist, limiting less physically fit cyclists from reaching high speeds. Furthermore, differences in equipment such as cargo bikes or racing bikes can significantly reduce or increase the desired speed of a cyclist.

Another difference is that overtaking is predominantly unproblematic outside rush hours for cyclists in Copenhagen as all bicycle paths are wide enough to fit at least two bicycles. However, due to the high frequency of overtakings, it may occur that the overtaking lane is occupied by a relatively slow cyclist overtaking an even slower cyclist. This makes it possible for

* Corresponding author.
E-mail address: madsp@dtu.dk (M. Paulsen).

https://doi.org/10.1016/j.trb.2019.07.002
0191-2615/© 2019 Elsevier Ltd. All rights reserved.

other cyclists to cause delays to fellow cyclists even at moderate traffic intensities. Including such bicycle interactions within a fast bicycle simulation model is the primary aim of this paper.

The current literature on explicit modelling of bicycle traffic can be divided into two main categories: Heterogeneous and homogeneous modelling.

The literature on modelling of heterogeneous bicycle traffic has had a strong dominance of Cellular Automata (CA) (Wolfram, 1986) based approaches. Whereas original CA applications on traffic simulation such as Nagel and Schreckenberg (1992) only allowed simulation of homogeneous traffic, newer multi-value CA approaches have allowed simultaneous simulation of inhomogeneous traffic through discrete classes with different maximum speeds. Initially proposed for mixed traffic of cars and motorcycles in Lan and Chang (2005), it was later applied to mixed bicycle traffic by Jia et al. (2007). Since then a whole range of studies have considered mixed bicycle traffic (Gould and Karner, 2009; Zhao et al., 2013; Jin et al., 2015a; Shan et al., 2015; Zhou et al., 2015; Li and Fang, 2017; Xue et al., 2017) or mixed traffic with bicycles alongside other modes (Vasic and Ruskin, 2012; Luo et al., 2015) using similar and extended CA approaches.

Although the CA approaches are capable of modelling different classes, the discrete nature of the methodology hinders including the entirety of the speed distribution. The CA methodology partitions links into cells with a minimum size requiring fully containing any of the simulated vehicles – cyclists in our case. Since a cyclist have to stay in the same cell or move an integer number of cells per time step, different cyclist types either have to be assumed equal or differing rather considerably in order to be simulated using such approach. Thus, CA is only a suitable approach for modelling heterogeneity when dealing with traffic that can be divided into a few clearly distinguishable classes where intra-class heterogeneity can be ignored completely. This inherently restricts the potential of such models for the purpose of our study.

Three methodologies for simulation of continuous heterogeneous bicycle traffic exist in the literature. Shen et al. (2011) proposed a cell-stream flow model, Liang et al. (2012, 2018) developed psychological-physiological force models, whereas Osowski and Waterson (2017) applied a social force model. These approaches are capable of realistically simulating heterogeneous bicycle traffic in great detail but are, however, computationally infeasible for full-scale applications as explicitly stated in Osowski and Waterson (2017). Although none of the other studies touch upon computation time, the detail level of these models strongly indicates that none of them would have the computational efficiency to be suitable for traffic assignment purposes.

As a side note, heterogeneous traffic has also been modelled without focusing (solely) on bicycle traffic. Multi-class heterogeneous highway traffic was modelled in Wong and Wong (2002) using an extended Lighthill–Whitham–Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956), whereas Tang et al. (2011) applied a macroscopic following model with honk effects to a mix of bicycle and pedestrian traffic.

The second main category of bicycle traffic modelling is homogeneous simulation of bicycle traffic. This has also been explored with CA approaches (Jiang et al., 2004; Liu et al., 2008; Jiang et al., 2017; Tang et al., 2018), but has more interestingly also received some attention with alternative methodologies. Zhang et al. (2013b) combined CA with gas dynamics models for high densities to model speed-density relationships for bicycle traffic. Zhao and Zhang (2017) introduced a robust unified follow-the-leader model framework capable of modelling homogeneous traffic of either pedestrians, bicycles, or cars. Although these studies generally benefit from more advanced continuous methods, they have the drawback of completely ignoring the inherent speed heterogeneity of cyclists.

Regarding data, some studies (Mai et al., 2013; Zhang et al., 2013a; Rui et al., 2014) have been analysing single-file controlled experiments to explain cyclists' behaviours. Additionally, some of the aforementioned bicycle models have been calibrated/validated by them (Liang et al., 2012; Xue et al., 2017; Zhao and Zhang, 2017), potentially leading to undesired biases as such experiments arguably differ from everyday bicycle traffic. First of all because overtaking is not considered in such controlled experiments. Secondly, because the cyclists have to ride in a circular or oval curve. Thirdly, since the cyclists of the controlled experiments are aware that they are participating in an experiment, and as such they might unconsciously behave differently from what they normally would in an uncontrolled environment. Collectively, these biases lead to experiments that are usually carried out using much lower speeds than the real-life counterparts they are seeking to model.

Hence, although calibrating a model using controlled experimental data is superior to no calibration, it still has some drawbacks compared to observed, uncontrolled data. Unfortunately, high quality data of observed, uncontrolled bicycle traffic is limited, although a few studies (Gould and Karner, 2009; Zhang et al., 2013b; Jin et al., 2015a; Li and Fang, 2017; Liang et al., 2018) have been able to collect such for calibrating their models.

Whereas forming fundamental diagrams require reliant aggregate data of very high traffic intensities, some characteristics of bicycling may, on the other hand, be derived by analysing bicycle traffic under low traffic intensities. Establishing a consistent bicycle simulation model based on assumptions that can be analysed at low traffic intensities is one of the aims of this study.

The study contributes to the literature by modelling explicit desired speeds for each cyclist drawn from a continuous distribution. Bicycle-to-bicycle interactions and overtaking are modelled implicitly using individualised speed-dependent headways for each cyclist and an intelligent lane choice mechanism. Furthermore, the implemented model is fast enough to be large-scale applicable, and thus has the potential to simulate the entirety of bicycle traffic within a metropolitan area.

The remainder of the paper is structured as follows. Bicycle traffic characteristics including speed and headway distance heterogeneity are analysed based on observed data in Section 2. Section 3 formulates the underlying assumptions for a bicycle simulation model capable of modelling inter-vehicular interactions of heterogeneous cyclists. Section 4 contains the

Fig. 1. An overview of the data area at Smallegade in Frederiksberg, Capital Region, Denmark.

design and results of small-scale experiments focusing on speed heterogeneity as well as aggregate fundamental diagrams accompanied by sensitivity analyses and data validation. Finally, a discussion of model limitations as well as directions for future work are found alongside the final conclusions in Section 5.

2. Characteristics of bicycle traffic

This section describes some key characteristics of bicycle traffic important for the later implementation of the bicycle traffic simulation model. The analysis is based on aerial video data from COWI Ltd.[1] processed using DataFromSky (RCE Systems R.S.O., 2014). The video was shot from the tower of Frederiksberg Town Hall situated in the Copenhagen enclave Frederiksberg, with the original purpose of evaluating the use of the large parking lot next to the town hall. Still, the video data includes all east-going (towards the camera) bicycle traffic of the street Smallegade from 2 p.m. to 5 p.m. with a total of 1168 cyclists. Their speed and acceleration for every cyclists have been calculated along 18 transversal gates each divided into a inner right (from the perspective of the cyclists) and an outer left part, see Fig. 1.

The average flow along this stretch is less than 390 cyclists per hour why the traffic intensity is not high enough for congestion effects to be relevant. This is on purpose though, as one of the key parameters that will be estimated in the following sub-sections (desired speed) would not be possible to estimate using congested data.

2.1. Speed heterogeneity

For each cyclist in the aerial video data the 90th percentile speed was used for determining their desired speed. Fig. 2 shows the distribution of the desired speeds across cyclists, alongside best-fit estimations of 11 candidate distributions. The mean value and median of the empirical distribution are 6.29 m/s and 6.21 m/s, respectively. Furthermore, the distribution has a variance of 1.23, an excess kurtosis of 1.26, and a skewness equal to 0.55, the latter two indicating that a normal disitribution may not be suitable.

The data reveals relatively high speeds compared to many other studies seen in the literature from other parts of the world. Allen et al. (1998) reviews several older studies from across the world, and shows that the majority of these lie between 12 km/h (3.33 m/s) and 20 km/h (5.5 m/s). Jensen et al. (2010) analysed bicycle speeds of shared bicycles in Lyon, France, and showed an average speed of 14.5 km/h (4.03 m/s). Bernardi and Rupi (2015) measured the average speed on three segregated bicycle lanes in Bologna, Italy, to be ranging from 14.6–18.9 km/h (4.1–5.3 m/s). Li and Fang (2017) showed an average speed of 2.67 m/s for cyclists at the Nanhu campus of Wuhan University of Technology, China. Finally, Flügel et al. (2017) found the average cycling speed in Oslo, Norway to be 16.85 km/h (4.68 m/s).

However, Buch and Greibe (2015) did a study in Copenhagen, Denmark, and found an average speed of 21.7 km/h (6.02 m/s) which is reasonably close to the data obtained for this study.

In order to obtain a continuous distribution with known probability density function, we proceed by searching for a theoretical distribution that matches the empirical distribution well. This is done by assessing 11 continuous candidate distributions for which best fit estimates have been found for every parameter in their parameter set using maximum likelihood

[1] https://www.cowi.com .

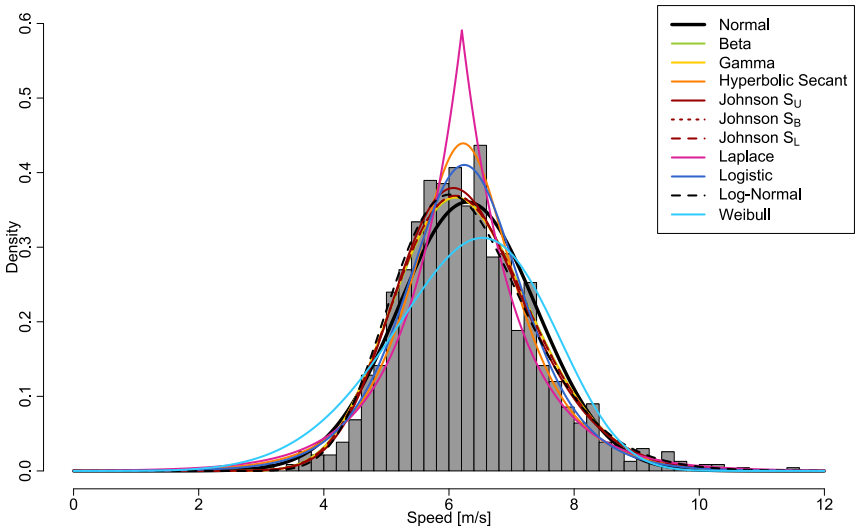


Fig. 2. Distribution of observed desired speeds and best-fit estimations of candidate distributions.

Table 1
Log-likelihoods (ℓ) as well as test statistics (D) and p-values for the Kolmogorov–Smirnov test for various candidate distributions.

Distribution	ℓ	D	p
Johnson's S_U	−1750.9	0.022	0.606
Johnson's S_L	−1756.8	0.029	0.286
Johnson's S_B	−1756.9	0.029	0.279
Logistic	−1757.7	0.028	0.310
Beta	−1758.4	0.031	0.200
Gamma	−1758.4	0.030	0.254
Hyperbolic Secant	−1760.9	0.032	0.181
Log-Normal	−1761.7	0.030	0.235
Normal	−1775.6	0.052	0.004
Laplace	−1782.5	0.042	0.032
Weibull	−1851.0	0.077	< 0.001

estimation. Table 1 contains the 11 distributions alongside their log-likelihood values (ℓ) and results from the Kolmogorov–Smirnov test (D and p).

Based on the log-likelihood, the Johnson's S_U distribution (Johnson, 1949) performs best, having a considerably higher log-likelihood than the following seven distributions which, on the other hand, do no differ much from each other. The final three distributions (Normal, Laplace, and Weibull) are clearly worse than the top eight distributions based on log-likelihood.

The Kolmogorov–Smirnov test calculates the largest vertical deviation (D) between the empirical and theoretical cumulative distribution function. Combined with the number of observations this number can be used to directly determine the p-value, i.e. the probability of the largest vertical deviation exceeding D, assuming that the empirical distribution was actually drawn from the theoretical distribution under consideration.

Results from the test shows that eight of the theoretical distributions do not deviate significantly from the empirical distribution when using a significance level of 5 %. The Johnson's S_U distribution gives the best results, but a simpler distribution such as the logistic distribution would also be suitable according to the test. It is worth emphasising that based on the Kolmogorov–Smirnov test the normal distribution – as one of only two distributions – can be ruled out as being an appropriate choice for any significance level higher than 0.4%.

Conclusively, as the Johnson's S_U both has the highest log-likelihood and performs the best in the Kolmogorov–Smirnov test, the analysis unanimously shows that it is the best suited distribution among the 11 candidate distributions.

2.2. Lane behaviour

The test street is a quite moderate Copenhagen bicycle path with a width of 1.85 m and a relatively low traffic intensity in the data collection period as stated earlier. We want to assess whether cyclists in the overtaking (left) lane on average

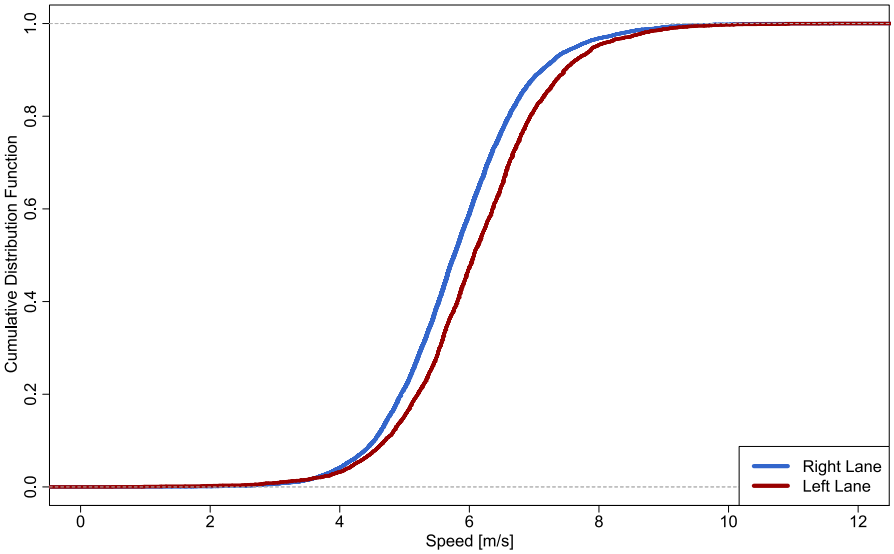


Fig. 3. Speed distribution of the right and left lane, respectively.

go faster than those in the inner (right) lane, despite knowing that this difference will be more evident on segments with higher traffic intensities. To do this the aerial data has been divided into cyclists riding on the right and left half of the bicycle path, respectively, and the speeds of the cyclists have been grouped by the lane at which the cyclists rode. The empirical cumulative distribution function of the speeds in each lane can be found in Fig. 3. It is seen that the two empirical distributions differ, and that the left lane seems to be the faster lane.

Whether the difference in mean speed across the two lanes is statistically significant has been tested by performing a two-sample z-test for the difference between means. The resulting z-statistic exceeds 15, meaning that the difference in mean speed across lanes is strongly significant using any reasonable significance level with the corresponding p-value being less than 10^{-50}.

This is not to say that lane formation behaviour of cyclists only occur when overtaking. Group behaviour where cyclists ride alongside each other – as implemented in Tang et al. (2018) – is also likely to occur. However, such abreast riding can arguably be assumed not to influence other cyclists as long as the grouping behaviour stops once a faster cyclist rings his/her bell.

2.3. Headways

The capacities of bicycle paths are highly dependent on how close cyclists are willing to ride to each other: the closer they ride, the higher the capacity. Because of this, determination of bicycle headways has already received some attention in the literature.

According to Andresen et al. (2014) – based on single-file track experiments – the headway distance, $d(v)$, is a function of the speed, v, having the functional form,

$$d(v) = \theta_0 + \theta_1 v, \tag{1}$$

with parameters $\theta_0 = 1.93$ being the sum of the average length of a bicycle ($\lambda^c = 1.73\,\text{m}$) and an additional safety distance of $\lambda^s = 0.2\,\text{m}$, and $\theta_1 = 0.78\,\text{s}$ being the assumed constant headway time between the front wheel of the following cyclist and the back wheel of the cyclist in front.

Hoogendoorn and Daamen (2016) investigated headways on a bicycle path in Delft in the Netherlands from photo-finish data. They use the approach of Wasielewski (1979) and Hoogendoorn (2005) and found that 54.1% of bicycles are constrained by the cyclist in front of them, and that the average headway time of these cyclists is 0.784 s with a standard deviation of 0.660 s.

In our video data we have supplementary information, as we can not only measure the headway distance but also the speed of every cyclist. This is helpful when identifying whether a cyclist is driving freely or being constrained, since large speed differences would indicate the former.

We filter the data such that we only consider data from cyclists that reasonably can be assumed to be following the bicycle in front of them. Since it is difficult to keep the exact same speed as the cyclist in front, in this analysis we allow

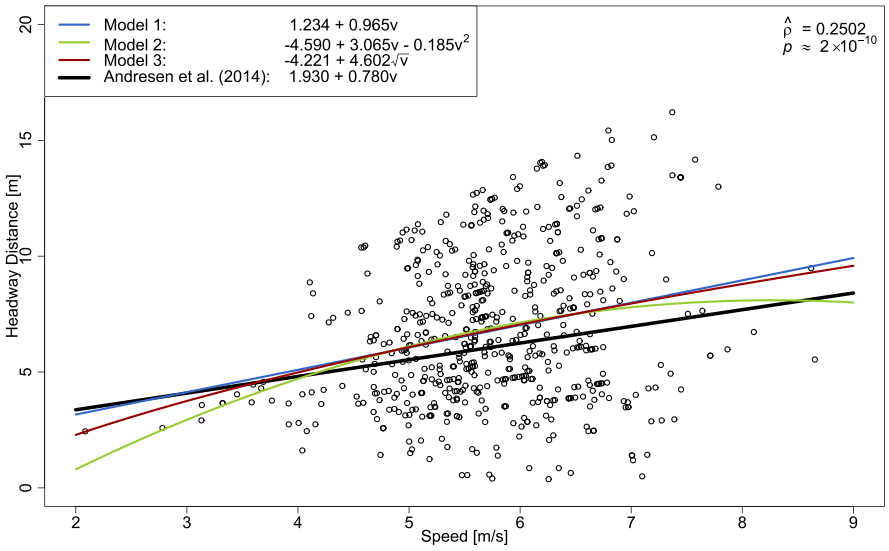


Fig. 4. Headway distances plotted against speed with three candidate trend lines and a line based on Andresen et al. (2014).

Table 2
Three linear regression models for headway distances based on speed with *t*-values in brackets compared to Andresen et al. (2014).

	Model 1	Model 2	Model 3	Andresen et al. (2014)
Intercept	1.234	−4.590	−4.221	1.930
	(1.398)	*(−1.367)*	*(−2.473)*	*(–)*
$\sqrt{\text{Speed}}$	–	–	4.602	–
	(–)	*(–)*	*(6.468)*	*(–)*
Speed	0.965	3.065	–	0.780
	(6.355)	*(2.603)*	*(–)*	*(–)*
Speed^2	–	−0.185	–	–
	(–)	*(−1.798)*	*(–)*	*(–)*
σ_ϵ	3.114	3.108	3.110	–
R^2	0.063	0.068	0.065	–

for an absolute speed difference threshold of 0.2 m/s above which cyclists are assumed to ride freely. This corresponds to riding freely if the difference in distance per second is larger than the desired longitudinal distance between two stationary bicycles (0.2 m) found in Andresen et al. (2014). Mohammed et al. (2019) identified a cluster characterised as following cyclists revealing an average speed difference of 0.93 m/s in the raw data and 0.76 m/s after bootstrapping, indicating that the value of 0.2 m/s indeed is quite strict.

Furthermore, as in Hoogendoorn and Daamen (2016), all observations with a headway time above a certain threshold are assumed to be freely moving why such observations are also discarded. In Hoogendoorn and Daamen (2016) the method incentivised not choosing a (too) low number, whereas doing such in this case would bias the analysis. A threshold of 2.3 s – as opposed to 3 s in Hoogendoorn and Daamen (2016) and 5 s in Mohammed et al. (2019) – has been used based on a flattening trend of the distribution of all headways in our dataset.

Finally, since observations of the same cyclist pair (at different gates) are highly correlated, only the median headway distance of all accepted instances of a cyclist pair is used. The final dataset consists of 607 headway distance observations from unique cyclist pairs. Due to the conservative thresholds of both speed differences and headway times we are rather confident that all non-following situations have been removed by the filtration.

Besides using the speeds to filter the headway observations, it can also be used to investigate the effect the speed has on the headway distance. Fig. 4 shows a scatter plot where the headway distance is plotted against the speed of the behind cyclist. Headway distances are found by multiplying the headway time with the speed of the cyclist.

In addition we fit a first and a second order polynomial as well as a square root model, see Table 2. Besides the parameter estimates and *t*-statistics, the table also contains the residual standard error (σ_ϵ) and coefficient of determination (R^2) for the proposed models. For comparison, the table also contains results found in Andresen et al. (2014).

The model obtained by using a first order polynomial has a steeper slope (headway time of 0.965 s) than in Andresen et al. (2014). However, the *t*-statistic of the null hypothesis that the true slope is equal to 0.780 s is 1.220, resulting in a two-sided *p*-value of 0.223 indicating that it is reasonably likely that the values from Andresen et al. (2014) could be true.

Intuitively, the headway distance should have a negative curvature as faster cyclists most likely are more experienced cyclists and as such would keep a lower headway time. In fact, this is partially supported by the second order polynomial fit which shows a negative second order term with a *t*-statistic of -1.80 corresponding to a one-sided *p*-value of 0.04. The square root model only provides a slight increase in R^2 compared to the first order polynomial from 0.063 to 0.068.

As the second order term in the second order polynomial is barely significant on a 5 % significance level, and the slope of the first order polynomial is not significantly different from the slope of Andresen et al. (2014), the choice is really between using the square root model or the originally proposed model by Andresen et al. (2014). Although relatively weak, there is some evidence towards the headway distance having a negative curvature, why the square root model seems to be the most appropriate headway distance function. It also has a slightly higher R^2 than the first order polynomial despite having the same number of parameters.

2.3.1. Headway heterogeneity

In order to account for the large heterogeneity in headway distances, the model is extended further to not only fit the mean value associated with a given a speed, but also the corresponding standard deviation. As the variation of the headway distances seems to be increasing for higher speeds, see Fig. 4, the standard deviation has to be modelled explicitly using a speed-dependent function. In a first model, letting $\mathcal{N}(\mu, \sigma^2)$ denote the normal distribution with mean μ and variance σ^2, we maintain normally distributed residuals and model the headway distances in the following way,

$$d(v) \sim \mathcal{N}\left(\theta_0 + \theta_1\sqrt{v},\ \left(\zeta_0 + \zeta_1\sqrt{v}\right)^2\right) = \theta_0 + \theta_1\sqrt{v} + \left(\zeta_0 + \zeta_1\sqrt{v}\right)\mathcal{N}\left(0, 1^2\right). \tag{2}$$

With this formulation the mean values are kept constant, whereas the standard deviation of the residuals is a first order polynomial of $\sqrt{v}$ parameterised by ζ_0 and ζ_1. This model, however, has no bounds on the headways as the normal distribution has infinite tails, which may provide unrealistic headway preferences when sampling from the distribution. In order to deal with this, two additional models based on beta distributed residuals are proposed.

The generalised beta distribution, which we denote by $\mathcal{B}_g(\alpha, \beta, a, b)$, is a regular beta distribution, $\mathcal{B}(\alpha, \beta)$, that is translated and scaled to have compact support $[a, b]$. This means that for a stochastic variable $X \sim \mathcal{B}_g(\alpha, \beta, a, b)$ the standardised variable $Z_{\mathcal{B}} = \frac{X-a}{b-a} \sim \mathcal{B}(\alpha, \beta)$. Using the notation $f(x|\alpha, \beta)$ for the probability density function of $\mathcal{B}(\alpha, \beta)$ and $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$, the probability density function of the generalised beta distribution, $\mathcal{B}_g(\alpha, \beta, a, b)$, is given by,

$$f_g(x|\alpha, \beta, a, b) = \begin{cases} \frac{1}{b-a} \times f\left(\frac{x-a}{b-a}\middle|\alpha, \beta\right) = \frac{1}{b-a}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{x-a}{b-a}\right)^{\alpha-1}\left(1-\frac{x-a}{b-a}\right)^{\beta-1}, & x \in [a, b] \\ 0, & x \notin [a, b]. \end{cases} \tag{3}$$

It is customary to require the expected value of the residuals to be 0. The easiest way to secure this, is to require the distribution to be symmetrical around 0. This can be achieved by requiring $\alpha = \beta, \alpha > 0$, and $-a = b = h, h > 0$. Since we still want to allow larger residuals for higher expected values of $d(v)$, h is parametrised by $h = \zeta_0 + \zeta_1\sqrt{v}$. We will denote such generalised symmetric beta distribution by $\mathcal{B}_s(\alpha, h)$. Its probability density function, f_s, is given by,

$$f_s(x|\alpha, h) = f_g(x|\alpha, \alpha, -h, h) = \frac{1}{2h} \times f\left(\frac{x+h}{2h}\middle|\alpha, \alpha\right), \quad \forall x \in \mathbb{R}, \tag{4}$$

Having introduced the above, three additional models focusing on modelling the residuals can be presented. Model 4 that uses the formulation from Eq. (2) as well as two models using Beta-distributed residuals; A symmetric Model 5 and a potentially asymmetric Model 6 defined by Eqs. (5) and (6), respectively,

$$d(v) \sim \theta_0 + \theta_1\sqrt{v} + \mathcal{B}_s(\alpha, h) = \theta_0 + \theta_1\sqrt{v} + 2h\left(\mathcal{B}(\alpha, \alpha) - \frac{1}{2}\right), \tag{5}$$

$$d(v) \sim \theta_0 + \theta_1\sqrt{v} + \mathcal{B}_g(\alpha, \beta, -h, h) = \theta_0 + \theta_1\sqrt{v} + 2h\left(\mathcal{B}(\alpha, \beta) - \frac{1}{2}\right). \tag{6}$$

For Model 4, 5, and 6 the parameters are estimated simultaneously using the L-BFGS-B algorithm (Byrd et al., 1995) in R (R Core Team, 2018). In order to secure positive standard deviations for the entire range of speeds in the headway data, $\sqrt{2}\theta_1 \geq -\theta_0$ and $\sqrt{2}\zeta_1 \geq -\zeta_0$ as well as negative values for θ_0 and ζ_0 were applied as constraints for the algorithm. Furthermore, Model 5 and 6 required $\alpha \geq 1$, and Model 6 additionally constrained $\beta \geq 1$.

For all of the models the parameters, their log-likelihood (ℓ), and Akaike Information Criterion (AIC) (Akaike, 1973) are presented in Table 3. Notice that ζ_0 of Model 3 is the previously implicitly estimated σ_ϵ from Table 2.

Firstly, it is seen that regarding the parameters concerned with the expected value of the headway distance, i.e. θ_0 and θ_1, the changes are almost indistinguishable when going from Model 3 to Model 4. When going onto Model 5 and Model 6 the changes persist to stay relatively minor.

Looking at the log-likelihood, it is seen that there is a considerable improvement when switching from a constant (Model 3) to a parameterised (Model 4) standard deviation. However, as such switch will always have a non-negative effect, it is

Table 3
Parameters, log-likelihoods (ℓ), and Akaike Information Criteria (AIC) of four models for headway distances using different functional forms for residuals.

	Model 3	Model 4	Model 5	Model 6
θ_0	−4.221	−4.234	−4.357	−4.539
θ_1	4.602	4.602	4.713	4.840
ζ_0	3.110	−4.397	−9.674	−0.841
ζ_1	–	3.109	6.841	6.959
α	–	–	1.865	1.910
β	–	–	–	2.033
ℓ	−1549.04	−1522.88	−1497.87	−1497.45
AIC	3104.08	3053.77	3005.74	3006.90

necessary to use some kind of measure to assess whether the improvement caused by the inclusion of the ζ_1-parameter is worthwhile. For nested models, as it is the case here, applying Wilks' theorem (Wilks, 1938), commonly known as a likelihood ratio test, is a frequently used method to assess improvements in likelihood. Letting Λ denote the likelihood ratio between Model 3 and Model 4, the test statistic can be calculated as follows,

$$D_\Lambda = -2\ln\Lambda = -2\times(-1549.04+1522.88) = 52.32.$$

Comparing that to a χ^2-distribution with one degree of freedom yields a p-value $< 10^{-12}$, showing that the improvement in likelihood is significant at any reasonable significance level.

When comparing Model 5 to Model 4, it is no longer possible to apply Wilks' theorem as neither model is a special instance of the other model. In such cases the Akaike Information Criterion (AIC) (Akaike, 1973) can be used instead. With k denoting the number of estimated parameters the AIC is defined as,

$$\text{AIC} = 2k - 2\ell, \tag{7}$$

and thus becomes higher when either the log-likelihood (ℓ) decreases or when the number of parameter increases. The AIC thus discourages overfitting, and the model with the lower AIC is considered the better option. We can easily see that Model 5 has a much lower AIC (3005.74) than Model 4 (3053.77), why Model 5 is superior to Model 4 based on AIC. The improvement is very similar to that when going from Model 3 to Model 4.

Comparing Model 5 to Model 6 we can see that without constraining β it is estimated to be a little larger than α, although the differences are rather small. We can once again apply Wilks' theorem, as Model 5 is a particular instance of Model 6 where $\beta = \alpha$. The test statistic $D_\Lambda = 0.835$ yields a p-value of 0.361, meaning that the improvement in likelihood is only significant when using a significance level above 36.1%. That Model 5 seems to be the better model is also supported by the AIC which is lower for Model 5. As such, there seems to be no statistical evidence against assuming $\alpha = \beta$ when modelling our residuals with a beta distribution with support $[-h, h]$, $h > 0$.

However, since all of the above model assessments are purely relative measures, it is sensible to test Model 5 against a baseline model. Only by doing so is it possible to evaluate whether Model 5 is actually a reasonable model, or just the best among a candidate set of models. A null model only containing the average value of headways, $\theta_0 = 6.788$ and a constant residual standard deviation $\zeta_0 = 3.211$ of the assumed normally distributed residual yields an AIC of 4211.66. The improvement in AIC of 1205.92 when going to Model 5 strongly suggests that adding the parameters used in Model 5 makes statistical sense.

Thus, Model 5 is chosen as the best suitable model for modelling headway distances; It has the best statistical measures, but also because it makes sense from a behavioural point of view to use one of the two bounded models (5 or 6), as using an unbounded model (3 or 4) theoretically allows extreme headway distances. This is particularly an issue when sampling from the distribution. Whereas the actual sampling procedure shall be explained in Section 3.3, for the sake of the argument Fig. 5 is presented here to show how the unbounded models 3 and 4 differs fundamentally from model 5 and 6.

The figure shows a surface plot illustrating the probability of obtaining a value that is further away from the expected value in the same direction as the value on the y-axis given some speed, v. In a more formal explanation, let D_v be the stochastic variable of the headway distance based on the speed v according to one of our proposed models, and let F_D be the theoretical cumulative distribution function of said stochastic variable. We then introduce the following measure, p, indicating the probability of getting a value that is larger/smaller than d given v,

$$p(d|v) = \min\{F_D(d|v), 1-F_D(d|v)\} = \frac{1}{2}\times P\Big(D_v > |d - E(D_v)| \;\Big|\; v\Big), \quad \forall d\in\mathbb{R}, \forall v \geq 2. \tag{8}$$

It is readily seen that Model 5 and 6 are almost indistinguishable, and that they are bounded very close to the empirical observations. This also means that sampling from such distribution will almost always provide reasonable values, although very small negative values can occur. Model 3 and 4, on the other hand, have substantial risks of sampling quite extreme negative as well as positive values. A probability of e.g. 10^{-5} might seem very small, but since the model is meant to be large-scale applicable and able to simulate several hundred thousands of cyclists, such probabilities will eventually come

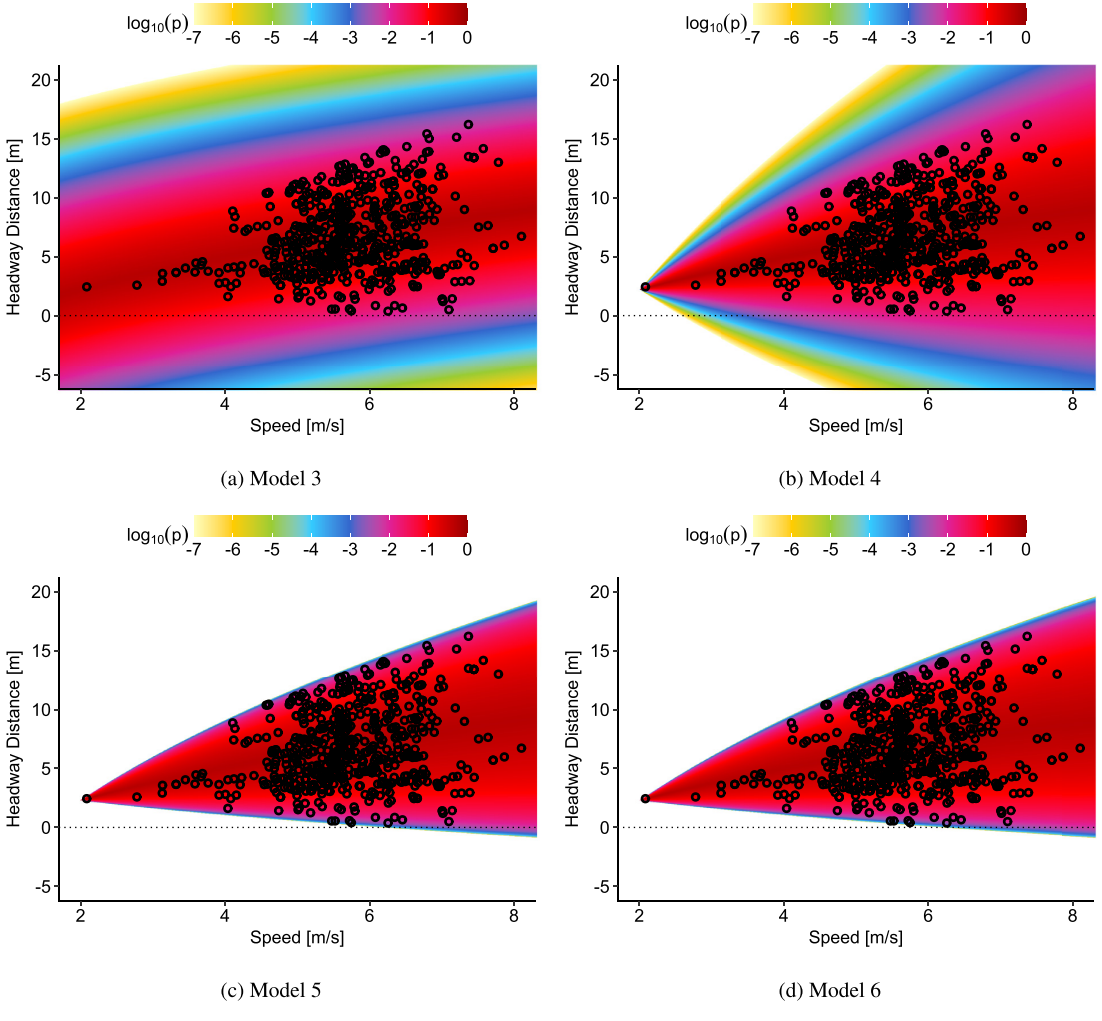


Fig. 5. Surface plots showing the probability of sampling more extreme values (see Eq. (8)) on a log-scale for Model 3 through 6.

into play. Finally, it can be seen that Model 3 does not take the increasing variance into account at all, and as such seems way off when sampling headway distances at low speeds.

Using Model 5 and the space mean speed $\bar{v}_0 = 6.104$, the expected headway time is $\frac{d(\bar{v}_0)-\lambda^c}{\bar{v}_0} = 0.910$ s. This seems reasonable although slightly above the headway times of 0.78 and 0.784 s found in Andresen et al. (2014) and Hoogendoorn and Daamen (2016), respectively.

3. Model

Whereas Section 2 was devoted to extracting certain characteristics of bicycle data from field data, this section aims at establishing a simulation framework for heterogeneous bicycle traffic into which such findings can be used as input. The proposed framework is based on the set of assumptions found below.

1. Every cyclist has an individual desired speed.
2. Bicycle paths can be divided into a number of pseudo-lanes.
3. Cyclists keep a headway distance to the cyclist in front of them.
4. Cyclists use the right-most pseudo-lane that can accommodate their desired speed.
   a. If no such pseudo-lane exists, they use the fastest possible pseudo-lane.
5. Cyclists stay in the same pseudo-lane for the entirety of the link.
6. Cyclists cannot enter a link if its area is fully occupied.

The remainder of this section is devoted to discussing each of the assumptions individually, before Section 3.7 summarises the input needed for the framework.

3.1. Desired speeds

The first assumption is that every cyclist has an individual desired speed. The desired speed can be interpreted as the speed a cyclist would ride by under free-flow conditions. In that sense it can also be understood as a personal maximum speed for that individual.

In accordance with findings in Section 2.1, we assume that cyclists' desired speeds follow a Johnson's S_U distribution which by denoting the inverse hyperbolic sine function by $\operatorname{arsinh}(x) = \ln(x + \sqrt{x^2+1})$ is defined as,

$$v_0^c \sim \mathcal{D} = \frac{\delta}{\lambda\sqrt{2\pi}\sqrt{1+(\frac{v_0^c-\xi}{\lambda})^2}} e^{-\frac{1}{2}(\gamma+\delta \operatorname{arsinh}(\frac{v_0^c-\xi}{\lambda}))^2}, \quad v_0^c \geq 2, \tag{9}$$

with parameters $\gamma = -2.75$, $\xi = 3.67$, $\delta = 4.07$, and $\lambda = 3.49$. Notice that the superscript c is used because any desired speed v_0^c is associated with and specific to one particular cyclist, c.

A uniform value for each cyclist, $u_c \in U(0, 1)$, can be transformed into a desired speed by the following modification of the quantile from the standard normal distribution $q_{u_c}, u_c \in]0, 1[$ readily available in most programming software,

$$v_0^c = \lambda \sinh\left(\frac{q_{u_c} - \gamma}{\delta}\right) + \xi, \quad u_c \in]0, 1[, \forall c \in C, \tag{10}$$

making the sampling of such values straightforward.

It is seen from Eq. (9) that we truncate the distribution at $v_0^{\min} = 2$ m/s in order to avoid desired speed that are too low (Johnson's S_U distribution has a natural lower bound at 0). In practice such values are rejected and replaced by a resampled value. The truncation is necessary because there exists a speed below which it is virtually impossible to keep the balance on a bicycle, but also because having too low values can almost shut down links entirely with the proposed model (see Section 3.4). It is worth noting that none of the estimations of the (fully) bounded distributions (Johnson's S_B and Beta) resulted in a non-negative lower bound, meaning that the value used for truncation had to be determined manually. The value of 2 m/s seems reasonable based on our data, as no cyclists were observed riding below this speed with the minimum observation being 2.67 m/s.

In the proposed model the desired speed is constant across links for each individual. However, the framework can easily be extended to have link specific desired speeds that for instance take the gradient, surface, or wind into account.

3.2. Infrastructure

The second assumption requires that every link of bicycle path, l, has an array of pseudo-lanes, Ψ^l. The number of pseudo-lanes, $|\Psi^l|$, contained in this array is based on the width of the bicycle path ω^l. The elements of Ψ^l are enumerated from right to left, i.e. from ψ_1^l to $\psi_{|\Psi^l|}^l$. In our model such pseudo-lanes are strictly obeyed in the sense that no interaction is assumed across pseudo-lanes.

Any function $f : \mathbb{R}^+ \to \mathbb{N}^+$ can be used for partitioning the link into lanes. We propose using a formula from an existing study (Buch and Greibe, 2015) who found the number of effective lanes – corresponding to $|\Psi^l|$ – on bicycle paths in

Copenhagen to be,

$$\left|\Psi^l\right| = 1 + \left\lfloor \frac{\omega^l - 0.4\,\mathrm{m}}{1.25\,\mathrm{m}} \right\rfloor. \tag{11}$$

The formula states that a new effective lane is obtained for every 1.25 m of width exceeding 0.4 m. Again this can easily be extended from a constant formula to also include link characteristics such as the gradient, wind, or surface type.

In the current implementation, however, the only additional attribute of a link of bicycle path is its length, λ^l. This is in contrast to traditional models for car traffic where for instance the capacity of the link has to be defined explicitly.

3.3. Headway distances

The third assumption assures that cyclists cannot overlap each other as they move through traffic. In fact, they even have to keep a minimum headway distance to the cyclist in front of them. The headway distance implemented in the model is based on the findings of Section 2.3, specifically Eq. (5).

Assuming that the residuals of the distance model arise from inter-personal preferences regarding the predefined headway distance, it makes sense to have the same cyclists consistently having a relatively short or long headway distance. This can be done by rewriting Eq. (5) by using the beta-distributed variable $Z_{\mathcal{B}} \sim \mathcal{B}(\alpha, \alpha)$,

$$d(v) \sim \theta_0 + \theta_1\sqrt{v} + 2\left(Z_{\mathcal{B}} - \frac{1}{2}\right) \times (\zeta_0 + \zeta_1\sqrt{v}), \quad v \geq 2. \tag{12}$$

Individual values of $z^c_{\mathcal{B}} \sim Z_{\mathcal{B}}$ can then be drawn for each cyclist $c \in C$ to find the cyclist-specific distance functions,

$$d^c(v) = \theta^c_0 + \theta^c_1\sqrt{v}, \quad v \geq 2, \tag{13}$$

where $\theta^c_0 = \theta_0 + 2\zeta_0(z^c_{\mathcal{B}} - \frac{1}{2})$ and $\theta^c_1 = \theta_1 + 2\zeta_1(z^c_{\mathcal{B}} - \frac{1}{2})$ are the individual-specific distance-function parameters. In this way, the population across all cyclists $c \in C$ approaches the distribution from Eq. (12) as $|C|$ gets large.

In order to sample from $Z_{\mathcal{B}}$, we use a rejection sampling algorithm where we for all cyclists $c \in C$ first draw a value from a uniform proposal distribution, i.e. $u_c \sim U(0, 1)$. We subsequently draw another uniform value $\acute{u}_c \sim U(0, 1)$, and assign u_c to $z^c_{\mathcal{B}}$ with probability p_a, i.e. if $\acute{u}_c < p_a$. The acceptance probability p_a can – by denoting the probability density function of $\mathcal{B}(\alpha, \beta)$ by f – be found to be,

$$p_a(u_c|\alpha, \alpha) = \frac{f(u_c|\alpha, \alpha)}{f\left(\frac{1}{2}\middle|\alpha, \alpha\right)} = \frac{u_c^{\alpha-1}(1-u_c)^{\alpha-1}}{\left(\frac{1}{2}\right)^{\alpha-1}\left(1-\frac{1}{2}\right)^{\alpha-1}} = (4u_c(1-u_c))^{\alpha-1}, \quad u_c \in [0, 1]. \tag{14}$$

If $\acute{u}_c \geq p_a$, on the other hand, the algorithm starts over until at some point $\acute{u}_c < p_a$ and $z^c_{\mathcal{B}}$ is assigned the latest value of u_c. It can be seen that the acceptance probability is exactly 1 at $u_c = \frac{1}{2}$, 0 at $u_c \in \{0, 1\}$, and that it is positive and strictly less than 1 all other places within the support, i.e. $p_a(.u_c|\alpha, \alpha) \in]0, 1[, \forall u_c \in]0, 1[\setminus \frac{1}{2}, \forall \alpha > 0, \forall c \in C$.

For the minimum desired speed, the corresponding expected headway distance is $E[d^c(v^{\min}_0)] = 2.31$ m, whereas the mean speed provides an expected distance of $E[d^c(\bar{v}^c_0)] = 7.29$ m. The number obtained by the formula is the headway distance including the length of a bicycle, which Andresen et al. (2014) found to be $\lambda^c = 1.73$ m on average.

The headway distance creates an inherent expected upper limit of the outgoing flow on a link as long as cyclists have finite speeds. By assuming no heterogeneity, i.e. $\theta^c_0 = \theta_0$ and $\theta^c_1 = \theta_1$ for all $c \in C$, Eq. (13) can be used to determine the theoretical expected maximum flow per lane at the mean speed,

$$C_l = \left|\Psi^l\right| \times \frac{3600}{\frac{E[d^c(\bar{v}_0)]}{\bar{v}_0}} = \left|\Psi^l\right| \times 3015.75, \quad l \in L, \tag{15}$$

meaning that in theory almost 3016 cyclists per lane can pass a cross section of the network per hour. The heterogeneity of cyclists will most likely reduce this number noticeably in practice, though.

3.4. Lane choice and speed determination

Section 3.1 described that every cyclist has a desired speed that he/she will ride by if the traffic conditions allow him/her to, whereas Section 3.2 introduced the concept of pseudo-lanes into which the traffic is assumed to be divided. Section 3.3 determined the distance needed between two consecutive cyclists in the same pseudo-lane. Combining all this allow us to specify the speed at which a cyclist will traverse a link. Apart from the desired speed of the cyclist the speed is solely a function of the arrival time and speed of the most recent cyclist that has entered the pseudo-lane that the cyclist choose to use.

As such the speed is a direct consequence of the lane choice which in the model is assumed to be performed at the moment the cyclist enters the link, see Section 3.5. A cyclist (c) starts considering the right-most lane ($i = 1$), and searches leftwards (increasing i's) in order to find a pseudo-lane (ψ^l_i) where he/she can go by his/her desired speed (v^c_0). He/she will choose the rightmost satisfactory lane; alternatively – if no lanes allow going by his/her desired speed – he/she chooses the lane with the highest possible speed ($\psi^l_{i\max}$).

Let t_s^l denote the time at which a cyclist reaches link l. A formal formulation of the lane choice process can then be described by Algorithm 1.

Algorithm 1 LaneChoice(c, l, t_s^l).

$\tilde{v}^{\max} \leftarrow 0$;
for i in 1 to $|\Psi^l|$ **do**
  if $\tilde{v}^c_{\psi_i^l}(t_s^l) \geq v_0^c$ **then return** ψ_i^l;
  else if $\tilde{v}^c_{\psi_i^l}(t_s^l) > \tilde{v}^{\max}$ **then**
    $\tilde{v}^{\max} \leftarrow \tilde{v}^c_{\psi_i^l}(t_s^l)$;
    $i^{\max} \leftarrow i$;
return $\psi^l_{i^{\max}}$;

Here $\tilde{v}^c_{\psi_i^l}(t_s^l)$ denotes the speed that the cyclist c would go by, if he/she chooses pseudo-lane ψ_i^l at time t_s^l. Notice that this expression is cyclist specific due to individualised desired speed and headway preferences explaining the superscript c. $\tilde{v}^{\max}$ is the highest possible speed among the pseudo-lanes, and can be obtained at $i^{\max}$'th pseudo-lane from the right, $\psi^l_{i^{\max}}$. Strictly speaking both of these should have an associated c in the notation, but this is omitted since they are only used locally.

The cyclist may choose a slower pseudo-lane, though. The speed assigned to a cyclist when entering link l at time t_s^l denoted by $\tilde{v}_l^c(t_s^l)$ internally includes the lane choice described in Algorithm 1. This speed will be equal to the desired speed if at least one of the pseudo-lanes allows such speed, but will be lower if none of the pseudo-lanes can accommodate the desired speed. In both cases $\tilde{v}_l^c(t_s^l)$ can be described by,

$$\tilde{v}_l^c\left(t_s^l\right) = \min\left\{ \max_{i \in 1,\ldots,\left|\Psi^l\right|} \left\{ \tilde{v}^c_{\psi_i^l}\left(t_s^l\right) \right\}, v_0^c \right\}, \tag{16}$$

with $\tilde{v}^c_{\psi_i^l}(t_s^l)$ still denoting the highest possible speed of pseudo-lane ψ_i^l at time t_s^l for cyclist c.

Let $\tau^c(v)$ be the speed dependent headway time from cyclist c to the cyclist in front of cyclist c. This can be derived using the headway distance function (Eq. (13)),

$$\tau^c(v) = \frac{d^c(v) - \lambda^c}{v} = \frac{\theta_0^c + \theta_1^c\sqrt{v} - \lambda^c}{v}. \tag{17}$$

Notice that v in Eq. (17) can denote any speed of cyclist c, and that any of such time, link, or pseudo-lane specific speeds will have more elaborate notation. Arguments with such detailed notation will be seen frequently as input for $\tau^c(\cdot)$ in the following.

Having introduced the above, we are now ready to present an explicit functional expression for the value of $\tilde{v}^c_{\psi_i^l}(t_s^l)$, arising directly by finding the maximum speed enforcing the headway time of Eq. (17) at the downstream end of the link. Let $t_e^{\psi_i^l}$ be the time at which the back wheel of the previous entrant of the i'th pseudo-lane of link l will leave the link[2] and let t_s^l denote the time at which the tip of the front wheel of the current cyclist reaches the start of link l. Then the current cyclist can at earliest reach the downstream end of the link at $t_e^{\psi_i^l} + \tau^c(\tilde{v}^c_{\psi_i^l}(t_s^l))$ as any earlier exit time would violate the headway time constraint. As speed is defined as the quotient between the distance travelled and the elapsed time, the maximum possible speed of pseudo-lane ψ_i^l for cyclist c can be found to be,

$$\tilde{v}^c_{\psi_i^l}(t_s^l) = \begin{cases} \dfrac{\lambda^l}{t_e^{\psi_i^l} + \tau^c(\tilde{v}^c_{\psi_i^l}(t_s^l)) - t_s^l}, & t_s^l < t_e^{\psi_i^l} + \tau^c(\tilde{v}^c_{\psi_i^l}(t_s^l)) \\ \infty, & t_s^l \geq t_e^{\psi_i^l} + \tau^c(\tilde{v}^c_{\psi_i^l}(t_s^l)) \end{cases} \tag{18}$$

It is worth noting that although the headway constraint will always be enforced at the downstream end of the link, it can be violated at the upstream end of the link if the cyclist arrives shortly after the previous cyclist. However, we find it behaviourally realistic that the cyclist will accept such violation and adapt his/her speed so that the preferred headway distance is met at the downstream end of the link (the next point of evaluation, see Section 3.5) as opposed to instantaneously braking heavily to a full stop or an extremely low speed.

[2] When calculating this number, it is necessary to assume that the previous cyclist will maintain the same speed at the first part (a distance of λ^c = 1.73 m) on the following link. However, the actual speed that the previous cyclist will travel on this link may differ.

As both the function value and domain boundary of $\tilde{v}^c_{\psi_i^l}(t_s^l)$ are functions of $\tilde{v}^c_{\psi_i^l}(t_s^l)$ itself, it turns out that it is beneficial to split Eq. (18) into three separate cases as closed-form expressions with constant domain boundaries can be found when dealing with each case separately. The cases are unified in Section 3.4.4.

The general restricted case

When $t_s^l < t_e^{\psi_i^l} + \tau^c(\tilde{v}^c_{\psi_i^l}(t_s^l))$ it is necessary to expand $\tau^c(\tilde{v}^c_{\psi_i^l}(t_s^l))$ by inserting Eq. (17) into Eq. (18),

$$0 = \frac{(t_e^{\psi_i^l} - t_s^l)^2}{\theta_1^{c^2}} \tilde{v}^c_{\psi_i^l}(t_s^l)^2 + \frac{2(t_e^{\psi_i^l} - t_s^l)(\theta_0^c - \lambda^c - \lambda^l) - \theta_1^{c^2}}{\theta_1^{c^2}} \tilde{v}^c_{\psi_i^l}(t_s^l) + \frac{(\theta_0^c - \lambda^c - \lambda^l)^2}{\theta_1^{c^2}}. \tag{19}$$

This is a second order polynomial of $\tilde{v}^c_{\psi_i^l}(t_s^l)$ with the following solution for $\theta_1^{c^2} + 4(t_e^{\psi_i^l} - t_s^l)(\lambda^c + \lambda^l - \theta_0^c) \geq 0$ and $t_e^{\psi_i^l} \neq t_s^l$,

$$\tilde{v}^c_{\psi_i^l}(t_s^l) = \frac{\theta_1^{c^2} + 2(t_e^{\psi_i^l} - t_s^l)(\lambda^c + \lambda^l - \theta_0^c) - \theta_1^c \sqrt{\theta_1^{c^2} + 4(t_e^{\psi_i^l} - t_s^l)(\lambda^c + \lambda^l - \theta_0^c)}}{2(t_e^{\psi_i^l} - t_s^l)^2}. \tag{20}$$

As mentioned, for this to be valid it is necessary but not sufficient for the radicant of the square-root to be non-negative, i.e.

$$\theta_1^{c^2} + 4\left(t_e^{\psi_i^l} - t_s^l\right)\left(\lambda^c + \lambda^l - \theta_0^c\right) \geq 0,$$

which by assuming positive $\lambda^c + \lambda^l - \theta_0^c$ can be rewritten as,

$$t_s^l \leq t_e^{\psi_i^l} + \frac{\theta_1^{c^2}}{4\left(\lambda^c + \lambda^l - \theta_0^c\right)}. \tag{21}$$

In the two following cases we will determine what happens when $t_s^l = t_e^{\psi_i^l}$ and $t_s^l > t_e^{\psi_i^l} + \frac{\theta_1^{c^2}}{4(\lambda^c + \lambda^l - \theta_0^c)}$.

The singularity case

When $t_s^l = t_e^{\psi_i^l}$, the main denominator of Eq. (21) becomes zero. In this case we need to insert $t_s^l = t_e^{\psi_i^l}$ directly into the first subfunction of Eq. (18) which yields,

$$\tilde{v}^c_{\psi_i^l}(t_s^l) = \frac{\lambda^l}{\tau^c(\tilde{v}^c_{\psi_i^l}(t_s^l))} = \frac{\lambda^l}{\frac{\theta_0^c + \theta_1^c \sqrt{\tilde{v}^c_{\psi_i^l}(t_s^l)} - \lambda^c}{\tilde{v}^c_{\psi_i^l}(t_s^l)}}, \quad t_s^l = t_e^{\psi_i^l}.$$

This can be solved straightforwardly to obtain,

$$\tilde{v}^c_{\psi_i^l}(t_s^l) = \left(\frac{\lambda^l + \lambda^c - \theta_0^c}{\theta_1^c}\right)^2, \quad t_s^l = t_e^{\psi_i^l}. \tag{22}$$

The practically unrestricted case

We now turn to the cases in which Eq. (21) is not fulfilled, i.e. Eq. (20) have no real solutions. It turns out that if we insert the upper bound from Eq. (21) into Eq. (20) we can obtain a lower bound of $\tilde{v}^c_{\psi_i^l}(t_s^l)$ for $t_s^l > t_e^{\psi_i^l} + \frac{\theta_1^{c^2}}{4(\lambda^c + \lambda^l - \theta_0^c)}$,

$$\tilde{v}^c_{\psi_i^l}(t_s^l) \geq 4\left(\frac{\lambda^c + \lambda^l - \theta_0^c}{\theta_1^c}\right)^2, \quad t_s^l > t_e^{\psi_i^l} + \frac{\theta_1^{c^2}}{4(\lambda^c + \lambda^l - \theta_0^c)}. \tag{23}$$

This can be shown to be large enough to largely exceed any reasonable desired speed for realistic values of $\lambda^c, \lambda^l, \theta_0^c$, and θ_1^c. In worst case, the cyclist has drawn and accepted a $z_{\mathcal{B}} = 1$ in Eq. (13), resulting in the parameters $\theta_0^c = -14.03$ and $\theta_1^c = 11.55$. Then using the smallest link length λ^l that will be presented in this study, 20 m, and the constant $\lambda^c = 1.73$, the obtained value will be,

$$\tilde{v}^c_{\psi_i^l}(t_s^l) = 4\left(\frac{1.73 + 20.00 + 14.03}{11.55}\right)^2 = 38.32,$$

corresponding to approximately 138 km/h, which clearly is way beyond any reasonable desired speed for a cyclist. Thus, for any t_s^l exceeding $t_e^{\psi_i^l} + \frac{\theta_1^{c^2}}{4(\lambda^c + \lambda^l - \theta_0^c)}$, we can simply use the lower bound value obtained by Eq. (23).

The situations where $\tilde{v}^c_{\psi^l_i}(t^l_s)$ would not be sufficiently high are if θ^c_1 is very high or if the link length, λ^l, is very short. In such cases it would be necessary to look beyond the downstream end of the link to include the length of the headway distance. This is of course undesirable. However, having very short links or headway distance that increase immensely with increased speeds are in themselves inappropriate, why we consider this as a purely hypothetical limitation of the model. For all relevant practical purposes of the model the $\tilde{v}^c_{\psi^l_i}(t^l_s)$ from Eq. (23) will be higher than any reasonable desired speed of a cyclist.

This means that although t^l_s may be so late that there is no restriction on speed, i.e. $\tilde{v}^c_{\psi^l_i}(t^l_s) = \infty$, we do not have to address such cases when using our model, as the value obtained by using the result from Eq. (23) will end up assigning the same speeds to cyclists in practice. This comes in handy as the domain of Eq. (23) is given through constants, whereas the domain of the last subfunction of Eq. (18) is a function of its value.

Unified expression

Summarising the results from the three cases, Eq. (18) can be expanded to the following piecewise function,

$$\tilde{v}^c_{\psi^l_i}(t^l_s) = \begin{cases} \frac{\theta_1^{c2} + 2(t_e^{\psi^l_i} - t^l_s)(\lambda^c + \lambda^l - \theta_0^c) - \theta_1^c \sqrt{\theta_1^{c2} + 4(t_e^{\psi^l_i} - t^l_s)(\lambda^c + \lambda^l - \theta_0^c)}}{2(t_e^{\psi^l_i} - t^l_s)^2}, & t^l_s < t_e^{\psi^l_i} + \frac{\theta_1^{c2}}{4(\lambda^c + \lambda^l - \theta_0^c)} \wedge t^l_s \neq t_e^{\psi^l_i} \\ (\frac{\lambda^c + \lambda^l - \theta_0^c}{\theta_1^c})^2, & t^l_s = t_e^{\psi^l_i} \\ 4(\frac{\lambda^c + \lambda^l - \theta_0^c}{\theta_1^c})^2, & t^l_s > t_e^{\psi^l_i} + \frac{\theta_1^{c2}}{4(\lambda^c + \lambda^l - \theta_0^c)} \end{cases} \tag{24}$$

Notice that contrary to Eq. (18) the domains and expressions of all three subfunctions are defined solely by constants. The function can be shown to be a monotonously increasing function of t^l_s over the domain of its first subfunction. It is also continuous across the singularity found in $t^l_s = t_e^{\psi^l_i}$. Finally it is continuous in the limit leading to the domain of the third function and constant in the remainder of the domain. As such, the function in Eq. (24) is a monotonously non-decreasing function of t^l_s.

The speed assigned to a cyclist, $\tilde{v}^c_l(t^l_s)$, is used for calculating when he/she is expected to leave the link, but is actually only an upper bound for the speed. The actual speed the cyclist had on the link (v^c_l) can only be calculated retrospectively when it is known exactly when the cyclist has entered the next link, as this may occur later than expected if the next link is fully congested, see Eq. (26) in Section 3.6.

3.5. Longitudinal horizon

As it was evident from Section 3.4, the lane choices have a large influence on how a cyclist move through the network. The fifth assumption states that such choices are only made when entering a new link. This assumption is arguably the most behaviourally questionable. Normally lane choices will happen dynamically such that a cyclist will pull towards the right when finishing overtaking.

Whereas choosing a lane at the beginning of every link may be behaviourally inappropriate, it allows for a very fast implementation as the model can be implemented in an event-based fashion with events only occurring when cyclists leaves or enters a link, making the model scale extremely well.

Note that the lane choice frequency can be adjusted by transforming each link of the network into several smaller links. Although this does not make the longitudinal horizon of lane choices dynamic, it does allow it to have an upper bound as opposed to the fully link dependent horizon achieved by using the raw network.

3.6. Spatial restrictions and spill-back

So far the model has a limitation on the number of cyclists that can pass over a cross section due to the headway time between cyclists. However, until now there is no guarantee that the link can actually contain every bicycle that is allocated to the link.

In order to prevent this a sixth and final assumption is added to the model. It secures that a cyclist, c, cannot enter any pseudo-lane of the link, if he/she would make the occupied length/area of the link exceed the actual length/area of the link, i.e.

$$\begin{cases} \text{Can enter}, & d^c(\tilde{v}^c_l(t^l_s)) \leq L^l - \Omega^l \\ \text{Is denied}, & d^c(\tilde{v}^c_l(t^l_s)) > L^l - \Omega^l \end{cases}, \tag{25}$$

with $L^l = |\Psi^l|\lambda^l$ being the product of the number of pseudo-lanes ($|\Psi^l|$) and the length of the link (λ^l), Ω^l being the space occupied by cyclists on the link, $d^c(v^c(l, t^l_s))$ being the headway distance from Eq. (13), and $\tilde{v}^c_l(t^l_s)$ being the speed assigned to cyclist c when entering link l at time t^l_s as defined in Eq. (16). By summing over $d^c(\tilde{v}^c_l(t^l_s))$ for all cyclists that are currently on the link, Ω^l can be found.

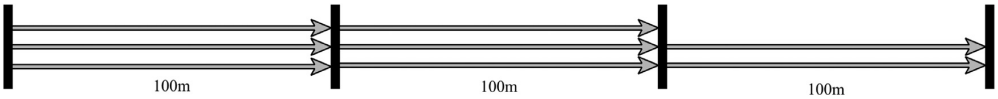


Fig. 6. The toy-network used for small-scale experiments of the proposed model.

When a cyclist is denied, the queue in which he/she is positioned will be processed once a cyclist has successfully left the downstream link. This phenomenon can of course move upstream and cause spill-back over multiple links if the traffic intensity is sufficiently high.

Once the cyclist leaves a link the actual speed of a cyclist on a past link can be found by dividing the length of the link with the time spent on the link. The time waiting to enter a link counts as time spent on the previous link in this regard. This can be formulated mathematically as such: Let λ^l denote the length of link l, t_l^c denote the time at which a cyclist is allowed entrance at link l, and $t_{l^+}^c$ denote the time at which the cyclist is allowed to enter the following link, l^+, then the actual speed at link l is,

$$v_l^c = \frac{\lambda^l}{t_{l^+}^c - t_l^c} \tag{26}$$

This implies that the speed assigned by Eq. (16) is only upper bound of the actual speed the cyclists traverse the link with. The actual speed will be lower if the cyclist cannot enter the following link immediately, whereas the headway distances are always based on the provisionally assigned speed computed by Eq. (16).

3.7. Model input

One of the benefits of the proposed model is that it requires very little input. As any other model simulating a network it requires a set of links L. As opposed to other models, however, for any link $l \in L$ it only requires the width $\omega^l \in \mathbb{R}^+$ and length $\lambda^l \in \mathbb{R}^+$.

In order to have traffic to simulate it also needs demand. It requires the path – an ordered subset of L – to be known for every cyclist c in the model population C. Every such path r_c, $c \in C$ must also have an associated starting time $t^{r_c} \in \mathbb{R}^+$.

Apart from this, the model requires three non-trivial things for input. The first is a lane partitioning formula $|\Psi^l| : \mathbb{R}^+ \to \mathbb{N}^+$ which can divide any link $l \in L$ into a positive integer value of pseudo-lanes. In our proposed implementation it is link indifferent, but it can easily be extended to include input of higher dimensions to account for additional link characteristics. The second is a desired speed distribution $\mathcal{D} : [0, 1] \to \mathbb{R}^+$, which can assign a desired speed v_c to all cyclists $c \in C$ based on a uniformly random number. Finally, the model requires a formula to define the headway distance $d : \mathbb{R}^+ \to \mathbb{R}^+$, which preferably should be a function of speed. Notice that the model also supports individualised headway distance as used in this paper, but that such specification is optional.

4. Experiments

This section is devoted to assessing the capability of the model to give meaningful results when faced with some small-scale experiments. The experimental design is introduced in Section 4.1, before investigating the disaggregate and aggregate properties of the model in Section 4.2. The most questionable assumption concerning the longitudinal horizon of lane choice (see Section 3.5) is analysed in Section 4.3 where simulations with a homogeneous setup are also analysed. Finally, the model is validated against real-life observations in Section 4.4.

4.1. Experiment design

The network used for small-scale tests is the toy network presented in Fig. 6. It consists of three 100 m long links in serial ($\omega^L = [100, 100, 100]$), which are three, three, and two metres wide, respectively, i.e. $\lambda^L = [3, 3, 2]$. Using $|\Psi^l|$ as given by Eq. (11), the number of pseudo-lanes are $|\boldsymbol{\Psi}|^L = [3, 3, 2]$. All cyclists have to parse all three links, i.e. $r_c = (l_1, l_2, l_3), \forall c \in C$.

Such bottlenecks are quite common along major bicycle corridors in Copenhagen. Along these there are generally room for three efficient lanes, but sometimes, due to the street being especially narrow at some places, the bicycle path also has to narrow to two lanes. The bottleneck setup has been chosen because it is has been a frequently used example since Vickrey (1969), but also since it can be difficult for bicycle traffic to reach a congested regime otherwise (Agarwal et al., 2015).

In order to investigate the relationships between flows, speeds, and densities, simulations are conducted using varying levels of traffic intensity. The intensity is changed by controlling the number of cyclists ($|C|$) that have to ride through the network. A total of 199 experiments are run with in-flows on the first link ranging from $|C| = 50$ to $|C| = 10{,}000$ cyclists per hour with a constant increment between experiments of 50 cyclists per hour.

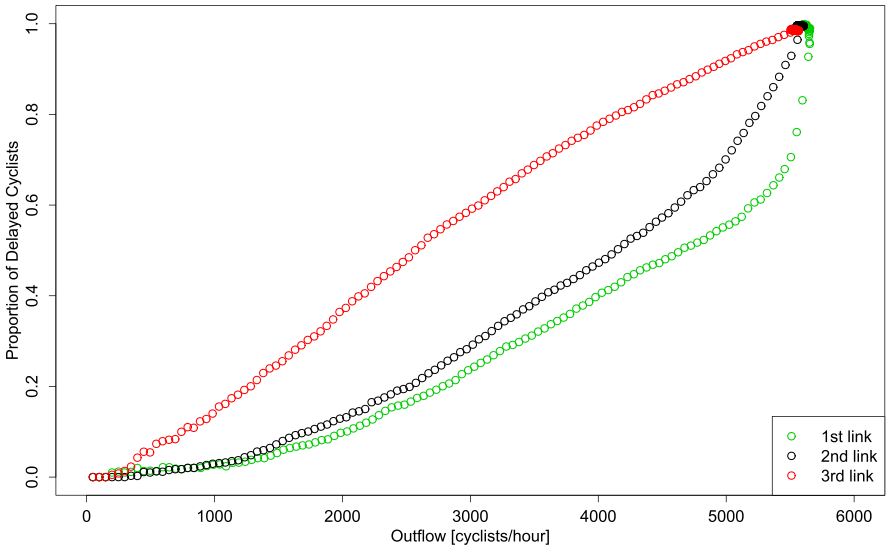


Fig. 7. Proportion of delayed cyclists on each of the three links as a function of outflow.

Every cyclist is assigned a uniformly distributed arrival time between 0 and 3600 s ($t^{r_c} \sim U(0, 3600), \forall c \in C$) using the same fixed seed for all experiments. Likewise, also with a fixed seed, every cyclist is given a desired speed drawn from the Johnson's S_U distribution from Eq. (10), i.e. $v_0^c \sim \mathcal{D}, \forall c \in C$.

In all experiments the network is simulated for 3600 s, meaning that all cyclists will have entered the beginning of the network. There is, however, no guarantee that the network is fully emptied when the simulation period ends.

All small-scale experiments as well as the actual model implementation were carried out in Java. Concerning the computational efficiency, the experiments showed that the running time was proportional to the number of times a cyclist entered a link, using approximately 1 microsecond (10^{-6} s) per such event. This has also proven to hold in pseudo-large-scale experiments on random, spatially unfounded networks. Assuming link lengths of 100 m and an average trip length of 5 km, this corresponds to simulating 1 million bicycle trips per minute, indicating that the model is indeed fast enough to be large-scale applicable in a traffic assignment model.

4.2. Results

In this section we will take a look at the results of the small scale experiments. The section is twofold with the first focusing on the speed heterogeneous properties of the model. The second part focuses on aggregate verification and which patterns are seen in the corresponding fundamental diagrams.

For all results figures, links are enumerated in increasing order from upstream to downstream, and each link keeps its own unique colour throughout the results section.

4.2.1. Speed heterogeneous properties

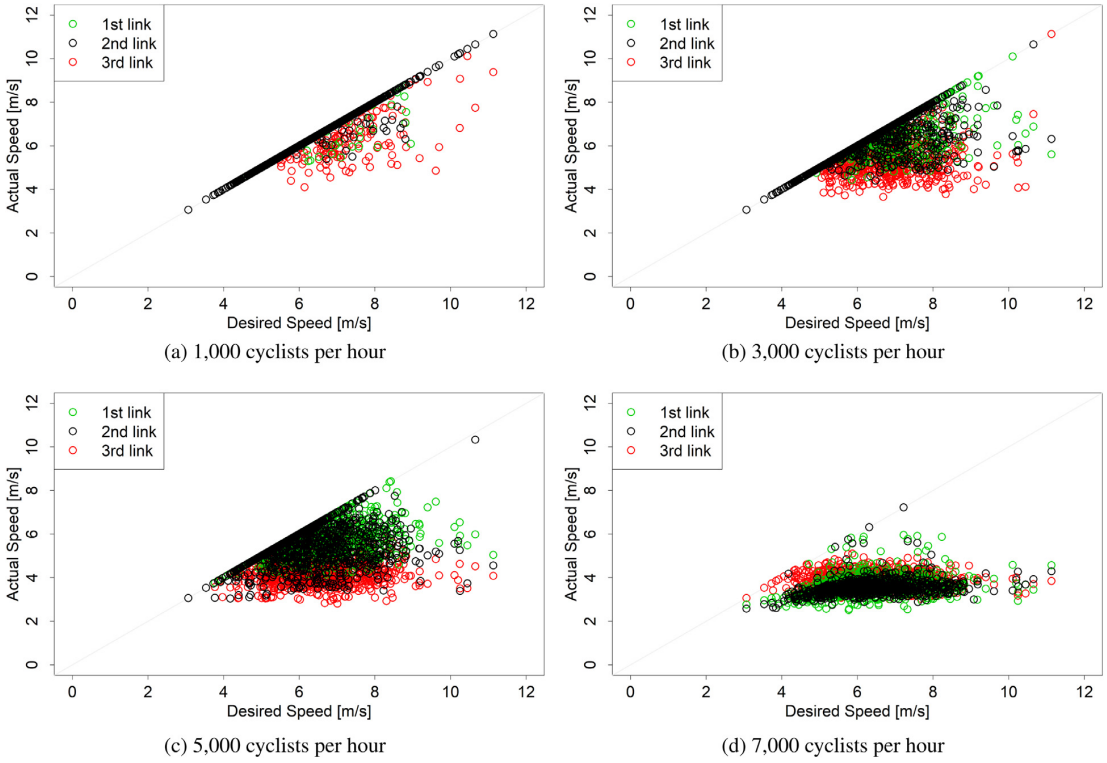
A central element in traffic simulation is the ability to delay users of the network. Fig. 7 shows the proportion of slowed down cyclists on each of the three links given various outflows. A cyclist is defined as being slowed down if he/she has an actual speed strictly lower than his/her desired speed on a link. Outflows are the actual empirical outflows found in the experiments, which often will be lower than the values of $|C|$ under which the experiments were conducted.

Firstly, it is seen that the model is in fact capable of causing delays for some cyclists under moderate traffic intensities, whereas other cyclists are not delayed. At flows of about one third of the capacity, more or less half of the cyclists will be delayed on the final narrow link. This is a feature that macroscopic models cannot include.

Secondly, it is seen that for the first and second link, the proportion of delayed cyclists increases slower for than for the third link when considering low flows. However, since the curves for the two first links are convex, they eventually catch up and also reach a point where everyone will be delayed.

Fig. 8 is constructed in order to show which kind of cyclists get delayed at different traffic intensities. The plots show 1000 actual speeds (from cyclists with ids 1 through 1000) plotted against the desired speeds for four different experiments using different input flows ($|C|$) to the first link.

Each of the four plots represent different phases. The plot based on 1000 cyclists represent the first phase where delays are rare and only occurs for relatively fast cyclists. The delays of this phase happens most often on the third link, as this is



(a) 1,000 cyclists per hour

(b) 3,000 cyclists per hour

(c) 5,000 cyclists per hour

(d) 7,000 cyclists per hour

Fig. 8. Actual speeds of cyclists with id 1 through 1000 as a function of desired speeds for various traffic intensities grouped by link.

only has two lanes and thus a limited opportunity of overtaking. The traffic is not high enough for the bottleneck to be an issue.

The second phase can be seen in the plot based on 3000 cyclists. Under such conditions it is quite normal for cyclists of moderate and fast desired speeds to be delayed. In this phase the delays are frequents on all three links but with the slowest speeds still being found on the last link.

In the third phase represented by 5000 cyclists really high speeds are almost impossible to achieve. Furthermore, the bottleneck is starting to become a serious problem and many cyclist – even the ones with very low desired speeds – experience delays. Finally, the scatters are also shifted downwards a bit (slower actual speeds) compared to the scatters using 3000 cyclists.

In the fourth phase 7000 cyclists are trying to get through the network. This is clearly more than the network can handle, and a serious breakdown occurs on the second and the first link due to the bottleneck. This shows that spill-back effects are clearly captured by the model. Although the traffic fully breaks down to a jam speed of approximate 2.5–4 m/s before the bottleneck, the speeds on the following third link seems almost indistinguishable from what was seen using 5000 cyclists.

To summarise the general effect of delays, the speed distribution for the third link for all experiments where $|C| \leq 7500$ can be found in Fig. 9. Values above 7500 cyclists have not been included in the plot as no noteworthy changes occur beyond this point, probably due to the capacity being reached. It is clearly seen that the distribution becomes steeper, i.e. more homogeneous, as the traffic intensity increases. This is due to the fact that high speeds are no longer possible due to congestion, and that these cyclists have to go slower than they desire. However, even at the very large traffic intensities there is still some heterogeneity present.

Overall the model seems to have the disaggregate properties that were expected from it regarding modelling of speed heterogeneous cyclists.

4.2.2. Aggregate results

We now move on to see how the model performs on an aggregate level. The disaggregate results of the experiments can be aggregated into well-known measures (fundamental diagrams) that are generally used for assessing traffic.

We start off by presenting the speed-flow curves which can be found in Fig. 10, where the individual measurements are enriched with a best fit BPR-curve (Bureau of Public Roads, 1964). Normally the formula describes travel time, but it can be

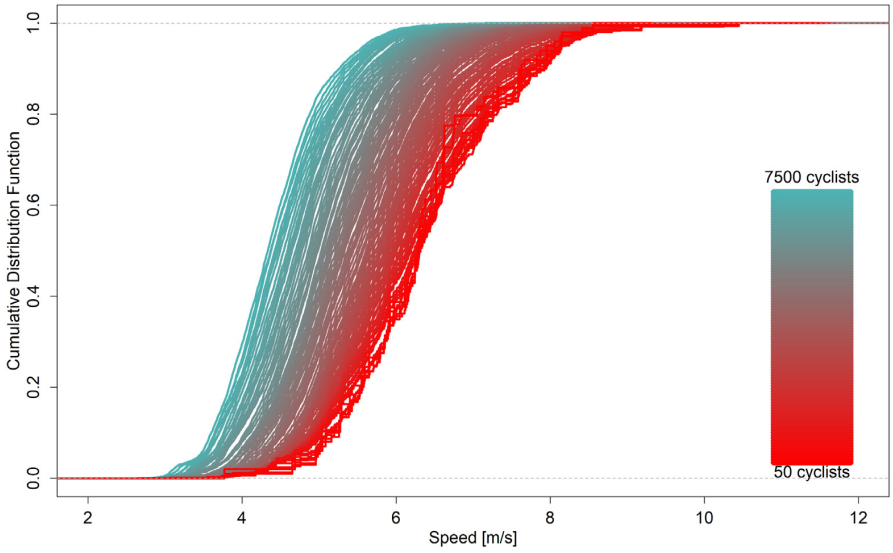


Fig. 9. The cumulative speed distribution for the third link given input flows on the first link ranging from 50 to 7500 cyclists per hour.

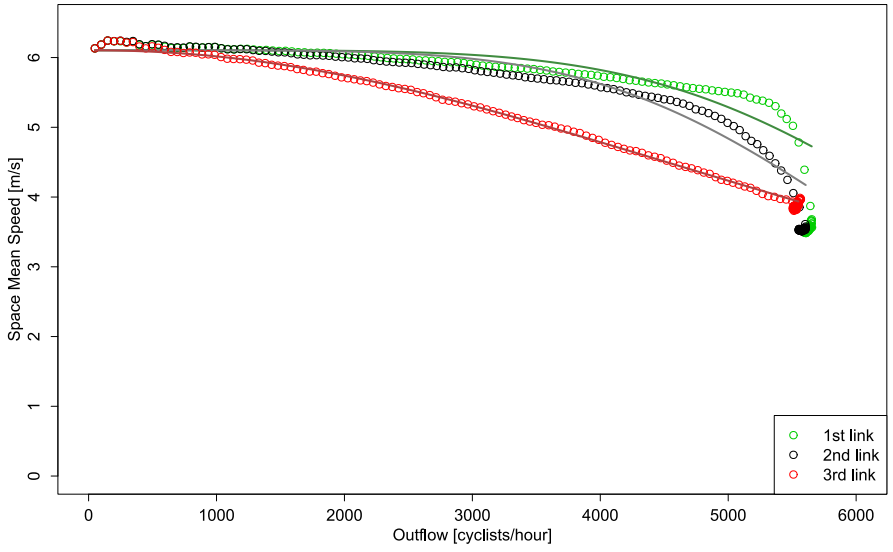


Fig. 10. Space mean speeds for each of the three links as a function of outflow and corresponding best fit BPR curves.

transformed to determine the speed v_l of a link l through the mean speed free flow travel time $\bar{v}_0$, the flow q, the capacity parameter C_l (a soft constraint which can be exceeded), as well as the remaining two parameters α_l and β_l,

$$v_l(q) = \frac{\bar{v}_0}{1 + \alpha_l(\frac{q}{C_l})^{\beta_l}}. \tag{27}$$

The fits are based on the monotonically non-decreasing part of the flows, i.e. for all $|C|$'s from 50 until an increasing $|C|$ no longer provides an increase in the outflow of that link. The α_l, β_l, and C_l parameters are estimated using ordinary least squares, and the parameters for the two relevant links - the first and the second - were found to be 0.43 and 0.74 for α_l, 5.19 and 4.99 for β_l, and 6094.18 and 6155.59 for C_l, respectively.

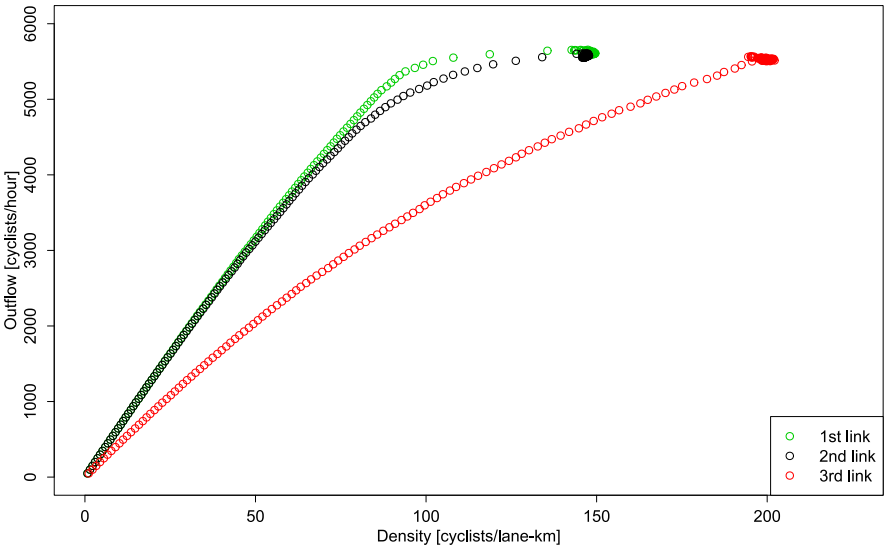


Fig. 11. Outflow for the three links as a function of density.

It is seen that on the link following the bottleneck the measurements coincide surprisingly well with the BPR function. However, on the two constrained links leading to the bottleneck the speed-flow relation cannot reasonably be assumed to follow the same form as the BPR curve. Both curves have a long slightly decreasing linear beginning after which the speeds starts to drop rapidly with a more sudden change than predicted by the BPR curves.

The maximum observed outflow of the third link (and thus also the two previous in this case) is found to be 5606 cyclists per hour, corresponding to 2803 cyclists per hour per pseudo-lane. This is equivalent to between 1933 and 3398 cyclists per hour per meter (in width), as two pseudo-lanes occur for bicycle paths with widths ranging from 1.65 m and 2.9 m using the number of efficient lanes of Buch and Greibe (2015), i.e. Eq. (11).

Previous studies have also dealt with determining the maximum outflow of cyclists per hour per meter. Gould and Karner (2009) found capacities to be in the range from 1395 to 1641 cyclists per hour per meter through simulation. Jin et al. (2015b) and Zhou and Xu (2015) estimated capacities with mean values of 1800 and 2512 cyclists per hour per meter, respectively. Hoogendoorn and Daamen (2016) estimated the capacity to be 1531 cyclists per hour per meter, but emphasised that it was underestimated due to bidirectional traffic. Liang et al. (2018) found that the capacity is just above 2000 cyclists per hour per meter, and that the marginal effect of width is diminishing slightly.

Even though the range of our capacity estimate is quite wide, the found capacity still seems to be comparable to what is seen elsewhere in the literature, although possibly in the high end.

The flow-density relation is found in Fig. 11. The concavity of the third link is barely visible, as opposed to the clear concave tendencies of the second and first links. The concave form is as expected, and it is seen that the flow of the second link greatly exceeds that of the third link until the maximum flow is reached. This will naturally lead to rapid increase in density on the second link, however without ever reaching the higher maximum density found for the third link.

Earlier studies using on-site data (Gould and Karner, 2009; Jin et al., 2015b; Liang et al., 2018) supports that the flow is a concave, increasing function of the density for low and medium densities. However, even though all three studies extrapolates these results to imply a decreasing part of the curve, the empirical evidence to support this claim is only convincing in Liang et al. (2018). Our model seems to be monotonically non-decreasing for the entire observed range of densities, seemingly contradicting the Chinese field observations. It is still formally unknown how the real curve is supposed to look like for saturated bicycle traffic in a Danish context, although unsaturated Danish observations will be presented in Section 4.4. It does however contradict with what is already known for other modes of transport (Hoogendoorn and Knoop, 2013).

Finally, in Fig. 12 we consider speed-density functions. Once again an almost straight and decreasing line is seen for the third link, whereas the other two links show some sign of negative curvature. This shows that the speed is less vulnerable at low densities, but that it is very volatile for medium to high densities. At jam density, however, the speed seems to be constant at about 3.6 m/s (12.8 km/h). This is significantly higher than in Liang et al. (2018), where the jam speed from Chinese observed data is found to be approximately 0.8 m/s. However, Gould and Karner (2009) shows a jam speed of approximately 2.2 m/s which is somewhat closer to the values obtained in this study, indicating that the results could be plausible. Especially when keeping in mind that the speeds are generally faster in Denmark as outlined in Section 2.1.

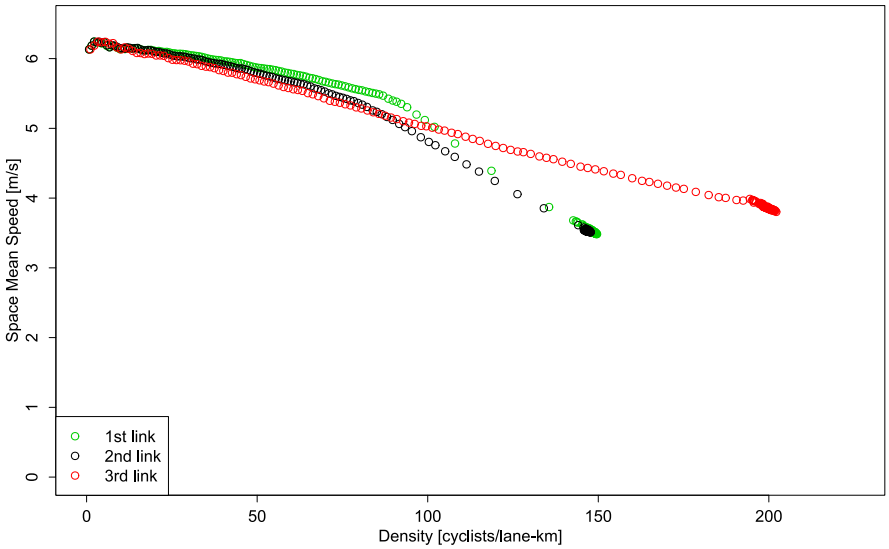


Fig. 12. Space mean speeds for each of the three links as a function of density.

4.3. Sensitivity analyses

In this section two different kinds of sensitivity analyses will be carried out. Section 4.3.1 runs simulation tests where heterogeneity is ignored, whereas Section 4.3.2 focuses on link lengths and the effect their inherent predefined longitudinal horizon of lane choices have.

4.3.1. Ignoring heterogeneity

One of the main drivers of the proposed methodology is its capability to include heterogeneity in desired speeds as well as for headway distances. This section seeks to investigate how the model reacts when assuming strictly homogeneous cyclists instead.

In order to apply homogeneity, all cyclists have to be assigned the same desired speed and headway preference. The used desired speed is the space mean speed of the estimated Johnson S_u distribution, see Section 2.1, namely by $\bar{v}_0 = 6.104$ m/s. The corresponding headway parameters are based on Model 5 (Section 2.3.1), but whilst assuming $z_B^c = \frac{1}{2}, \forall c \in C$, in Eq. (12) corresponding to assigning $\theta_0^c = \theta_0$ and $\theta_1^c = \theta_1$ for every cyclist $c \in C$. Running the model using the input parameters presented above results in the speed-flow curve presented in Fig. 13.

It can be seen that when using homogeneous input, the straight part of the speed flow curve becomes completely flat, and remains so until an average outflow of about 4000 cyclists per hour. Above this, the congested regime is entered for the first two links, and the curvature fits very well with the best fit BPR-curves also drawn in the figure. In fact, the results obtained by the model could in this case easily and almost entirely without any loss of information be modelled through the corresponding BPR-curve. This is in contrast to the heterogeneous counterpart, Fig. 10, where substituting the simulated results with a BPR curve would alter the results considerably. In the homogeneous case, the α_l parameters of the BPR curve for the first and the second link – the two relevant links – are 14.17 and 56.19, respectively, the β_l parameters are 56.19 and 15.23, and the capacity parameters C_l are 6303.98 and 6057.30.

It should also be noted that even though we see the largest reduction of speed before the bottleneck in both the homogeneous and heterogeneous case, the heterogeneity actually causes speed reductions after the bottleneck too. This is an important feature of the model – and of bicycle traffic in general – that is completely ignored when assuming homogeneity.

Due to the homogeneity the cyclists can be “packed” more tightly than when dealing with heterogeneous cyclists and thus provide a higher max flow, equivalent to what is known for railway traffic for instance. Assuming homogeneity increases the maximum flow from 5717 to 5886, meaning that the realised capacity of bicycle paths would be overestimated when assuming homogeneity. It also increases the jam speed from 12.8 m/s to 14.5 m/s.

Conclusively, when ignoring heterogeneity it seems as if the model behaves very similarly to traditional flow-based models. This is a nice feature since the model then appropriately can be used on any distribution of desired speeds and headways – even when the variance of these are small.

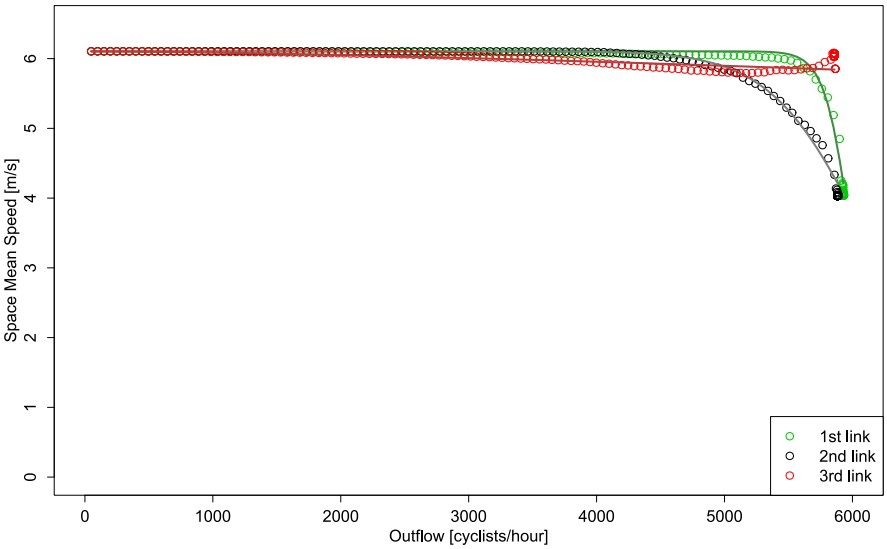


Fig. 13. Space mean speed as a function of outflow when assuming full homogeneity and corresponding best fit BPR curves.

4.3.2. Link lengths

The results presented so far have all arisen from a setup with 100 m long links. As introduced in Section 3.5, in the proposed model the link length directly determines the longitudinal horizon for the choice of lanes which subsequently determines the speed. As an indirect consequence it also determines the frequency of lane choices and hence how often the speed can be changed. In this section, we will perform a sensitivity analysis that explores the consequences of adjusting this by repeating the experiments with links lengths of 20, 50, 200, and 500 m, respectively.

We first evaluate how the proportions of delayed cyclists are affected by the link length. The relevant plots can be found in Fig. 14. The model can be seen to keep its capability of delaying cyclists under moderate traffic intensities regardless of the link length. The curvature of the lines are influenced considerably by the link lengths though, with longer links leading to more concave functions and vice versa. The effect on the narrower third link is especially notable. Here the limited possibility of overtaking due to only having two lanes becomes paramount when the link lengths get very long, whereas the problem is minor when having a short link length. In fact, when having a very short link length of 20 m the proportion of delayed cyclists on the third link will never exceed 95 % as opposed to the other investigated link lengths.

Fig. 15 shows the effect on the speed distributions for the last link. It generally shows that the longer the link, the prompter the effect kicks in. Using a very long link length (500 m) seems to be a little too extreme, causing a dramatic speed reduction even for low traffic intensities. Reducing the links to a very short length (20 m), on the other hand, seems to invoke a more speed heterogeneous (flatter) distribution where the effect of increased traffic intensity is somewhat smaller.

Finally, we evaluate how the link length affects one of the fundamental diagrams, namely the speed-flow relation (see Fig. 16). When using a link length of 20 m or 50 m the curvature of the second link is rather soft, never allowing the flow to be large enough to force the curve to be really steep. In these two cases the corresponding BPR-curves fit tremendously well, with the parameters for the second link being $(\alpha, \beta) = (0.91, 4.29)$ and $(0.91, 6.30)$ for 20 m and 50 m, respectively. For the longer links, the final part of the curves are very steep, eventually curving inwards towards the y-axis as known from speed-flow diagrams of other modes. Interestingly, when having long links (200 m and 500 m) the effect of the restricted possibility to overtake exceeds that of the bottleneck right until the flow capacity is reached. For link lengths of 50 m this happens slightly before the capacity is reached, whereas it happens considerably earlier when using really short links lengths of 20 m where the bottleneck effect is more prominent and dominates the effect of restricted overtaking for high traffic intensities.

Another interesting finding is that the maximum flow is slightly higher when using a link length of 50 or 200 m, than when using 20 m. This seems counter-intuitive at first, as having link lengths of 20 m would improve the conditions for overtaking as lane choices happen much more frequently. However, the links also becomes so small, that the surplus space that is not used – because adding another cyclist to the link would make the occupied area of the link exceed the actual area, see Section 3.6 – constitutes a considerable part of the area. This means that the links will be poorly utilised in such cases, which most likely can explain the surprisingly lower flows using very short links. This is supported by the fact that the maximum average density found in the experiments is much lower for these short links (133 cyclists per lane-km), as opposed to 141, 148, 155, and 162 cyclists per lane-km for links of length 50, 100, 200, and 500 m, respectively.

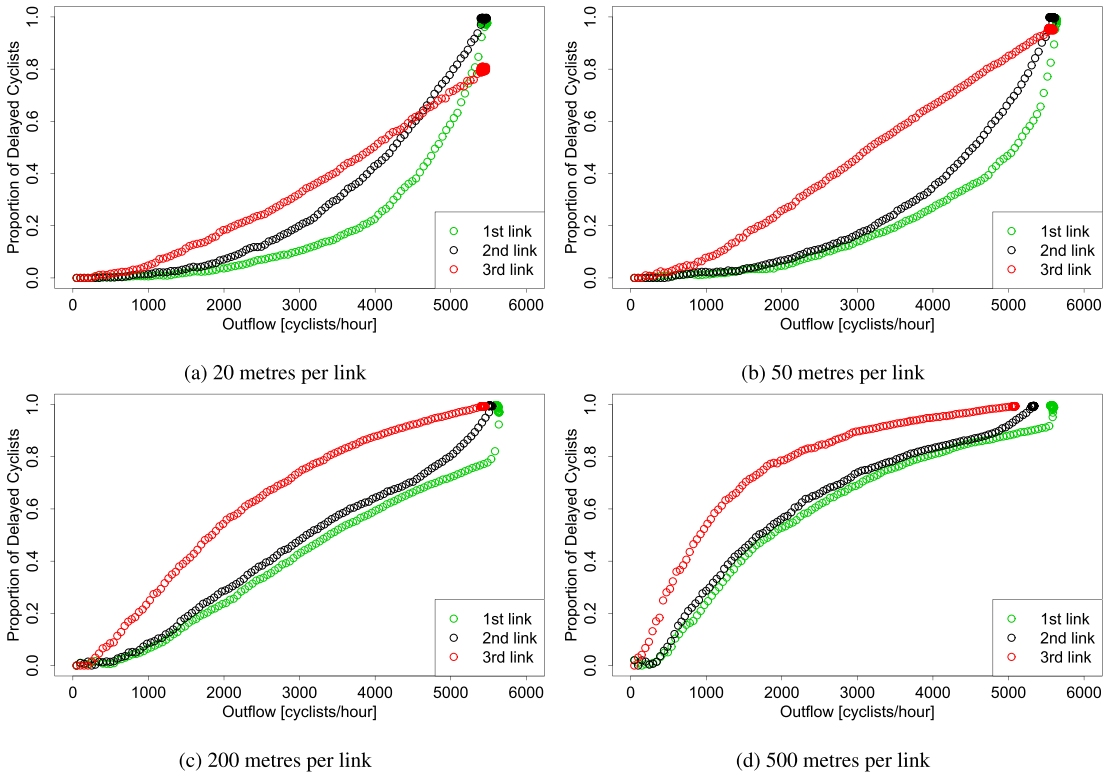


Fig. 14. Sensitivity analysis of slowed down proportion as a function of outflow for various link lengths.

Although the model has been found to be rather sensitive to changes in link length, it seems to behave reasonably as long as the link lengths do not become too short or too long. In general, using $\lambda^l = 50, \forall l \in L$ provide results which intuitively seem plausible.

4.4. Model validation

In order to validate the model, on-site observations have been collected from Queen Louise's Bridge in central Copenhagen, using the same video-approach as in Section 2. Queen Louise's Bridge is the street with the highest maximum bicycle flow in Copenhagen with 3763 cyclists between 8 a.m. and 9 a.m. in the direction towards the city centre according to City of Copenhagen (2018). This has been the primary reason for choosing this particular street. A single frame of the video data that covered a time period from 7.30 a.m to 8.30 a.m. on a weekday morning can be found in Fig. 17.

The bicycle path on the street is 3.9 m wide, which corresponds to three efficient lanes according to Eq. (11). As opposed to our test network there is no bottleneck following Queen Louise's Bridge. Because of this, we have generated new simulation results using a modified network where all of the three links have three efficient lanes (the last link only had two lanes in the test network, see Fig. 6). Notice that according to Eq. (11) widths of both 3.0 and 3.9 m corresponds to three efficient lanes. This allows the results from the model to be compared directly with the empirically collected observations.

The video data has been subdivided into time intervals of 30 s in which flows as well as average speeds and densities are calculated. Jin et al. (2015b) uses a temporal division of the collected data of 15 s but states that finding the optimal time interval is still a topic of further research. With our data we have found that using a time interval of 30 s, as also seen in Gould and Karner (2009), seems as a reasonable interval.

Two cross sections are drawn across the bicycle path 10 and 20 m away from the camera, respectively. Although a sample area length of 10 m may seem small, it is actually larger than all measurements used in Gould and Karner (2009) and Jin et al. (2015b) which use lengths down to 2.9 m. The flow is determined by the number of cyclists crossing the second cross section within the time interval, and adjusted to be in the unit cyclist per lane per hour. Speeds for every cyclist are calculated as the distance divided by the time between the two cross sections. The average speed are subsequently calculated by the harmonic mean of the speeds of all cyclists crossing the second cross section within the time interval.

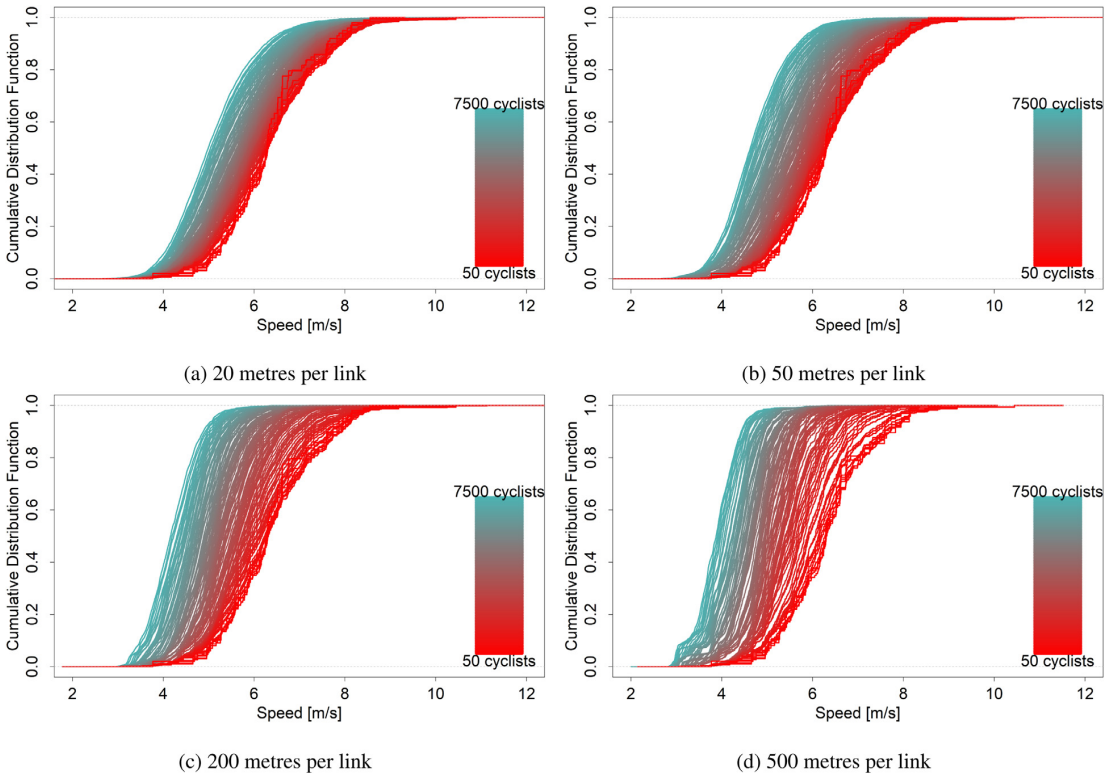


Fig. 15. Sensitivity analysis of the speed distribution for the last link as a function of cyclists inflows on the first link for various link lengths.

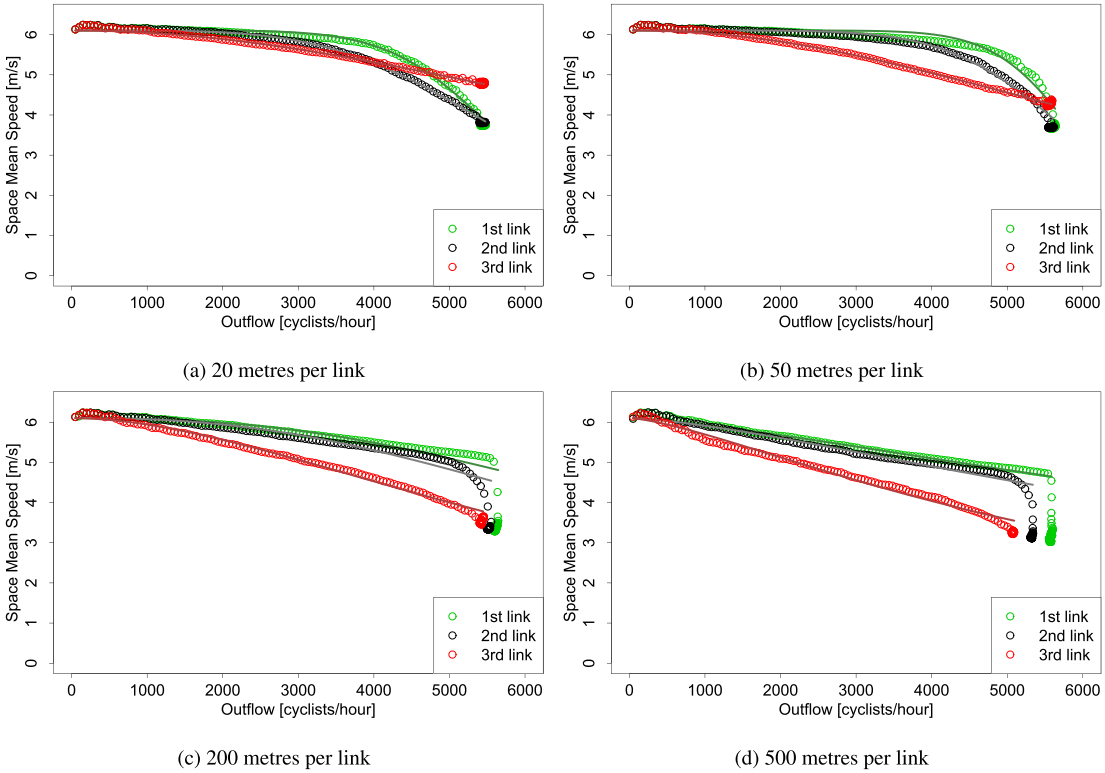
For determining densities, each time interval is subdivided into slices of 0.1 s. In each of such slices, the number of cyclists between the two cross sections is counted, and the average density of a time interval is then defined as the average number of cyclists in the area per time slice divided by the distance between the two cross sections and the number of lanes.

The corresponding fundamental diagrams emerging from the observed data and for the non-bottleneck simulations can be found in Fig. 18. Firstly, the plots show that without the bottleneck setup it is very hard to reach the saturated regime of the curves. This is a general problem for bicycle traffic (as also mentioned in Agarwal et al., 2015), and we see the problem here for both the simulations and for the empirical data. However, as the street from which the data is collected is the street with the highest traffic intensity for bicycles in all of Copenhagen it probably still is the best obtainable data to compare against.

The speed-flow and speed-density curves show that our simulations have speeds that are a little too high (roughly 10%) for low flows and densities compared to the observed data. This could indicate that there are some biases in either of the two video datasets that we have not corrected for, e.g. wind, gradient, or quality of the asphalt. Despite the starting point of the observations being a little too low for both the speed-flow and speed-density diagrams, we do observe a decreasing trend on the right half of the data that is very similar to that of the simulations.

Regarding the flow-density diagram we see surprisingly similar results for the empirical data and the simulations across the entire spectrum of densities. However, on this part of the curve the tendency is almost strictly linear, and only the very beginning of a concave tendency can be seen from the data. This is in line with what is seen for low densities for empirical data of traffic that can be modelled appropriately by a BPR curve (Nielsen and Jørgensen, 2008). The simulations do have a concave form although it is far less evident than when using the bottleneck setup, see Fig. 10.

Overall, although our model seems to predict a little too high speeds for freespeed-like situations the simulated results are very close to the observed data for high traffic intensities. This is in particular striking when keeping in mind that the model is calibrated using external data collected at lower traffic intensities and at a different location, i.e. it is not calibrated using the data it is compared against in contrast to most other models on bicycle traffic. This is a deliberate choice since one of the aims of this study is to model bicycle traffic at intensities exceeding what is currently observable at bicycle paths. As such, it is fully expected that our current model has some deviations relative to the empirical data. The deviations between



(a) 20 metres per link

(b) 50 metres per link

(c) 200 metres per link

(d) 500 metres per link

Fig. 16. Sensitivity analysis of speed-flow relations for varying link lengths.

Fig. 17. A single frame from the video data collected at Queen Louise's Bridge in Copenhagen.

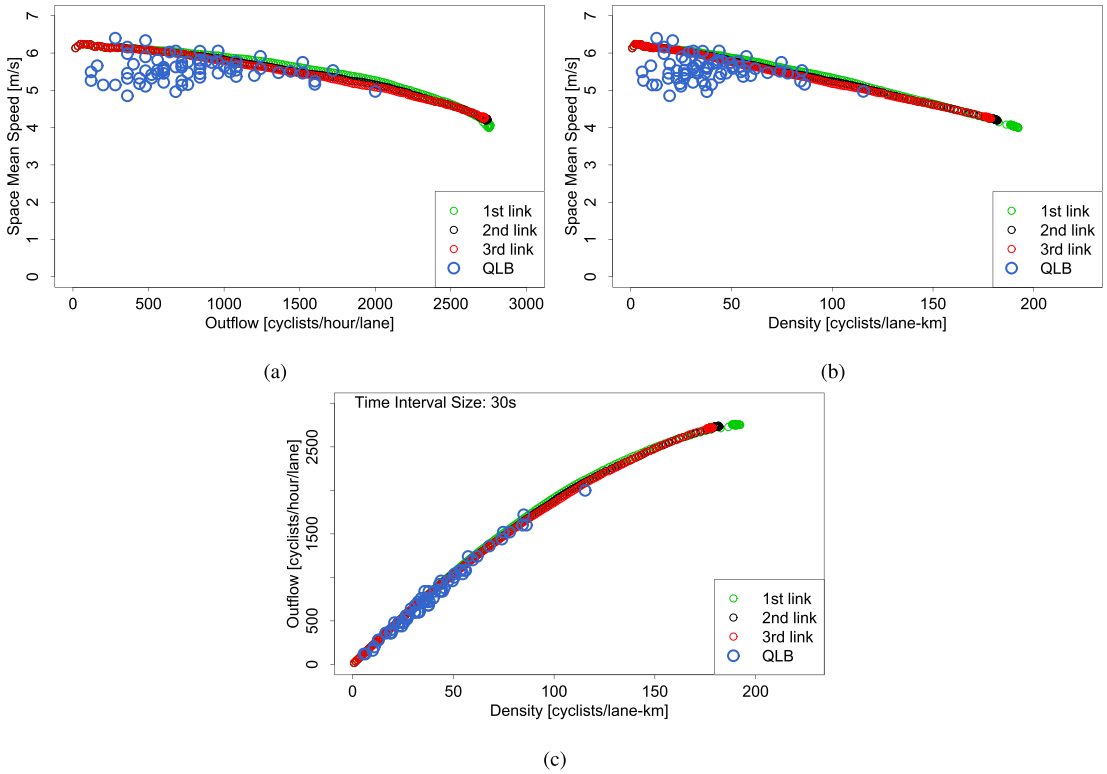


Fig. 18. Fundamental diagrams of non-bottleneck simulations with three links each with three efficient lanes as well on-site observations from Queen Louise's Bridge (QLB).

the model and observations may later be minimised by a slight adjustment/recalibration of the input distributions, meaning that these minor discrepancies should not be considered as evidence against the proposed methodology as a whole.

It has obviously not been possible to perform a fully disaggregate evaluation of on-site observations. The desired speed can only be revealed in uncongested situations and the fundamental diagrams require congestion to be established. As such there is no straightforward way to acquire both types of information simultaneously through the available video data.

5. Conclusions and future work

The purpose of this paper was to develop a bicycle traffic simulation model capable of handling cyclist speed heterogeneity and its inherent interactions between cyclists. Individual desired speeds from Johnson's S_U distribution as well as individual beta-distributed headway distance parameters were assigned for each cyclist. Using simple assumptions on lane formation and choices on multi-lane bicycle paths, the simulation results of a bottleneck test network revealed that differences in desired speeds and headway distances are indeed one possible way to explain and model delays on bicycle paths. Even though the likelihood of such delays increases with the traffic intensity, delays also occur at moderate flow levels – especially for cyclists with high desired speeds due to the implicit modelling of overtaking restrictions. Although the assumptions underlying the model are simple, the resulting fundamental diagrams showed reasonable resemblance to previous empirical analyses of bicycle traffic as well as existing simulation-based analyses found in literature. Also the resulting lane capacities of the model were reasonably close to what has previously been found in the literature.

Furthermore, the model was also validated against data from the location with the highest possible bicycle traffic intensity in Copenhagen. However, it was found that under such non-bottleneck setup it is very difficult for bicycle traffic to reach a fully saturated regime. It neither happened in our model nor in the empirical data, why the model validation was only possible for the unsaturated part of the fundamental diagrams. Although the baseline speeds were a little too high in the simulation results compared to the empirical observations, simulation model resulted in fundamental diagrams that fit the real-life observations well for moderate and high traffic intensities.

It was also shown that accounting for heterogeneity altered the fundamental diagrams by causing speed reductions at an earlier stage than when using a fully homogeneous setup. This suggests that such heterogeneity should not be ignored when modelling bicycle traffic.

The disaggregate aspects of the model that allow delays to occur under low traffic intensities could unfortunately not be validated using the acquired video data, since it is impossible to derive desired speeds of cyclists in congested situations. It could be interesting to look further into how to collect data to support such analysis in future work, for instance through providing GPS trackers to a small portion of the cyclists that are known to regularly pass through the video analysis area.

Link lengths play an important role in the proposed model, as they determine the longitudinal horizon for lane choices. Therefore, a sensitivity analysis was carried out concerning the link lengths. Overall, it seemed as if the model behaved reasonably when varying the link lengths. It did, however, also show evident changes in the curves, and that it seemed as if having too short or too long links can make the model behave undesirably. The model was relatively stable in the interval from 50 m to 100 m, though, with both values providing appropriate results.

Besides seemingly being theoretically well-founded, the proposed model also proved to scale well, with low running times increasing linearly with the number of cyclist/link interactions. This shows that the model is most likely large-scale applicable, and will be suitable for implementing in a traffic simulation for an entire metropolitan area. The current study only deals with link modelling, though. In order to have a model that also works realistically when cyclists have routes that cross each other and those of other modes, a proper intersection modelling that can be integrated with the model proposed in this study needs to be developed too.

The model can be used for various performance and policy analysis purposes. For instance, it can be used to provide insight into determining at which point the current width of bicycle paths will no longer provide sufficient capacity for the traffic demand and hence should be extended to optimise traffic flow. It can also be used to investigate the consequences of the introduction of a new type of mode on the bicycle paths, e.g. a large share of high-speed pedelecs. This would make the desired speed distribution highly bipolar, making the model even more relevant to use and general methods based on averages even more inappropriate. The model would also be able to predict to which degree the pedelecs would be able to reach their full speed potential when riding on the bicycle paths.

Most importantly, however, the theoretical foundation developed in this paper serves as an important step towards allowing bicycle traffic to be modelled at a similar level-of-detail to that of car traffic including detailed congestion patterns. This allows for the travel time of the two modes to be compared fairly in the analyses of future scenarios. Moreover, it would be possible to more accurately investigate the effects on major investments in bicycular infrastructure as a way to increase urban mobility in metropolitan areas.

Acknowledgements

The authors would like to sincerely thank Rasmus Albrink from RAW Mobility Ltd[3] (previously working for COWI Ltd[4]) for collecting, preprocessing, and discussing relevant aspects of the obtained video data. COWI Ltd kindly allowed us to use their data from Smallegade for analysing and estimating key model parameters from empirical observations. RAW Mobility Ltd collected the video data from Queen Louise's Bridge used to establish the on-site fundamental diagram for congested bicycle traffic, which we are also very thankful for. We would also like to thank the two anonymous referees for their highly constructive criticism which helped improve the quality of the paper.

References

Agarwal, A., Zilske, M., Rao, K.R., Nagel, K., 2015. An elegant and computationally efficient approach for heterogeneous traffic modelling using agent based simulation. Procedia Comput. Sci. 52, 962–967. doi:10.1016/j.procs.2015.05.173.

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, N.B., Csaki, F. (Eds.), Proceedings of the 2nd International Symposium on Information Theory, pp. 267–281. Akadamiai Kiado, Budapest.

Allen, D., Rouphail, N., Hummer, J., Milazzo, J., 1998. Operational analysis of uninterrupted bicycle facilities. Transport. Res. Rec. 1636 (98), 29–36. doi:10.3141/1636-05.

Andresen, E., Chraibi, M., Seyfried, A., Huber, F., 2014. Basic driving dynamics of cyclists. In: Behrisch, M., Krajzewicz, D., Weber, M. (Eds.), Simulation of Urban Mobility. Springer, Berlin, Heidelberg, pp. 18–32. doi:10.1007/978-3-662-45079-6_2.

Bernardi, S., Rupi, F., 2015. An analysis of bicycle travel speed and disturbances on off-street and on-street facilities. Transport. Res. Procedia 5, 82–94. doi:10.1016/j.trpro.2015.01.004.

Buch, T.S., Greibe, P., 2015. Analysis of bicycle traffic on one-way bicycle tracks of different width. In: European Transport Conference 2015. Association for European Transport (AET), Frankfurt, Germany. https://aetransport.org/en-gb/past-etc-papers/conference-papers-2015?abstractId=4418&state=b.

Bureau of Public Roads, 1964. Traffic Assignment Manual for Application with a Large, High Speed Computer. U.S. Dept. of Commerce, Urban Planning Division, Washington D.C.. https://catalog.hathitrust.org/Record/000968330.

Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput. 16 (5), 1190–1208. doi:10.1137/0916069.

City of Copenhagen, 2018. Trafiktal [Data Sheet]. https://data.kk.dk/dataset/trafiktal/resource/50f7a383-653a-4860-bb4e-306f221a2d2a

Flügel, S., Hulleberg, N., Fyhri, A., Weber, C., Ævarsson, G., 2017. Empirical speed models for cycling in the Oslo road network. Transportation doi:10.1007/s11116-017-9841-8.

Gould, G., Karner, A., 2009. Modeling bicycle facility operation. Transport. Res. Rec. 2140 (1), 157–164. doi:10.3141/2140-17.

Hoogendoorn, S., Daamen, W., 2016. Bicycle headway modeling and its applications. Transport. Res. Rec. 2587, 34–40. doi:10.3141/2587-05.

[3] https://www.rawmobility.dk.

[4] https://www.cowi.com.

Hoogendoorn, S.P., 2005. Unified approach to estimating free speed distributions. Transport. Res. Part B 39 (8), 709–727. doi:10.1016/j.trb.2004.09.001.
Hoogendoorn, S.P., Knoop, V., 2013. Traffic flow theory and modelling. In: van Wee, B., Annema, J.A., Banister, D. (Eds.), The Transport System and Transport Policy: An introduction. Edward Elgar Publishing Ltd, pp. 125–159.
Jensen, P., Rouquier, J.-B., Ovtracht, N., Robardet, C., 2010. Characterizing the speed and paths of shared bicycle use in Lyon. Transport. Res. Part D 15 (8), 522–524. doi:10.1016/j.trd.2010.07.002.
Jia, B., Li, X.-G., Jiang, R., Gao, Z.-Y., 2007. Multi-value cellular automata model for mixed bicycle flow. Eur. Phys. J. B 56 (3), 247–252. doi:10.1140/epjb/e2007-00116-5.
Jiang, R., Hu, M.-B., Wu, Q.-S., Song, W.-G., 2017. Traffic dynamics of bicycle flow: experiment and modeling. Transport. Sci. 51 (3), 998–1008. doi:10.1287/trsc.2016.0690.
Jiang, R., Jia, B., Wu, Q.-S., 2004. Stochastic multi-value cellular automata models for bicycle flow. J. Phys. A 37 (6), 2063–2072. doi:10.1088/0305-4470/37/6/007.
Jin, S., Qu, X., Xu, C., Ma, D., Wang, D., 2015. An improved multi-value cellular automata model for heterogeneous bicycle traffic flow. Phys. Lett. A 379 (39), 2409–2416. doi:10.1016/j.physleta.2015.07.031.
Jin, S., Qu, X., Zhou, D., Xu, C., Ma, D., Wang, D., 2015. Estimating cycleway capacity and bicycle equivalent unit for electric bicycles. Transport. Res. Part A 77, 225–248. doi:10.1016/j.tra.2015.04.013.
Johnson, N.L., 1949. Systems of frequency curves generated by methods of translation. Biometrika 36 (1/2), 149–176. doi:10.2307/2332539.
Lan, L.W., Chang, C.-W., 2005. Inhomogeneous cellular automata modeling for mixed traffic with cars and motorcycles. J. Adv. Transport. 39 (3), 323–349. doi:10.1002/atr.5670390307.
Li, B.-X., Fang, L., 2017. An improved multi-value cellular automata model for mixed bicycle traffic flow on campus. In: Proceedings of the 2nd Annual International Conference on Electronics, Electrical Engineering and Information Science (EEEIS 2016). Atlantis Press, Paris, France doi:10.2991/eeeis-16.2017.107.
Liang, X., Mao, B., Xu, Q., 2012. Psychological-physical force model for bicycle dynamics. J. Transport. Syst. Eng.Inf. Technol. 12 (2), 91–97. doi:10.1016/S1570-6672(11)60197-9.
Liang, X., Xie, M., Jia, X., 2018. New microscopic dynamic model for bicyclists' riding strategies. J. Transport. Eng. Part A 144 (8), 1–16. doi:10.1061/JTEPBS.0000148.
Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads. Proc. R. Soc. A 229 (1178), 317–345. doi:10.1098/rspa.1955.0089.
Liu, H., Wang, H., Feng, Y., 2008. Bicycle flow modeling based on cellular automata. In: Proceedings of the 27th Chinese Control Conference. IEEE, Kunming, Yunna, China, pp. 527–531. doi:10.1109/CHICC.2008.4605463.
Luo, Y., Jia, B., Liu, J., Lam, W.H.K., Li, X., Gao, Z., 2015. Modeling the interactions between car and bicycle in heterogeneous traffic. J. Adv. Transport. 49 (1), 29–47. doi:10.1002/atr.1257.
Mai, X., Lv, W., Wei, X., Song, W., Jiang, R., 2013. Analyzing the characteristics of unidirectional bicycle movement around a track based on digital image processing. Procedia Eng. 62, 519–524. doi:10.1016/J.PROENG.2013.08.095.
Mohammed, H., Bigazzi, A.Y., Sayed, T., 2019. Characterization of bicycle following and overtaking maneuvers on cycling paths. Transport. Res. Part C 98, 139–151. doi:10.1016/j.trc.2018.11.012.
Nagel, K., Schreckenberg, M., 1992. A cellular automaton model for freeway traffic. J. Phys. I 2 (12), 2221–2229. doi:10.1051/jp1:1992277.
Nielsen, O.A., Jørgensen, R., 2008. Estimation of speedflow and flowdensity relations on the motorway network in the greater Copenhagen region. IET Intell. Transport Syst. 2 (2), 120. doi:10.1049/iet-its:20070024.
Osowski, C., Waterson, B., 2017. Establishing the validity of cycle path capacity assumptions in the Highway Capacity Manual. Int. J. Sustain. Transport. 11 (6), 422–432. doi:10.1080/15568318.2016.1266424.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. http://www.r-project.org/.
RCE Systems R.S.O., 2014. Data From Sky [Traffic Monitoring Tool]. http://www.datafromsky.com.
Richards, P.I., 1956. Shock waves on the highway. Oper. Res. 4 (1), 42–51. doi:10.1287/opre.4.1.42.
Rui, J., Mao-Bin, H., Qing-Song, W., Song, W.-G., 2017. Traffic dynamics of bicycle flow: Experiment and modeling. Transport. Sci. 51 (3), 998–1008. doi:10.1287/trsc.2016.0690.
Shan, X., Li, Z., Chen, X., Ye, J., 2015. A modified cellular automaton approach for mixed bicycle traffic flow modeling. Discrete Dyn. Nat. Soc. 2015, 1–11. doi:10.1155/2015/213204.
Shen, F., Wang, H., Liu, H., 2011. A mesoscopic model for bicycle flow. In: Proceedings of the 30th Chinese Control Conference, pp. 5574–5577. Yantai, China. https://ieeexplore.ieee.org/document/6001058/.
Tang, T.-Q., Huang, H.J., Shang, H.Y., 2011. A macro model for bicycle flow and pedestrian flow with the consideration of the honk effects. Int. J. Modern Phys. B 25 (32), 4471–4479. doi:10.1142/S0217979211059462.
Tang, T.Q., Rui, Y.X., Zhang, J., Shang, H.Y., 2018. A cellular automation model accounting for bicycle's group behavior. Physica A 492, 1782–1797. doi:10.1016/j.physa.2017.11.097.
Vasic, J., Ruskin, H.J., 2012. Cellular automata simulation of traffic including cars and bicycles. Physica A 391 (8), 2720–2729. doi:10.1016/j.physa.2011.12.018.
Vickrey, W.S., 1969. Congestion theory and transport investment. Am. Econ. Rev. 59 (2). 251–160. http://www.jstor.org/stable/1823678.
Wasielewski, P., 1979. Car-following headways on freeways interpreted by the semi-poisson headway distribution model. Transport. Sci. 13 (1), 36–55. doi:10.1287/trsc.13.1.36.
Wilks, S.S., 1938. The large-sample distribution of the likelihood ratio for testing composite hypotheses. Ann. Math. Stat. 9 (1), 60–62. doi:10.1214/aoms/1177732360.
Wolfram, S., 1986. Theory and Applications of Cellular Automata: Including Selected Papers, 1983–1986. World Scientific, Cambridge, Massachusetts.
Wong, G., Wong, S., 2002. A multi-class traffic flow model an extension of LWR model with heterogeneous drivers. Transport. Res. Part A 36 (9), 827–841. doi:10.1016/S0965-8564(01)00042-8.
Xue, S., Jia, B., Jiang, R., Li, X., Shan, J., 2017. An improved Burgers cellular automaton model for bicycle flow. Physica A 487, 164–177. doi:10.1016/j.physa.2017.05.036.
Zhang, J., Mehner, W., Andresen, E., Holl, S., Boltes, M., Schadschneider, A., 2013. Comparative analysis of pedestrian, bicycle and car traffic moving in circuits. Procedia 104, 1130–1138. doi:10.1016/J.SBSPRO.2013.11.209.
Zhang, S., Ren, G., Yang, R., 2013. Simulation model of speeddensity characteristics for mixed bicycle flow. Comparison between cellular automata model and gas dynamics model. Physica A 392 (20), 5110–5118. doi:10.1016/j.physa.2013.06.019.
Zhao, D., Wang, W., Li, C., Li, Z., Fu, P., Hu, X., 2013. Modeling of passing events in mixed bicycle traffic with cellular automata. Transport. Res. Rec. 2387 (1), 26–34. doi:10.3141/2387-04.
Zhao, Y., Zhang, H.M., 2017. A unified follow-the-leader model for vehicle, bicycle and pedestrian traffic. Transport. Res. Part B 105, 315–327. doi:10.1016/j.trb.2017.09.004.
Zhou, D., Jin, S., Ma, D., Wang, D., 2015. Modeling mixed bicycle traffic flow: a comparative study on the cellular automata approach. Discrete Dyn. Nat. Soc. 2015, 1–11. doi:10.1155/2015/420581.
Zhou, D., Xu, C., 2015. Estimating capacity of bicycle path on urban roads in Hangzhou, China. In: Proceedings of the 94th Annual Meeting of the Transportation Research Board. Yantai, China. https://trid.trb.org/view/1337236.

5 Paper 4: Large-scale assignment of congested bicycle traffic using speed heterogeneous agents

The following pages contain the article:

M. Paulsen and K. Nagel (2019). "Large-Scale Assignment of Congested Bicycle Traffic Using Speed Heterogeneous Agents". In: *Procedia Computer Science* 151, pp. 820–825. DOI: 10.1016/j.procs.2019.04.112. URL: https://linkinghub.elsevier.com/retrieve/pii/S1877050919305769.

Please cite accordingly.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 151 (2019) 820–825

Procedia Computer Science

www.elsevier.com/locate/procedia

The 8th International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications (ABMTRANS 2019), April 29 - May 2, 2019, Leuven, Belgium

Large-Scale Assignment of Congested Bicycle Traffic Using Speed Heterogeneous Agents

Mads Paulsen[a,*], Kai Nagel[b]

[a]*Technical University of Denmark, Department of Management Engineering, Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark*
[b]*Technische Universität Berlin, Transport Systems Planning and Telematics, Salzufer 17-19, 10587 Berlin, Germany*

Abstract

Despite requiring less space than most other modes of transport, bicycle traffic will also be prone to congestion when the traffic volume is sufficiently large. Such congestion can eventually influence the route choices of cyclists using the network. In this study we model bicycle congestion on a detailed network of the greater Copenhagen area by assigning an entire day of bicycle traffic using a recently developed method for dynamic network loading of speed heterogeneous multi-lane bicycle traffic. The model iteratively assigns appropriate routes for more than a million bicycle trips in the demand sensitive network, and with computation times of less than 15 minutes per iteration the proposed model proves to be large-scale applicable. This makes it the first dedicated bicycle traffic assignment model to account for congestion. The results indicate that the solid bicycle infrastructure of Copenhagen and cyclists' willingness to change routes are key to keeping travel times low for cyclists.

© 2019 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the Conference Program Chairs.

Keywords: Bicycle traffic assignment, bicycle congestion modelling, multi-agent simulation, speed heterogeneity

1. Introduction

Although congestion on bicycle tracks may seem unlikely many places in the world, in some cities (e.g. Copenhagen) the traffic level of bicycles has reached a level which leads to considerable bicycle congestion. The congestion influences the travel time, and may also influence cyclists' route choice.

For decades the literature[1] has been concerned with how the level-of-service for cyclists is influenced by the car traffic volume. However, including the flow of bicycles as a parameter of attractiveness of a route has only received little attention in the literature. Two stated preferences studies[2,3] found that cyclists generally have disutility towards increasing bicycle flow on bicycle lanes in Nanjing, China. This indicates that the more cyclists using a link, the less attractive the link will become. In traffic assignment for motorised traffic such disutility can be explained by

* Corresponding author. Tel.: +45 45 25 65 98.
E-mail address: madsp@dtu.dk

1877-0509 © 2019 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the Conference Program Chairs.
10.1016/j.procs.2019.04.112

increased travel time modelled with a flow-dependent travel time function for each link[4]. However, for bicycle traffic assignment only a single model[5] has so far incorporated a mechanism that makes links with high flows less attractive.

Primarily dealing with determining the static link characteristics influencing bicycle route choice, a wide range of dedicated bicycle route choice models exist[6] – some of which also perform one-shot assignments of bicycle flows[7,8]. Other studies[9,10,11,12,13] have focused on assigning bicycles traffic onto networks including various link characteristics, but without any feedback from the network such as lowered safety or speeds influencing the route choice, although one of the studies[12] mentions it as a necessary extension. Another study[14] solely dealing with route choice also argues – as a suggestion for future research – that bicycle route choice models ought to incorporate congestion effects from traffic assignment models with dedicated bicycle volume-delay functions.

Nevertheless, until now the literature only contains one example[5] of such bicycle traffic assignment considering congestion effects. That study focuses on mixed traffic in Patna, India and assigns mixed traffic onto a multi-modal, shared network whilst taking the maximum speeds of trucks, cars, motorbikes, and bicycles into account. However, due to the chaotic Indian traffic scheme with mixed traffic being far from what is seen in a city with plenty of dedicated bicycle infrastructure, there still is a need for a dedicated bicycle traffic assignment model that can model congestion in segregated bicycle traffic.

The aim of this study is to fill this gap by assigning bicycle traffic onto a network at a large scale using an iterative feedback loop between the supply and the demand, as it is commonly known from motorised traffic. Based on a recently developed network loading model for heterogeneous bicycle traffic[15] this study includes detailed congestion patterns on link level, and forms a complete agent-based traffic assignment model for bicycles by allowing every cyclist to adjust his/her route according to the congested network.

The study does not deal with travel time delays caused by conflicting traffic at intersections nor traffic signals. We acknowledge that such delays cause substantial added travel time for cyclists. However, in order to keep focus on the large-scale applicability and consequences of the implemented network loading model such delays have been considered to be beyond the scope of this paper.

The remainder of the paper is structured as follows. Section 2 contains the applied methodology including a brief summary of the used network loading model[15]. The data used for the case study is described in Section 3, before presenting the results of said case study in Section 4. Finally, a discussion of the realism and the limitations of the model is found in Section 5 alongside suggestions for future research.

2. Methodology

In order to have an assignment model for bicycle traffic, it is necessary to be able to model how the demand influences the supply and vice versa. Plenty of route choice models for bicycle traffic exist[6], meaning that it is known to a large extend how the supply influences the demand for any link in the network. Until recently, however, it has been ignored how the supply is influenced by the demand predicted by such route choice models.

The network loading used in the assignment model of this study is a MATSim[16]-implementation of a recently proposed methodology[15] for dynamic, large-scale applicable network loading of bicycle traffic. The model is based on individual desired speeds for every cyclist corresponding to an individualised free-flow speed. It supports multi-lane links with each link having a predefined number of pseudo-lanes, allowing faster cyclists to overtake slower cyclists by placing themselves in the outermost lane(s) as long as there is room. If a cyclist can no longer choose a pseudo-lane that will satisfy his/her desired speed, the cyclist will choose the fastest available pseudo-lane. Furthermore, each cyclist has an individualised, speed dependent headway distance, that he/she must keep to the cyclist in front.

The speed of a cyclist entering a lane can be determined at link entrance solely based on his/her desired speed and headway preferences, as well as the time at which the previous entrant of the selected pseudo-lane is going to be leaving and entered the link, as neither overtaking nor lane changing are allowed within a lane. Because delays are based on the lack of opportunity to overtake, it also means that cyclists with high desired speeds have a higher tendency to be delayed, whereas it is virtually impossible for the overall slowest cyclist to be delayed. The methodology yields reasonable aggregate fundamental diagrams[15] as long as there is sufficient speed heterogeneity, and as such is best suited for agent-based transport modelling where individualised preferences can be assigned directly to each agent.

The route choice model used in this study is a simple multinomial logit model where the free flow travel time and congested travel time of alternatives are the only two variables used in the disutility function. The congested travel

time, i.e. the difference between the actual travel time and the free flow travel time, is penalised 50 % harder than free flow travel time, in line with previous findings for car users in the area[17].

Each agent has a maximum number of five plans in his/her choice set, and after each iteration the performed plan is assigned a score based on the obtained travel time in the previous iteration. New plans are added to the choice set by performing a shortest path search in a network with empirical average travel times in time bins of 15 minutes based on the previous iteration. However, during the shortest path search the link cost of every link is agent-specific by using the maximum between the empirical (time binned) travel time and the free flow travel time based on that agent's desired speed.

Although the framework allows extension to include other link (to link) attributes previously seen in the literature such as signals, slope, surface, and left/right turns, for the time being only travel time is used in order to focus on the congestion caused by the network loading model.

The traffic assignment model has been implemented in the open-source and agent-based transport simulation software MATSim[16] by replacing parts of the default mobility simulator with the model from Paulsen et al. (2018)[15], alongside minor changes in the routing in order to account for individual desired speeds.

3. Case Study

In order to test the realism and the large-scale applicability of the model a case study for the greater Copenhagen area is carried out. The demand is based on the Copenhagen Model for Person Activity Scheduling (COMPAS)[18], producing daily activity plans for a synthetic population of the area. Only persons with a bicycle trip during his/her day have been used for this study, resulting in a population of 547,085 persons with daily activity plans forming a total of 1,082,958 bicycle trips.

Based on aerial data of observed, uncongested bicycle traffic on Smallegade in Frederiksberg, every individual is assigned a normally distributed headway distance parameter and a desired speed (v_0 in m/s) based on Johnson's S_U distribution[19] with the estimated parameters $\gamma = -2.75$, $\xi = 3.67$, $\delta = 4.07$, and $\lambda = 3.49$,[15]

$$v_0 \sim \frac{\delta}{\lambda\sqrt{2\pi}\sqrt{1+\left(\frac{v_0-\xi}{\lambda}\right)^2}}\, e^{-\frac{1}{2}\left(\gamma+\delta\sinh^{-1}\left(\frac{v_0-\xi}{\lambda}\right)\right)^2}, \tag{1}$$

inferring a mean speed and standard deviation of roughly 22 km/h and 4 km/h, respectively. This distribution has the highest likelihood and lowest Kolmogorov-Snirnov statistic among 11 candidate distributions with estimated parameters, and does not deviate significantly from the empirical distribution from Smallegade according to the Kolmogorov-Smirnov Goodness-of-Fit test (p-value above 0.6).[15]

The network is based on OpenStreetMap (OSM)[20] and is included in MATSim by altering the default OSM network reader[21]. All link types that generally allow bicycle traffic have been included in the model unless the road explicitly stated that bicycle traffic was not allowed. Furthermore, any road where bicycle traffic is explicitly mentioned as allowed or designated has also been included in the network.

The network loading model[15] only needs the lengths and widths of the bicycle infrastructure of links. Lengths are always available and can always be extracted directly from the OSM data. Widths also play a vital part for the methodology, but are rarely available. On the few links where widths (ω^l) were available, the number of lanes were determined based on the widths according to a Danish study[22], that found the number of efficient lanes (Ψ^l) of bicycle traffic to be determined by,

$$\Psi^l = 1 + \left\lfloor \frac{\omega^l - 0.4\,m}{1.25\,m} \right\rfloor. \tag{2}$$

Widths were manually added to crucial arterial roads of Copenhagen – especially the *cycle superhighways* – before extracting the data from OSM. This was done in order to secure these links being capable of handing large amount of bicycle traffic. Where widths were not available, the type of bicycle infrastructure determined the number of lanes. Explicit cycleways where motorised traffic is not allowed, as well as roads with bicycles riding in lanes or tracks were given two pseudo-lanes, since practically all of such infrastructure in Denmark would be wide enough to form two efficient lanes. Roads with no information about cycleways and footpaths where bicycle traffic is allowed were only given a single lane, as the presence of cars and pedestrians in such cases will limit the possibility to overtake.

Table 1. Average travel time and congested time according to five different scenarios.

Scenario	Avg. Travel Time Per Trip	Avg. Congested Travel Time Per Trip
Unlimited Capacity (Desired Speed)	16.43 min	0.00 min
Actual Infrastructure (it. 0)	16.73 min	0.29 min
Actual Infrastructure (it. 200)	16.67 min	0.12 min
Single-Lane Infrastructure (it. 0)	18.76 min	2.33 min
Single-Lane Infrastructure (it. 200)	17.70 min	0.81 min

4. Results

The assignment model was run for the population and network described in Section 3 for a total of 200 iterations with fixed choice sets after 160 iterations. The average computation time per iteration was 14.9 minutes on a single node of a high performance computer with two 2.8 GHz deca-core processors with 120 GB RAM.

Table 1 shows the average travel time and congested travel time for five different scenarios. It is seen that in the very first iteration, corresponding to no one adapting their routes according to the network performance, using the actual infrastructure a cyclist would on an average get his/her average trip travel time prolonged by approximately 18 seconds of congested travel corresponding to just under 2 % of the total travel time. Due to the possibility to choose alternative routes, this number is eventually reduced so that the average congested travel time is 7 seconds (roughly 1 %) while lowering the overall travel time by a similar amount.

This may be interpreted as bicycle congestion being a non-issue. On the other hand, it can also be seen as an indication of how well the profound bicycle infrastructure found in Copenhagen is relieving congestion. This is supported by the two final rows of Table 1, which are based on model runs where all links only have a single pseudo-lane, thus radically reducing the possibility to overtake. With such an infrastructural setup, using distance based shortest path for every cyclist results in a travel time that is prolonged by an average of 2.03 minutes per trip (12 %). In such a scenario large benefits can be made by adapting the route, and eventually the excess travel time can be reduced to 1.03 minutes (6.2 %) with 48 seconds of the travel time being congested time, more than six times as much than with the actual infrastructure. It is worth noting that these numbers are average values across an area larger than the entire Capital Region of Denmark which also covers a lot of smaller towns and rural areas where bicycle congestion is practically non-existing. This means that the congested travel time of trips in the central parts of Copenhagen will be considerably higher than the average values across the entire area presented in the table.

In order to investigate how the route choice is affected by congestion in the central part of Copenhagen, a difference map based on the morning peak hour from 7am to 8am between the actual congested network and a free-flow assignment is presented in in Figure 1. At first it might be hard to see any clear patterns in the map, but including some local knowledge lets the map be understood more easily. In the south-eastern part of the map it is seen that a substantial amount of cyclists are avoiding Knippelsbro. More than 100 cyclists per direction deviate onto the Inner Harbour bridge, and for the busy north-western direction the route across Langebro also becomes a reasonable alternative. This indeed seems plausible as Torvegade (the street linking to Knippelsbro from south-east) has a relatively moderate bicycle infrastructure that only allows two efficient lanes per direction, although having one of the highest daily bicycle counts. In the other end of the city centre, it is seen that Queen Louise's Bridge is gaining cyclists due to congestion from competing bridges. This seems intuitive as this bridge and Nørrebrogade (the street linking to Queen Louise's Bridge from north-west) forms a corridor that predominantly have three efficient lanes following a radical renovation in 2011, and is considered as the main arterial road for bicycle traffic connecting westbound suburbs.

5. Conclusions & Future Work

In this study a traffic assignment model for bicycle traffic was proposed using a detailed, demand sensitive network loading model[15]. A large-scale case study was conducted for the greater Copenhagen area with running times proving that the presented methodology is in fact large-scale applicable.

The results have shown that based on the proposed model, congestion related route changes does seem to happen in Copenhagen with some cyclists avoiding the most congested links in favour of links with particularly wide bicycle

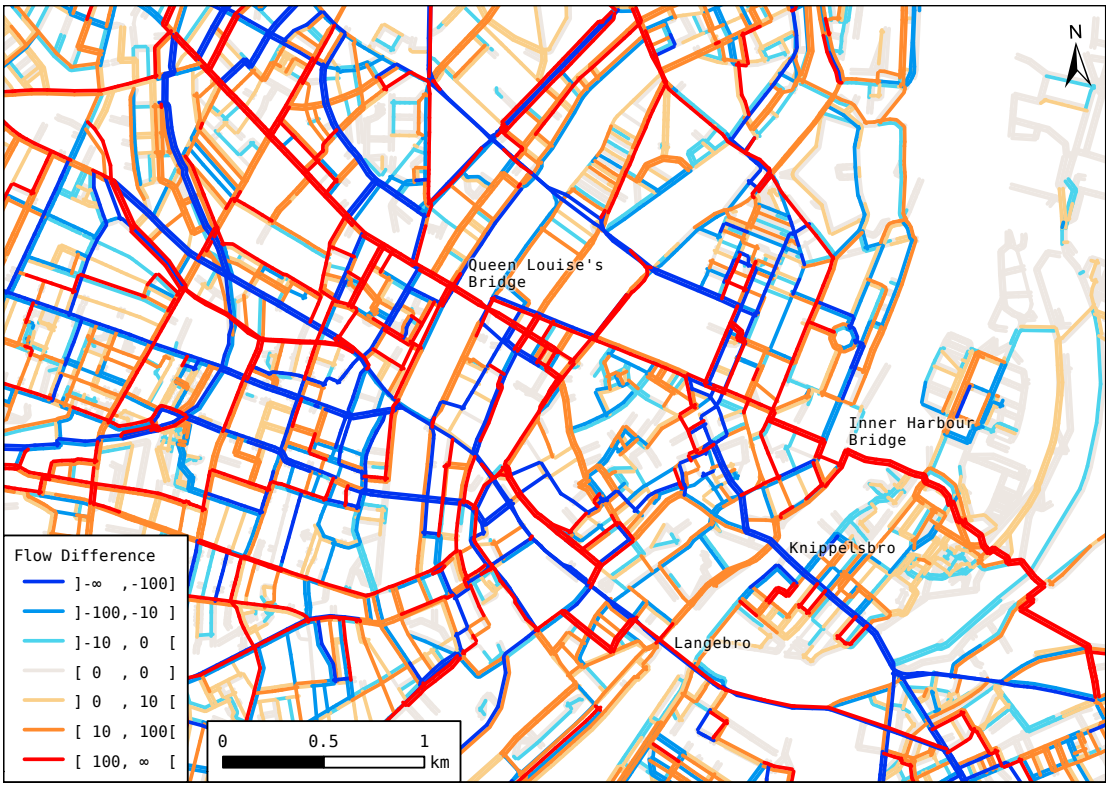


Fig. 1. Map with selected points of interests showing the difference in flow in the morning peak hour from 7 to 8 between modelled selected routes and routes found by using a free-flow, distance-based assignment.

tracks. The results also showed that having a comprehensive, decent bicycle infrastructure can greatly reduce the travel time of cyclists. However, this effect is likely overestimated with the current study as the single-lane scenario implies no opportunity to overtake altogether. In reality cyclists do tend to find a way of overtaking, why the predicted travel times will probably be pessimistic for the fastest cyclists.

Even though the travel time gains are possibly overestimated, the study only includes travel time as a parameter. In reality, better infrastructure is inherently attractive for cyclists [1], and will also improve the perceived safety of cyclists. Parameters that this study do not deal with explicitly, although it would be interesting to include in future research.

Furthermore, the methodology applied in this study only deals with delays on the links themselves. However, the majority of excess travel time for cyclists is possibly induced at intersections. In the morning peak hour certain core links in Copenhagen have bicycle queues that are often longer than what can be emptied in a single signal cycle. Delays can also happen when cyclists are waiting for through-going traffic when making a left turn.

Such delays can be included directly in the mobility simulation by concurrently simulating motor vehicles and modelling right of way and signals at intersections. This would also capture how cyclists delay the remainder of traffic, meaning that it would be very suitable for project appraisal purposes. Although relevant, implementing such methodology would be a serious extension and possibly be on the limit on what is computationally feasible.

A pragmatic alternative could possibly be to include the number of turns and signals as parameters in the route choice model [23] and/or to apply fixed travel time penalties for every signal, left turn, and right turn, respectively. The literature on bicycle route choice modelling also contains additional link-specific parameters such as land-use [14], surface [14], and slope [24], which straightforwardly can and should be added to the model in future work, although slope is not a huge issue in the greater Copenhagen area.

References

1. Antonakos. Environmental and travel preferences of cyclists. *Transportation Research Record* 1994;(1438):25–33. URL: https://trid.trb.org/view/413764.
2. Li, Wang, Liu, Ragland. Physical environments influencing bicyclists' perception of comfort on separated and on-street bicycle facilities. *Transportation Research Part D: Transport and Environment* 2012;**17**(3):256–261. URL: http://linkinghub.elsevier.com/retrieve/pii/S1361920911001556. doi:10.1016/j.trd.2011.12.001.
3. Bai, Liu, Chan, Li. Estimating level of service of mid-block bicycle lanes considering mixed traffic flow. *Transportation Research Part A: Policy and Practice* 2017;**101**:203–217. URL: https://linkinghub.elsevier.com/retrieve/pii/S0965856416300374. doi:10.1016/j.tra.2017.04.031.
4. Sheffi. *Urban transportation networks: Equilibrium analysis with mathematical programming methods*. 1. Englewood Cliffs, NJ, USA: Prentice-Hall; 1985. ISBN 0139397299. URL: http://web.mit.edu/sheffi/www/selectedMedia/sheffi_urban_trans_networks.pdf.
5. Agarwal, Ziemke, Nagel. Bicycle superhighway: An environmentally sustainable policy for urban transport. VSP Working Paper 17-16, TU Berlin, Transport Systems Planning and Transport Telematics; 2017. URL: https://www.researchgate.net/publication/320136429_Bicycle_superhighway_An_environmentally_sustainable_policy_for_urban_transport.
6. Pritchard. Revealed Preference Methods for Studying Bicycle Route Choice—A Systematic Review. *International Journal of Environmental Research and Public Health* 2018;**15**(3):470. URL: http://www.mdpi.com/1660-4601/15/3/470. doi:10.3390/ijerph15030470.
7. Hood, Sall, Charlton. A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters* 2011;**3**(1):63–75. URL: http://www.tandfonline.com/doi/full/10.3328/TL.2011.03.01.63-75. doi:10.3328/TL.2011.03.01.63-75.
8. Zimmermann, Mai, Frejinger. Bike route choice modeling using GPS data without choice sets of paths. *Transportation Research Part C: Emerging Technologies* 2017;**75**:183–196. URL: https://linkinghub.elsevier.com/retrieve/pii/S0968090X16302637. doi:10.1016/j.trc.2016.12.009.
9. Ehrgott, Wang, Raith, van Houtte. A bi-objective cyclist route choice model. *Transportation Research Part A: Policy and Practice* 2012;**46**(4):652–663. URL: http://linkinghub.elsevier.com/retrieve/pii/S0965856411001844. doi:10.1016/j.tra.2011.11.015.
10. Jacyna, Wasiak, Kłodawski, Gołębiowski. Modelling of Bicycle Traffic in the Cities Using VISUM. *Procedia Engineering* 2017;**187**:435–441. URL: https://www.sciencedirect.com/science/article/pii/S1877705817319276. doi:10.1016/j.proeng.2017.04.397.
11. Ziemke, Metzler, Nagel. Modeling bicycle traffic in an agent-based transport simulation. *Procedia Computer Science* 2017;**109**(00):923–928. URL: https://linkinghub.elsevier.com/retrieve/pii/S1877050917311043. doi:10.1016/j.procs.2017.05.424.
12. Ryu, Chen, Su, Choi. Two-Stage Bicycle Traffic Assignment Model. *Journal of Transportation Engineering, Part A: Systems* 2018;**144**(2):04017079. URL: http://ascelibrary.org/doi/10.1061/JTEPBS.0000108. doi:10.1061/JTEPBS.0000108.
13. Ziemke, Metzler, Nagel. Bicycle traffic and its interaction with motorized traffic in an agent-based transport simulation framework. *Future Generation Computer Systems* 2018;URL: https://linkinghub.elsevier.com/retrieve/pii/S0167739X17320447. doi:10.1016/j.future.2018.11.005.
14. Prato, Halldórsdóttir, Nielsen. Evaluation of land-use and transport network effects on cyclists' route choices in the Copenhagen Region in value-of-distance space. *International Journal of Sustainable Transportation* 2018;:1–12URL: https://www.tandfonline.com/doi/full/10.1080/15568318.2018.1437236. doi:10.1080/15568318.2018.1437236.
15. Paulsen, Rasmussen, Nielsen. Fast or Forced to Follow: A Speed-Heterogeneous Approach to Congested Multi-Lane Bicycle Traffic Simulation. Manuscript submitted for publication; 2018.
16. Horni, Nagel, Axhausen, editors. *The Multi-Agent Transport Simulation MATSim*. London: Ubiquity Press; 2016. URL: https://www.ubiquitypress.com/site/books/10.5334/baw/. doi:10.5334/baw.
17. Prato, Rasmussen, Nielsen. Estimating Value of Congestion and of Reliability from Observation of Route Choice Behavior of Car Drivers. *Transportation Research Record: Journal of the Transportation Research Board* 2014;**2412**(1):20–27. URL: http://journals.sagepub.com/doi/10.3141/2412-03. doi:10.3141/2412-03.
18. Prato, Rasmussen, Nielsen, Watling. A disaggregate pseudo-dynamic assignment for the activity-based model of the Greater Copenhagen Area. In: Joao, editor. *13th World Conference on Transport Research (WCTR)*. Rio de Janeiro, Brazil: Federal University of Rio de Janeiro; 2013, p. 1–19.
19. Johnson. Systems of Frequency Curves Generated by Methods of Translation. *Biometrika* 1949;**36**(1/2):149–176. URL: https://www.jstor.org/stable/2332539?origin=crossref. doi:10.2307/2332539.
20. OpenStreetMap. Accessed 13 December 2018. URL: https://www.openstreetmap.org.
21. Zilske, Neumann, Nagel. OpenStreetMap For Traffic Simulation. In: Schmidt, Gartner, editors. *Proceedings of the 1st European State of the Map – OpenStreetMap conference, no. 11-10*. 2011, p. 126–134.
22. Buch, Greibe. Analysis of Bicycle Traffic on One-Way Bicycle Tracks of Different Width. In: *European Transport Conference 2015*. Frankfurt, Germany: Association for European Transport (AET); 2015, URL: https://aetransport.org/en-gb/past-etc-papers/conference-papers-2015?abstractId=4418&state=b.
23. Broach, Dill, Gliebe. Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice* 2012;**46**(10):1730–1740. URL: http://linkinghub.elsevier.com/retrieve/pii/S0965856412001164. doi:10.1016/j.tra.2012.07.005.
24. Menghini, Carrasco, Schüssler, Axhausen. Route choice of cyclists in Zurich. *Transportation Research Part A: Policy and Practice* 2010;**44**(9):754–765. URL: http://linkinghub.elsevier.com/retrieve/pii/S0965856410001187. doi:10.1016/j.tra.2010.07.008.

6 Paper 5: Including right-of-way in a joint large-scale agent-based dynamic traffic assignment model for car and bicycle traffic

The following pages contain the article:

M. Paulsen, T. K. Rasmussen, and O. A. Nielsen (2020). "Including right-of-way in a joint large-scale agent-based dynamic traffic assignment model for car and bicycle traffic". Manuscript submitted to Networks and Spatial Economics.

Please cite accordingly.

Networks and Spatial Economics manuscript No.
(will be inserted by the editor)

Including Right-of-Way in a Joint Large-Scale Agent-Based Dynamic Traffic Assignment Model for Cars and Bicycles

Mads Paulsen[1] · Thomas Kjær Rasmussen[1] · Otto Anker Nielsen[1]

Received: / Accepted:

Abstract Intersections typically account for a substantial part of the total travel time in urban areas, and an even higher share of the congested travel time, especially for bicycle traffic. Nevertheless, delays caused by yielding for cyclists or cars at intersections have previously not been modelled in large-scale bicycle traffic assignment models. This study proposes a computationally efficient large-scale applicable methodology for explicitly modelling yielding for conflicting movements at multi-modal intersections in an agent-based traffic assignment model. Nodes representing the intersections are classified into five different node types which simulate potential movements across nodes differently while obeying right-of-way and preventing simultaneous conflicting movements. The methodology is implemented within a joint assignment model capable of modelling on-link congestion of both car and bicycle traffic, and is applied to a large-scale case study of a Metropolitan Copenhagen network with 144,060 nodes and 572,935 links. The MATSim case study with 4,593,059 trips shows manageable computation times similar to when not modelling right-of-way at intersections. Especially for car traffic, yielding at intersections imposes considerable excess travel time. The effects are larger for trips going to the central part of the city where the inter-modal impact of conflicting bicycle traffic is identified as the dominant source of added travel time. The study finds that failing to model conflicting moves at intersections generally underestimates travel times and causes too much traffic to go through the urban core, highlighting the importance of joint modelling of intersections.

M. Paulsen
E-mail: madsp@dtu.dk
ORCiD: 0000-0002-1445-0876

T.K. Rasmussen
E-mail: tkra@dtu.dk

O.A. Nielsen
E-mail: oani@dtu.dk

[1]Department of Management, Technoology and Economics, Technical University of Denmark
Bygningstorvet 116, 2800 Kgs. Lyngby, Denmark

Manuscript submitted to Networks and Spatial Economics

Keywords Intersection modelling · Multi-modal traffic · Bicycle traffic · Large-scale traffic assignment · Agent-based simulation

1 Introduction

One of the most important purposes of traffic assignment models is to give a realistic representation of congestion for all relevant modes. As such, in urban contexts it is relevant to include bicycle traffic, but due to fundamental differences between car traffic and bicycle traffic, the underlying methodologies to model congestion for these modes may differ (Paulsen et al., 2019). In cities with a high level of segregation between modes, this may be dealt with by implementing separate methods for modelling on-link travel times for car and bicycle traffic, respectively. However, a joint model including both car and bicycle traffic is necessary when modelling intersections, as using two separate models would ignore the considerable interactions occurring between the two types of traffic due to yielding and right-of-way. Proposing such a model and integrating it in a joint traffic assignment model for otherwise separated car and bicycle traffic is the purpose of this paper.

In the past two decades, dynamic traffic assignment has been an important topic in research (Peeta and Ziliaskopoulos, 2001; Chiu et al., 2011). However, until recently neither static nor dynamic large-scale traffic assignment models have been able to model congestion for dedicated bicycle traffic realistically. A few studies have modelled congestion for bicycle traffic, but through congestion effects arising when sharing road infrastructure with motorised transport modes along links. Agarwal et al. (2019) modelled congestion by agent-based simulation of mixed traffic in the Indian city of Patna. Wierbos et al. (2019) proposed a flow-based approach for modelling mixed car and bicycle traffic suitable for integration in traffic assignment models, however without doing so. Furthermore, neither methodologies are suitable for cities with a high level of segregation between bicycle and car traffic as they do not model congestion on dedicated bicycle infrastructure.

Paulsen et al. (2019) introduced an agent-based network loading model for large-scale modelling of congestion on bicycle paths based on speed heterogeneity. Later, Paulsen and Nagel (2019) integrated the methodology in a full traffic assignment model with feedback mechanisms between demand and supply, and applied it to a large-scale bicycle network of Metropolitan Copenhagen. However, without modelling excess travel time arising from interactions between vehicles at intersections.

Such delays at intersections typically constitute a considerable amount of the total travel time in the network, especially for bicycle traffic where delays on the links themselves are limited due to good opportunities to overtake. Modelling intersections is particularly relevant in urban areas with a high concentration of bicycle traffic, as the presence of bicycles does not only cause delays for other bicycles, but also for turning car traffic. Likewise, car traffic also delays cyclists having to cross larger roads. In cities with a high level of segregation between modes, inter-modal delays can be ignored on links as the two traffic streams are completely separated, but are still important to model at intersections in order to obtain realistic travel times for both modes.

Microsimulation tools such as AIMSUN (Dandl et al., 2017), SUMO (Krajzewicz et al., 2014), and VISSIM (Fellendorf, 1994) are capable of modelling

delays at intersections for both car and bicycle traffic in great detail, but are not large-scale applicable.

Large-scale flow-based car traffic assignment models have in the last twenty or so years often included some form of intersection delays (Nielsen et al., 1998a). Such models can indirectly include the effect of having to yield for bicycle traffic, by reducing the capacity of relevant turn movements by applying relevant adjustment factors. Estimating such adjustment factors for bicycle traffic have received quite some attention in the literature (Allen et al., 1998; Brilon and Miltner, 2005; Chen et al., 2007; Li et al., 2009, 2011; Guo et al., 2012; Chen et al., 2014; Preethi and Ashalatha, 2018), and is a common way to include car travel time delays at intersections caused by prioritised traffic in flow-based models. Similar adjustments can be done for other types non-motorised traffic (Mondal and Gupta, 2020) such as pedestrians (Niittymäki and Pursula, 1997; Milazzo et al., 1998; Chen et al., 2008, 2015; Roshani and Bargegol, 2017).

Although such methods can be used for modelling car traffic delays caused by bicycle traffic, they are not suitable for modelling the opposite case where cyclists have to yield for car traffic, as the large heterogeneity among cyclists makes flow-based approaches inappropriate (Paulsen et al., 2019). As such, using flow-based methods is not a viable option when wanting to perform realistic traffic assignment for both car and bicycle traffic while modelling their interaction at intersections. The only existing large-scale applicable network loading model capable of modelling congestion on dedicated bicycle infrastructure, is the agent-based methodology from Paulsen et al. (2019) applied in Paulsen and Nagel (2019). As such, in order to be able to model on-link congestion, any proposed intersection model that can capture the delays caused by yielding at intersections need also to be agent-based for the two to be compatible.

This study contribute to literature by modelling delays caused by conflicting movements at intersections directly in the mobility simulation of a large-scale agent-based traffic assignment model for joint car and bicycle traffic. Modelling such conflicts is not new in itself, e.g. Dandl et al. (2017), neither is it new to model an entire metropolitan area in an agent-based transport simulation setup, e.g. Raney et al. (2003). But doing both simultaneously, i.e. modelling conflicts at intersections while also simulating and modelling on-link congestion as well as route choices and traffic assignment of both bicycle and car traffic of an entire metropolitan area is. The study does this by extending the open-source agent-based transport simulator MATSim (Horni et al., 2016). As MATSim is already capable of simulating a large geographical area in feasible time, the specific objective of this study is to replace the existing, simplistic intersection model (node model) of MATSim with a detailed one obeying multi-modal right-of-way at intersections, and apply it to large-scale case study with a large proportion of bicycle traffic.

The remainder of the paper is structured as follows. Section 2 describes how the nodes of network are classified into five different types, whereas Section 3 presents how vehicles and bicycles are modelled when travelling across such nodes. A case study of Metropolitan Copenhagen is presented in Section 4 alongside results. Section 5 discusses the findings and outlines directions for future research. Finally, the conclusions of the study is summarised in Section 6.

2 Node classification

The node model that will be presented in Section 3, distinguishes between different node types, see Figure 1. Therefore, in this section the method for classifying each node of the network into each of the five node types is presented. In line with earlier studies such as Nielsen et al. (1998b) this process is fully automated, and is carried out in the pre-processing stage before the actual simulation.

Instead of basing the node types on each individual link of the node, this method is based on link *bundles* in order to make the model less dependent on the exact network digitalisation. A bundle, D, is defined as a direction from/to the node which consists of up to four unidirectional links, but at most one of the following listed in counterclockwise order: Bicycle out-link, car out-link, car in-link, and bicycle in-link. If there is more than 15° between two consecutive links in a bundle, the bundle is split into multiple bundles.

In the implementation of the present paper, we base our network on OpenStreetMap (OpenStreetMap, 2020). The *size* of a bundle is based on the link of the bundle with the largest OpenStreetMap *highway type* with the number of lanes of this link as a potential tie breaker. However, these could as well have been based on any (combination of) ordinal attribute(s) that could be a proxy for the size of the road. Alternatively, a value could be assigned manually to (specific) links of the network if needed. Notice that in our implementation, we have not distinguished between highway types with a "_link" suffix ("motorway_link", "tertiary_link", etc.) and their counterparts without such suffix ("motorway", "tertiary", etc.) when determining priorities nor when categorising the nodes.

By defining a *bundle path* as a possibly bi-directional path connecting two different bundles across a node having their size determined by the smallest of the two bundles, we formally define the five node types used in the study:

- *Merging nodes:* Two large one-way car in-links merge into a single large one-way car out-link.
- *Diverging nodes:* A single large one-way car in-link diverges into two (or more) large one-way car out-links.
- *Right priority nodes:* All bundle paths have the same size.
- *Directed priority nodes:* The size of a single bundle path dominates the size of all other bundle paths.
- *Anti-priority nodes:* Bundle path sizes differ, but there is no single dominant bundle path.

We again refer to Figure 1 for examples of each of the node types. A more detailed description of each of the node types and how their categorisation is determined is presented in the following paragraphs. Some node types distinguish between *primary* and *secondary* bundles. How these are determined is also included in the following when needed.

A requirement for being a *merging* or *diverging node*, respectively, is that the smallest link leading to/from the node is of highway type "primary_link" or higher, and that it only has a single out-link (merging node) or in-link (diverging node). This means that they are almost exclusively used in conjunction with motorways, where merging nodes often represent ramps leading into the motorway and diverging nodes often represent ramps leading away from the motorway. For merging nodes the primary bundle is defined as the bundle including the largest of the two

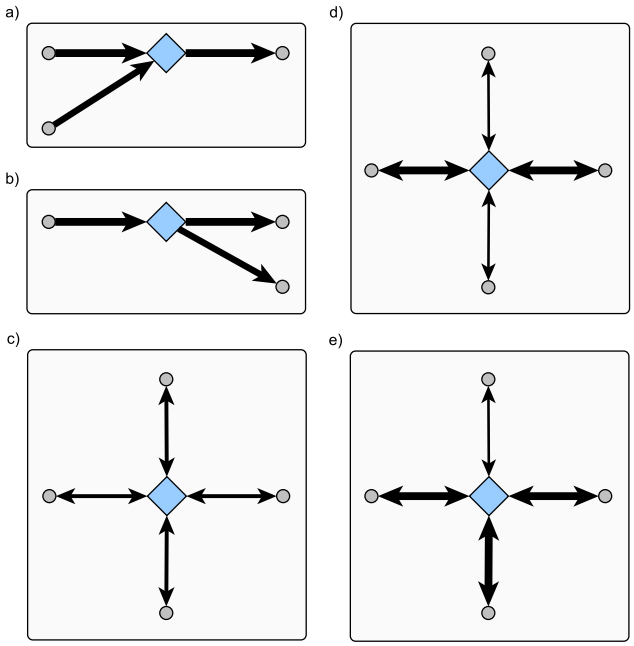


Fig. 1 Examples of each of the node types: a) Merging node, b) Diverging node, c) Right priority node, d) Directed priority node, and e) Anti-priority node. Each line represents a *bundle* with *size* proportional to the width of the line, and with arrowheads indicating the possible travel directions within the bundle.

links based on highway type and number of lanes. If these are tied, the primary bundle is selected arbitrarily among the bundles containing in-links. For diverging nodes there is only one in-link and there is no distinction between the out-links as none of the possible movements are conflicting.

Nodes that are neither merging nor diverging nodes and where all bundles have the same size are categorised as *right priority nodes*. The vast majority of these nodes occur when residential roads cross each other, but are also used when larger roads of equal size intersect. A special case are dead-end nodes with only one bundle – these are also categorised as right priority nodes.

If two bundles are larger than all the other bundles and the node is not a merging nor a diverging node, then the node is a *directed priority node*. The two bundles of the dominant bundle path are declared as *primary bundles*. Consequently, all other bundles are *secondary bundles*. A special case is nodes with exactly two bundles. These are always categorised as directed priority nodes (with no secondary bundles).

If none of the above criteria have been fulfilled, it means that at least two bundle sizes differ but that the node has no single dominant bundle path, and it is

neither a merging nor a diverging node. For these cases the *anti-priority* nodes are used. Although the anti-priority node is much like a right priority node it also has some elements of the directed priority node, as bundles are divided into priority bundles and secondary bundles. The largest bundle and all other bundles sharing the size of the largest bundle constitute the set of primary bundles, whereas the remaining bundles are secondary bundles. The anti-priority nodes are often seen in cases where removing the minimum bundle would cause the node to become a right priority node.

3 Simulation model

In this section we will present the right-of-way node model developed for this study for implementation in MATSim (Horni et al., 2016). The reason for implementing the methodology in MATSim is threefold. Firstly, because the software is open source, secondly it is widely used for a wide range of scenarios across the world, and thirdly because the methodology for simulating bicycle traffic (Paulsen et al., 2019) has previously been implemented in MATSim (Paulsen and Nagel, 2019).

Since the original node model of MATSim (Flötteröd, 2016) is used as baseline in our study, we first introduce the original MATSim node model in Section 3.1, before presenting our proposed right-of-way model in Section 3.2.

3.1 The original MATSim node model

The mobility simulation in MATSim (Horni et al., 2016) consists of two separate models; a link model and a node model (Flötteröd, 2016). The link model determines at link entry at which time a vehicle will be ready to leave the link again, at which point the vehicle is moved onto the *buffer* of the link. The node model, on the other hand, is responsible for transferring vehicles from buffers and onto subsequent links.

In each time step (typically of $\tau = 1$ second) the node model process all links that has vehicles ready to move, $L \in \mathcal{L}_N^A$, of all active nodes, $N \in \mathcal{N}_A$. The processing order is chosen at random, but with high capacity links having a larger probability of being chosen first, as the probability of being chosen as the first (and subsequent) link to be processed is proportional to the flow capacity of the link, C_L^F. This means, that when going into a node from a link with a low capacity, chances are that other links will be processed first, often causing vehicles from low capacity links being placed behind vehicles from high capacity links on the following link.

Once the link is selected, all of its *buffers*, $B \in \mathcal{B}_L$, are processed. It is worth mentioning that although the basic setup of MATSim used in this study only has a single buffer per link, MATSim does offer configurations with explicit use of lanes (Grether, 2014; Grether and Thunig, 2016; Thunig et al., 2019), potentially increasing the number of buffers for some links. The processing order of buffers (links) is quite relevant in MATSim, as the node model offers no guarantee that a vehicle can be moved over to the desired link, as the receiving link has to have available space in order for this to happen. The storage capacity is simply defined as the area the link (measured in lane metres), found by multiplying the length

Algorithm 1 Original MATSim node model with spatial queue simulation

for all active nodes $N \in \mathcal{N}_a$ **do**
  Determine set of links remaining to be processed, $\mathcal{L}_N$.
  while $\mathcal{L}_N \neq \emptyset$ **do** *# Links left to process*
    Draw L^* stochastically from $\mathcal{L}_N$ with probability $\propto C^F_{L^*}$ and remove it.
    for $B \in L^*$ **do** *# All buffers of link*
      while $B \neq \emptyset$ **do** *# Buffer is not empty*
        $V \leftarrow B$.`getFirstVehicle()`
        $R_V \leftarrow V$.`getNextLink()` *# Get receiving link*
        # Receiving link accepts vehicle
        if R_V.`getSpareStorageCap()` $\vee\, t > T_B + \varsigma$ **then**
          B.`removeFirst()` *# Remove vehicle from buffer*
          $T_B \leftarrow t$ *# Update buffer time*
          R_V.`receive(`V`)` *# Move vehicle to receiving link*
        else *# Receiving link rejects vehicle*
          break *# Move on to next buffer*

of the link with its number of lanes. A standard car (one person car unit (pcu)) is assumed to take up 7.5 lane metres of space, which is subtracted/added to the spare storage capacity of the link when a vehicle enters/leaves the link. The area occupied by complex by cyclists is more complex and agent-specific, see Paulsen et al. (2019) for details.

When using more advanced link transmission models than the spatial queue model with link dependent back wave speed used here (Gawron, 1998), e.g. spatial queues with backward-travelling holes (Agarwal et al., 2018) or a Newell (Newell, 1993) kinematic wave representation (Flötteröd, 2016), additional restrictions may apply, but are not included in the present study.

Even when the restrictions for accepting a vehicle is not met by the receiving link, i.e. when R_V.`getSpareStorageCap()` ≤ 0, MATSim offers an alternative opening for allowing vehicles to enter the link anyway. A processed vehicle is allowed entry if it has been at the front of the buffer for more than a so-called "stuck time" (or "squeeze time" (Axhausen et al., 2015)) in this paper denoted by ς. This is included in MATSim in order to prevent persistent grid-lock situations to occur (Charypar, 2008; Rieser and Nagel, 2008; Axhausen et al., 2015). By denoting the time at which the vehicle became the front vehicle of the buffer by T_B, and the current simulation time as t, this can be formulated as $t > T_B + \varsigma$.

If neither of the conditions allowing the vehicle to move onto the receiving link are met, the vehicle and all behind vehicles in the same buffer have to wait (at least) until the next time step.

An algorithmic representation of the original node model with the most simple configurations for MATSim is given Algorithm 1.

The bicycle simulation methodology from Paulsen et al. (2019) for simulated bicycle traffic on dedicated bicycle infrastructure was implemented as an extension to MATSim in Paulsen and Nagel (2019) using separate links for bicycles. Since bicycle tracks are predominantly situated in conjunction with roads allowing car traffic, the car and bicycle infrastructure share the vast majority of their nodes. However, since only the sending links and receiving links are taken into account in

the original node model, there will be no inter-modal interactions at intersections as the sending and receiving links for bicycle traffic are completely disjoint.

As indicated, different versions of the node model can be achieved by using explicit lanes (Grether, 2014; Grether and Thunig, 2016; Thunig et al., 2019) or by using more advanced link transmission models (Agarwal et al., 2018; Flötteröd, 2016) than the spatial queue model used here. However, none of them makes it possible to take into account conflicting movements at unsignalised intersections. As such, although the original node model does have the potential to cause vehicle delays at intersections, it does not take the specific turn movements into account, meaning that there is no distinction between conflicting and non-conflicting movements.

3.2 The right-of-way node model

We now turn to present the right-of-way node mode formulated for this study. Section 3.2.1 presents the overall concept of the model, before proposed buffers are introduced in Section 3.2.2, and the order in which they are processed for the various node types are considered in Section 3.2.3 and following sections (3.2.4-3.2.8. Finally, Section 3.2.9 deals with changes to MATSim (Horni et al., 2016) that does not directly concern the mobility simulation.

3.2.1 Main algorithm

The node model proposed in this study is outlined in Algorithm 2. The model differs considerably from the original one (Algorithm 1), both regarding determining the processing order of links, Ω, determined in line 2, but especially concerning the actual mechanism for deciding whether vehicles are allowed to move across the node (line 9). We first consider the latter, before sections 3.2.3-3.2.8 address the former.

Whenever a buffer is processed, the corresponding movement, m, of the front vehicle, V, is determined. A subscript V, i.e. m_V, has been omitted in the notation in order to avoid nested subscripts in other variables. Each of such movements, m, has an associated time, t_m, (with an initial value of -1) which needs to be exceeded by the simulation time, t, before the movement can take place. As such, for any potential movement, m, it is checked whether any previous movements have temporarily disallowed m. Having $t > t_m$ is a necessary but not sufficient requirement for the movement of m, as the requirements of the original node model still applies. That is besides $t > t_m$ there has to be available space on the receiving link, or the vehicle has to have been in front of the buffer for more than ς seconds, see Section 3.1 and Algorithm 1.

Any conducted movement m, forces any conflicting move, n, to be unavailable until a certain time denoted by t_n. Formally, if m is conducted then for all conflicting moves, $n \in \mathcal{C}_m$, t_n is updated to the simulation time, i.e. $t_n \leftarrow t, \forall n \in \mathcal{C}_m$. If $t \leq t_m$, m is disallowed and no more vehicles from that buffer are processed in the time step.

Notice that cyclists may have to split up left turns in several partial steps. In such cases the algorithms handles this internally through the `getNextMove()` command, which would then return the movement to the temporary location instead

Algorithm 2 Proposed right-of-way node model

```
1: for all active nodes N ∈ N_a do
2:   Ω ← N.determineBufferOrder()
3:   while Ω ≠ ∅ do                                          # Buffers left
4:     B ← Ω.poll()                                         # Extract next buffer
5:     while B ≠ ∅ do                                       # Buffer is not empty
6:       V ← B.getFirstVehicle()
7:       m ← V.getNextMove()                                 # Determine move
8:                  # No conflicting moves and receiving link accepts vehicle
9:       if t > t_m ∧ (R_m.getSpareStorageCap() > 0 ∨ t > T_B + ς) then
10:        B.removeFirst()                                  # Remove vehicle from buffer
11:        T_B ← t                                          # Update buffer time
12:        for n ∈ C_m do                  # Make all conflicting moves unavailable
13:          t_n ← t
14:        R_m.receive(V)                                   # Move vehicle to receiving link
15:      else                                               # Receiving link rejects vehicle
16:        break                                            # Move on to next buffer
```

of the full left turn movement. Whereas the receiving link, R_m, is normally an out-link, in the case of initial steps of partial left turns for cyclists it is an in-link for which the cyclist is placed at the *front* of the relevant buffer of the queue – only changing T_B if the buffer is empty.

There are slight variations/additions in the bottom part of Algorithm 2 for different node types which if included in the general algorithm would make it overly complicated and difficult to read. The discrepancies between the stated algorithm and the actual code is clearly stated in the subsections concerning the relevant node types, Sections 3.2.4-3.2.8.

However, before going into detail with the processing order of buffers. the following subsection provides an overview of the buffer types used in the right-of-way setup.

3.2.2 Buffers

In the right-of-way node model we distinguish between vehicles wanting to turn left, and vehicles that do not. How left is defined is varies across node types. Depending on the size of the link, left turning vehicles may potentially be placed in another buffer than vehicles going elsewhere.

Three different types of buffers configurations are formulated:

- *Single buffer:* All vehicles regardless of movement is placed and stay in the same *main buffer* until they leave the link.
- *Two interacting buffers:* All vehicles are initially places in the *main buffer*, but left turning vehicles are eventually transferred to the *left buffer* when reaching the front of the main buffer.
- *Two separated buffers:* Left turning vehicles are completely separated from the other vehicles by being placed in a *left buffer*, whereas the other vehicles are placed in the *main buffer*.

The single buffer is used for all car links that only has a single lane and is of lower ranked highway type than "tertiary_link", whereas two interacting buffers are used for single lane car links of higher types. When a car link has more than two lanes the two separated buffers are used. For bicycles, the single buffer, two interacting buffers, and two separated buffers are used when links have one, two, and more than two lanes, respectively, where the number of lanes are determined according to the width of the bicycle infrastructure, see Buch and Greibe (2015) and Paulsen et al. (2019).

In-links of diverging nodes always have separated buffers, but will have as many buffers as they have out-links, potentially leading to more than two buffers.

The first and third buffer types are simple, whereas the second (two interacting buffers) needs further explanation. Like the third one, it consists of a main buffer and a left buffer, but left turning vehicles have to be at the front of the main buffer before they can be moved onto the left buffer. However, the left buffer has limited capacity (2 for cars, and 3 for bicycles), which when reached prevents left turning vehicles from being moved to the left buffer, potentially causing other vehicles to be trapped behind the queue of left turning vehicles. The capacity mimics that roads often widens a bit when approaching intersections, and that cars can then overtake on the right when only a few left turning cars are in front. Notice that a cyclists doing a partial left turn and being places at the front of the buffer is allowed to make the buffer go above capacity, as in real-life such cyclists are placed in the intersection rather than on the actual links.

In all cases the movement type (left turn or not left turn) is only determined at initial buffer insertion and is stored in the vehicle in order to avoid determining this redundantly at a later stage.

3.2.3 Buffer order determination

We now move on to how each of the node types determines the order in which the buffers must be processed.

It is worth recapping that D denotes a bundle, and consists of up to four links (in/out for car/bicycle) denoted by $L \in D$. However, when determining the buffer order of a node, only the buffers containing in-links are relevant, as uni-directional bundles leading away from the node has no way of leading traffic into the node.

For some node types the bundles are categorised into primary and secondary bundles, denoted by $\mathcal{D}_P$ and $\mathcal{D}_S$, see Section 2. Again, only bundles containing in-links need to be considered.

Furthermore, each buffer has up to four buffers, see Section 3.2.2. To make the algorithms easier to read, we assume that every bundle has get methods can access any of the four buffers, i.e. `getBicycleLeftBuffer()`, `getBicycleMainBuffer()`, `getCarMainBuffer()`, and `getCarLeftBuffer()`. The buffers are placed into the list of buffers, Ω, in order with the `addIfNotEmpty()` command placing an element at the end of the list, but only if the buffer is not empty.

3.2.4 Directed priority nodes

With the notation in place, we are ready to present how directed priority nodes determine the order of its buffers, see Algorithm 3. Notice that the determination

Algorithm 3 `determineBufferOrder()` for directed priority nodes

```
1: Ω ← ∅                                                    # Initialise buffer order
2: for D ∈ D_p do                                  # Add bicycle buffers from primary bundles
3:   Ω.addIfNotEmpty(D.getBicycleLeftBuffer())
4:   Ω.addIfNotEmpty(D.getBicycleMainBuffer())
5: for D ∈ D_p do                                  # Add car main buffers from primary bundles
6:   Ω.addIfNotEmpty(D.getCarMainBuffer())
7: for D ∈ D_p do                                  # Add car left buffers from primary bundles
8:   Ω.addIfNotEmpty(D.getCarLeftBuffer())
9: for D ∈ D_s do                                  # Add bicycle buffers from secondary bundles
10:  Ω.addIfNotEmpty(D.getBicycleLeftBuffer())
11:  Ω.addIfNotEmpty(D.getBicycleMainBuffer())
12: for D ∈ D_s do                                 # Add car main buffers from secondary bundles
13:  Ω.addIfNotEmpty(D.getCarMainBuffer())
14: for D ∈ D_s do                                 # Add car left buffers from secondary bundles
15:  Ω.addIfNotEmpty(D.getCarLeftBuffer())
16: return Ω
```

happens completely deterministically. Firstly, the bicycle buffers of the primary bundles are processed. Secondly, the main buffers of the car links of the primary bundles are added, before adding the corresponding left buffers. After this, the same process is followed for the secondary buffers $D \in \mathcal{D}_S$.

Notice that in order to not make the algorithmic representation in Algorithm 2 overly complicated, we have neglected the fact that – as the name indicates – bicycle links with two interacting buffers have buffers that interact. In the actual code there is interaction between the two, making the matter slightly more advanced than indicated by Algorithm 2. For instance, if the main buffer suddenly has a left turning cyclist the front while processing the main buffer, this cyclist is moved onto the left buffer if there is available space. If its capacity is already reached it will not process any more cyclists. However, it will then again try to move cyclists from the left buffer, and subsequently from the main buffer again. Thus, it is ensured that it processes as many cyclists as possible within the time step as long as the two buffers are not blocking each other.

Instead of only moving to the in-link of the next (in counterclockwise direction) bundle when performing partial left turns, cyclists at directed priority can proceed across several bundles at once, as long as they do not cross the priority path connecting the two primary bundles.

3.2.5 Right priority nodes

At right priority nodes no bundles have static priority over other bundles. However, if priority was given to a bundle in the previous time step, then this bundle keeps the priority in the following time step if it has any traffic that needs to be processed. The remainder of the buffers are then added in counterclockwise order.

The full algorithm for determining the buffer order can be found in Algorithm 4, with D_p denoting the bundle that had priority in the previous time step. Notice,

Algorithm 4 `determineBufferOrder()` for right priority nodes

```
1: Ω ← ∅                                                    # Initialise buffer order
2: if D_p ≠ ∅ then                                  # Priority from previous time step passed on
3:     D* ← D_p
4:     D_p ← ∅
5: else                                                      # No priority to pass on
6:     D* ← getRandomBundle()                               # Select priority at random
7: D ← D*                                          # D is now prioritised in this time step
8: repeat                                                  # Traverse bundles counterclockwise
9:     Ω.addIfNotEmpty(U.getBicycleLeftBuffer())
10:    Ω.addIfNotEmpty(U.getBicycleMainBuffer())
11:    Ω.addIfNotEmpty(U.getCarMainBuffer())
12:    Ω.addIfNotEmpty(U.getCarLeftBuffer())
13:    D ← getNextBundleCounterclockwise(D)
14: until D = D*
15: return Ω
```

that for this to work the general algorithm for processing nodes (Algorithm 2) also needs to be adjusted slightly following line 11 as D_p will be updated to be the bundle in which the sending buffer belongs, but only if this has not already been done in the time step.

In the case that the previous time step did not conduct any movements for the node, i.e. $D_p = \emptyset$, there is no predetermined priority on which to base the buffer order on. In such cases an initial bundle is randomly chosen before adding the remaining bundles – still in counterclockwise order.

For this node type (and anti-priority nodes) links with two interacting buffers are treated in the same way as explained for bicycle links with two interacting buffers at in Section 3.2.4, i.e. more interactively than hinted in Algorithm 2.

3.2.6 Anti-priority nodes

The buffering order for anti-priority nodes are very similar to those of right priority nodes, but the process falls in two parts as the bundles are grouped in primary bundles and secondary bundles. The algorithm (Algorithm 5) is equivalent to performing Algorithm 4 for right-priority nodes on the set of primary bundles first, and adding the output of doing it on the set of secondary bundles afterwards.

3.2.7 Merging nodes

Merging nodes are special in the sense that although a priority is often present, the merging situation is to some degree a mutual responsibility of the involved vehicles; the vehicles on the prioritised part bundle, D_p, (often the motorway) are supposed to make space for on-coming traffic from the secondary bundle, D_s (often a ramp). As this is not possible when the prioritised part is fully occupied, since the cars would have no possibility to make room for in-coming cars, when this is the case incoming traffic from ramps is restricted from moving.

Algorithm 5 `determineBufferOrder()` for anti-priority nodes

```
1: Ω ← ∅                                                    # Initialise buffer order
2: if D_p ≠ ∅ then                          # Priority from previous time step passed on
3:    D* ← D_p
4:    D_p ← ∅
5: else                                                     # No priority to pass on
6:    D* ← getRandomBundle()                                # Select priority at random
7: D ← D*                                          # D is now prioritised in this timestep
8: repeat                                       # Traverse primary bundles counterclockwise
9:    if D ∈ D_p then
10:      Ω.addIfNotEmpty(U.getBicycleLeftBuffer())
11:      Ω.addIfNotEmpty(U.getBicycleMainBuffer())
12:      Ω.addIfNotEmpty(U.getCarMainBuffer())
13:      Ω.addIfNotEmpty(U.getCarLeftBuffer())
14:   D ← getNextBundleCounterclockwise(D)
15: until D = D*
16: repeat                                    # Traverse secondary bundles counterclockswise
17:   if D ∈ D_s then
18:      Ω.addIfNotEmpty(U.getBicycleLeftBuffer())
19:      Ω.addIfNotEmpty(U.getBicycleMainBuffer())
20:      Ω.addIfNotEmpty(U.getCarMainBuffer())
21:      Ω.addIfNotEmpty(U.getCarLeftBuffer())
22:   D ← getNextBundleCounterclockwise(D)
23: until D = D*
24: return Ω
```

Algorithm 6 `determineBufferOrder()` for merging nodes

```
1:                                                          # Both buffers are empty
2: if D_p.getCarMainBuffer() = ∅ ∧ D_s.getCarMainBuffer() = ∅ then
3:    return ∅
4: else if D_p.getCarMainBuffer() = ∅ then                  # Primary buffer empty
5:    return D_s.getCarMainBuffer()                         # Return secondary buffer
6:                        # Secondary buffer empty or primary link is fully occupied
7: else if D_s.getCarMainBuffer() = ∅ ∨ D_p.getSpareStorageCap() ≤ 0 then
8:    return D_p.getCarMainBuffer()                         # Return primary buffer
9: else
10:   Ω ← ∅                                                 # Initialise buffer order
11:   u = rand()                                  # Choose buffer order stochastically
12:   if u < ϱ_p then          # Probability equal to proportion of total capacity
13:      Ω.add(D_p.getCarMainBuffer())
14:      Ω.add(D_s.getCarMainBuffer())
15:   else
16:      Ω.add(D_s.getCarMainBuffer())
17:      Ω.add(D_p.getCarMainBuffer())
18: return Ω
```

When the prioritised part does have space, due to merging being a mutual and cooperative tasks between the two streams of traffic, the first bundle to be processed is – just like in the original MATSim node model – randomly drawn with probability proportional to the flow capacity of the in-links of the bundles.

In Algorithm 6 the proportion of the total in-capacity (from in-links of D_p and D_s) made up by the in-link of the primary bundle, is denoted by ϱ_p.

Although `getSpareStorageCap()` is defined as associated to links, in this case we allow it to be used directly on the primary bundle, as the bundle is required have exactly one link, making the notation unambiguous.

3.2.8 Diverging nodes

The buffer order for diverging nodes is quite special as only one bundle is available. Furthermore, as no conflicting movements are possible, the buffer order does not matter why each of the buffers $B \in \mathcal{B}_D$, are simply processed in the order they are programmed, see Algorithm 7. As the only node, diverging nodes are – in the special case where more than two out-links exist – allowed to have more than two buffers.

Just like for merging nodes we do not have to consider bicycle links, as these are not allowed on these node types.

Algorithm 7 `determineBufferOrder()` for diverging nodes

```
1: Ω ← ∅                                          # Initialise buffer order
2: for D ∈ 𝒟 do                                   # Trivial as there is only one in-bundle
3:     for B ∈ ℬ_D do                             # Trivial as there is only one buffer
4:         Ω.addIfNotEmpty(B)
5: return Ω
```

3.2.9 Replanning

Due to right-of way, the travel times of going straight through the prioritised direction may differ considerably from a corresponding left turn for instance. Even more so for a cyclist turning right along a prioritised road instead of crossing it by turning left. In order to account for this, the normal shortest path searching from node to node is no longer valid. Instead, a link-to-link routing where shortest path searches are performed in an inverse network is used as suggested by Grether (2014). In the inverse network, every link of the original network becomes a node in the inverted network, and the links of the inverted network consists of all link-to-link combinations that share a node in the original network. The average travel times for each time bin stored of the new network is thus based on the travel times from entering one link to entering a new one, as opposed to the original routing where the travel time is the difference between time at link exit and link entry. This study uses the same setup for routing when using the right-of-way node model, and further discovers that in such case, all events of vehicles leaving links can be completely disregarded, i.e. neither created nor processed, since they are not used anyway. This gives a dramatic performance boost (see also Section 4.3.1).

A special case is when an agent is routing along a link-to-link combination in a time bin where there was no traffic in the previous iteration. Here the default implementation (Grether, 2014) simply provides the fully uncongested travel time of the in-link, but we have opted for a pessimistic approach, where the highest travel time associated with the in-link among the link-to-link combinations is used if such exist. This approach has been chosen in order to reduce the likelihood of agents underestimating the time spent yielding at intersections.

As in Paulsen and Nagel (2019) routing is individualised such that an agent never predicts the travel time between two links to be lower than the length of the first link divided by the maximum speed of the agent. This is relevant for trucks which are restricted to go at most 80 km/h, and for cyclists with low individual desired speeds.

4 Case study

In this section concerning the large-scale case study of Metropolitan Copenhagen, we first present the data used in the case study in Section 4.1. The configurations of the runs are given in Section 4.2, before Section 4.3 presents the results, including computational performance, aggregate statistics, and a spatial analysis of the impact of right-of-way.

4.1 Data

The network used in the case study, see Figure 2, was based on OpenStreetMap (2020), with the dedicated network reader from Zilske et al. (2011) used for creating car infrastructure, and a modified version of this for the bicycle infrastructure (also used in Paulsen and Nagel (2019)). Network links were merged where possible, i.e. where consecutive links had identical attributes (except for length) and the nodes connecting them did not connect any other links. As the methodology from Paulsen et al. (2019) is sensitive to link lengths, bicycle links with length $\lambda_l > 60$ metres were split into $\left\lceil \frac{\lambda_l}{60} \right\rceil$ sublinks of equal length. With no bicycle demand outside Metropolitan Copenhagen, all bicycle links outside this area were removed. Likewise, for car and truck traffic the outside area only includes port zone traffic, why all links with highway-type "secondary" or lower were removed to speed up computation but still allow port zone traffic to enter the area. The resulting network (see Figure 2) consisted of 144,060 nodes and 572,935 links, when counting bicycle links with multiple sublinks as a single link (see above).

The demand was based on two different but associated sources. The bicycle demand of 1,082,958 bicycle trips per day comes from the Copenhagen Model for Passenger Activity Scheduling (COMPAS) (Prato et al., 2013), whereas its successor the Copenhagen Greater Area Model for Passenger Transport (COMPASS) (Kjems and Paag, 2019) is the source of the 3,210,685 car trips and 299,416 truck trips per day. Notice that the car and truck demands are based on OD-matrices, why they are included as single trip agents, whereas the bicycle trips are associated with agents with actual day plans, but where all other trips have been ignored. In total the scenario consists of 4,593,059 trips.

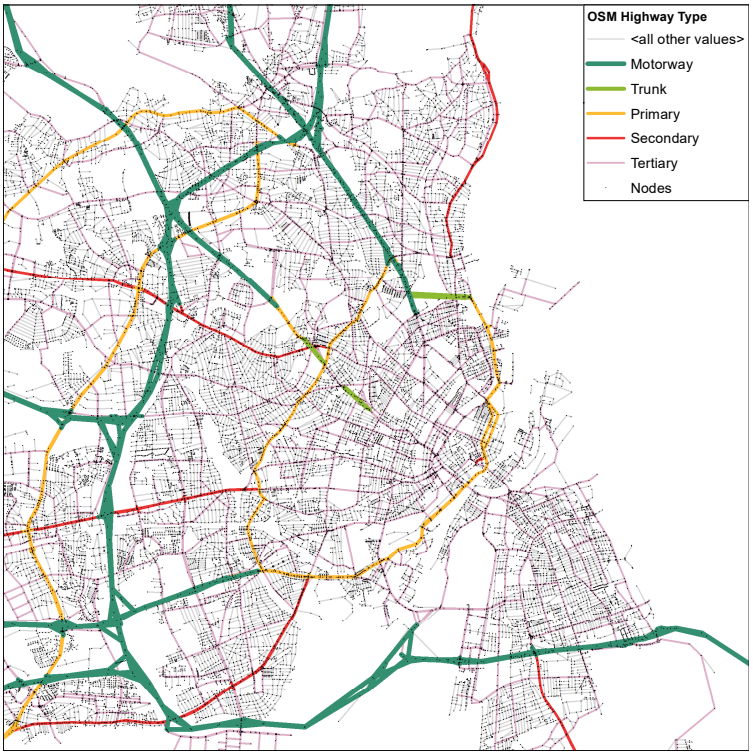


Fig. 2 Highway types and nodes for the part of the network in Greater Copenhagen where cars are allowed.

4.2 Configuration

The node classification and the node model introduced in Sections 2 and 3 have been implemented in MATSim (Horni et al., 2016). Existing approaches were used for simulating vehicles on the links. For car traffic, the standard MATSim spatial queue simulation (Gawron, 1998) was used. This simulation is a special instance of the simplified kinematic wave theory of Newell (1993) where each link has a backward wave speed such that the backward wave traverses the entire link in exactly one second (Flötteröd, 2016). Bicycles were simulated according to the methodology introduced in Paulsen et al. (2019) and implemented in MATSim in Paulsen and Nagel (2019).

Population downscaling, where only a subset (e.g. 10%) of the population is used, in turn reducing network capacities and letting each simulated entity represent multiple persons when analysing results, is often used in MATSim. It leads to serious computation time reductions, and recent literature has found the impact on

the final results to be minor (Llorca and Moeckel, 2019; Ben-Dor et al., 2020). In this study, however we have not used downscaling, as the bicycle simulation model from Paulsen et al. (2019) and the methodology we propose for right-of-way at intersections are designed for each cyclist/vehicle representing a single entity.

A stuck time (see Section 3.1) of $\varsigma = 5$ seconds has been chosen for this study. Axhausen et al. (2015) argue for lowering the stuck time to 10 seconds in order to circumvent large fluctuations in travel times between iterations, but as this was found to be insufficient in the present application (for both the original node model and the right-of-way node model), the parameter was additionally reduced to 5 seconds.

The assignment model was run for 150 iterations with a 5% rerouting rate between iterations, meaning that 5% of agents searched for a new shortest path in each iteration, except for the initial iteration for which all agents searched for the shortest path in the uncongested network. When agents search for a shortest path in subsequent iterations, a new plan containing the found shortest paths of all their trips is added to the choice set. The remaining 95% of agents were greedy in the sense that they chose the plan with the highest recorded score in their choice set.

Scores were solely based on travel times and consisted of a combination of *individualised free-flow travel time*, t_f, and congested travel time, t_c, with congested travel time being defined as the difference between the actual travel time, t_a and t_f, i.e. $t_c = t_a - t_f$. The scoring function is identical to the one used in Paulsen and Nagel (2019), with the score being a sum across all trips (legs) of the plan, $l \in \ell_p$,

$$S_p = \sum_{l \in \ell_p} \beta_f t_f^l + \beta_c t_c^l. \tag{1}$$

Note that individualised free-flow travel time varies across individuals when simulating cyclists, since these have desired speeds lower than the free-speed of the corresponding links, but also for trucks being routed at motorways due to their maximum allowed speed of 80 km/h. The chosen parameters for free flow travel time and the congested travel time were $\beta_t = -\frac{1}{60}$ per second and $\beta_c = -\frac{1}{40}$ per second, respectively.

After each iteration, the executed plan received a score based on eq. (1). As always in MATSim, plans that are not executed do not get their scores updated, as the exact values of travel times are unknown. Plans that were not executed, and thus not updated, for 20 consecutive iterations were removed from the choice set when choosing the plan for the next day. Likewise, plans with associated scores more than 30% worse than the best score in the choice set were also removed.

4.3 Results

4.3.1 Computation time and convergence

All scenarios were run on a high performance computer with two 2.8 GHz deca-core processors and 120 GB RAM using 20 threads. The computation times are reported in Table 1.

As can be seen, the proposed right-of-way (RoW) node model implementation only yields a slight increase in running time compared to the original node model.

Table 1 Average computation times per iteration for combinations of node model and included modes [min.]

	Base	RoW	
	5.15	5.42	*+ 5 %*
	18.59	18.92	*+ 2 %*
	21.77	25.19	*+ 16 %*

The largest increase in computation time is seen for the tri-modal setup (16 %), unsurprisingly, as the complexity of the processing of the intersections is higher when both bicycles and motorised traffic are included. Albeit the total computation times are large, they are still within the boundaries of what is manageable when modelling metropolitan areas, considering the large number of trips included in the case study.

It is worth mentioning that the right-of-way model saves a little time during the traffic simulation due to not having to create nor handle events when vehicles leave links, see Section 3.2.9. However, it uses considerably longer time on replanning, i.e. finding new shortest paths, as i) the graph in which the searches are performed is larger than for original node model, and ii) network inversion puts an additional overhead on the time it takes to retrieve average travel times for links and time bins of the previous iteration. This also means that if higher rerouting rates than 5% were used, the savings from omitting events when leaving links would no longer be close to balancing the increased computation time for the replanning phase.

When only simulating bicycle trips, the computation times show a dramatic improvement of 65 % compared to Paulsen and Nagel (2019). This is predominantly due to the network simplification described in Section 4.2, and partly due to using a lower rerouting rate.

Figures 3 and 4 show the convergence patterns for all scenarios, by plotting the average score of executed plans across all agents on a logarithmic y-axis. In order to do so, the average score across all agents in iteration i, $\bar{S}_i < 0$ has been transformed by $-\log_{10}(-\bar{S}_i)$ when plotting, but the axis labels have preserved the original unit of S_i.

It is clearly seen the absolute improvement from iteration 0 to 150 is is much lower for cyclists than for cars, even though the scores for cyclists is based on more than a single trip, which also causes their final scores to be lower than for cars and trucks. Disregarding that they converge to different scores because agents have different travel times, there seems to be no obvious differences between the original model and the proposed right-of-way model in terms of convergence speeds – except for the indirect effect of each iteration being computed faster.

The scores of the right-of-way model seems to have slightly larger variations between iterations before disallowing new routes to enter the choice set after iteration 120, though. Such variation is normal in MATSim, even after scores have more or less stabilised (Paulsen et al., 2018; Guggisberg, 2020). The fluctuations seem to be negligible at the aggregate level for the last 30 iterations why this is not deemed as a problem for the current study.

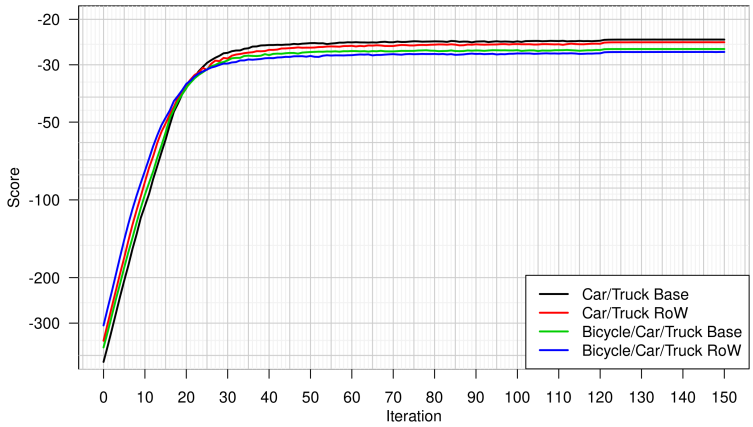


Fig. 3 Average executed score of agents for scenarios including car traffic.

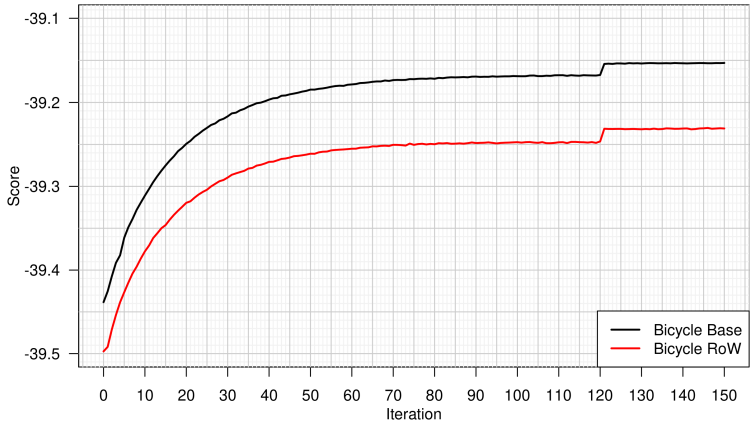


Fig. 4 Average executed score of agents for scenarios not including car traffic.

4.3.2 Aggregate statistics

The aggregate results of the simulations are summarised in Table 2. Each of the columns represents separate scenarios defined by their included modes and the applied node model. Notice that because bicycle traffic and car traffic are fully independent when using the original node model (Base), only one column showing the results of the simulation using all modes is needed for this model. Each of the row triplets provides the average free-flow, congested, and total travel time for the mode being shown at the left-hand side of the table.

By comparing the first column to the second and third column, respectively, the effect of the right-of-way (RoW) node model for isolated bicycle traffic and isolated truck and car traffic can be investigated. For bicycle traffic only minor increases in free-flow and total travel times are seen. On average cyclists, tends to cycle slightly longer routes but the effect is more or less negligible. The added total travel time corresponds to 20 % of the congested time of the original model, which is only slightly lower than for cars (24 %).

However, for cars the absolute changes are much larger, and not only the free-flow but also the congested travel time increases. This means that even though cars travel further (or on links with lower freespeed) in order to avoid congestion, they still end up spending more time in congestion that in the original node model. Interestingly, for trucks only the congested (and total) travel time increases. Generally, the results indicate that using the original node model underestimates travel times compared to the right-of-way model.

By comparing the second and third column, respectively, to the fourth column, the marginal effects of including the other mode(s) in the simulation can be found. The tendencies seems similar to those occurring when adding right-of-way to the uni/bi-modal setup, but a few points are worth mentioning.

Firstly, for all of the modes, the added travel times from including competing modes are less than or equal to the effect of including right-of-way without competing modes. This shows that, on average, intra-modal conflicts at intersections inflicts more added travel time than inter-modal conflicts. Secondly, it is found that motorised traffic witness larger added travel times when adding bicycles, than cyclists do when adding cars. This could indicate that right-of-way, as intended, is indeed favouring cyclists whenever practically possible.

Table 2 showed average travel time for the entire area across the entire day. However, this includes large portions of the days and geographical areas where congestion is rarely an issue. Instead, Table 3 provides the same measures, but where only trips departing between 7am and 8am with destination in Copenhagen or Frederiksberg Municipalities (the central most 95 square kilometres of the city)

Table 2 Average travel times per mode across all trips [min. / trip]

	Included modes				
	Node model	Base	RoW	RoW	RoW
	Free-flow travel time	15.12	15.15	—	15.17
	Congested travel time	0.15	0.15	—	0.16
	Total travel time	15.27	15.30	—	15.33
	Free-flow travel time	11.71	—	11.78	11.80
	Congested travel time	1.91	—	2.24	2.40
	Total travel time	13.62	—	14.02	14.20
	Free-flow travel time	13.79	—	13.78	13.80
	Congested travel time	1.47	—	1.88	2.07
	Total travel time	15.16	—	15.66	15.87

Table 3 Travel times for trips with destination in Copenhagen or Frederiksberg Municipalities departing between 8am and 9am – average across 60,202 bicycle trips, 40,082 car trips and 6,649 truck trips [min. / trip]

	Included modes				
	Node model	Base	RoW	RoW	RoW
	Free-flow travel time	13.87	13.92	—	13.97
	Congested travel time	0.35	0.37	—	0.40
	Total travel time	14.22	14.29	—	14.37
	Free-flow travel time	12.50	—	12.58	12.67
	Congested travel time	4.88	—	5.93	7.15
	Total travel time	17.38	—	18.51	19.82
	Free-flow travel time	8.87	—	8.90	8.98
	Congested travel time	2.55	—	3.26	4.30
	Total travel time	11.37	—	12.16	13.28

are included. Notice, that as opposed to the grand totals of Table 2 where car trips largely outnumber trips for other modes, the majority of the subset of trips used for calculating the numbers in Table 3 are bicycle trips (60,202 as opposed to 40,082 car trips and 6,649 truck trips).

The effects are more profound than for the total averages. This is as expected since the road network of the central municipalities is denser, and the traffic densities are higher. Again, the changes for bicycle traffic is limited, although a small increase in congested travel time is seen when adding right-of-way and cars to the scenario. For the motorised traffic we see that the marginal effect of modelling right-of-way is large even in the bi-modal setup. However, for all three modes the marginal effect of adding competing modes is the larger of the two effects.

This is in contrast to the overall averages across the entire day and area (Table 2, where intra-modal interactions were found to have a larger impact than inter-modal interactions. The findings highlight the importance of modelling conflicts of intersections in the central part of cities, as they – to a much larger extend than rural intersections – influence travel time considerably, especially when including a large share of bicycle traffic.

The average travel times for bicycle trips to Copenhagen is in line with the Danish National Travel Survey (Technical University of Denmark, 2020), while the modelled average travel times for cars are somewhat below. This is could due to having too short trips in the input demand, but may also be due to specific parameters of the model requiring additional calibration. Truck trips with destination in Copenhagen are very short, which can possibly be explained by a large proportion of short delivery trips.

4.3.3 Spatial analysis

Until now the focus has solely been the effects on average travel times. In this spatial analysis, the differences in flow across the different scenarios are examined.

In Figure 5, the link flow differences for cars and trucks between the run with the original node model and right-of-way scenario without bicycles are shown. Tendencies have been made more clear by using Inverse Squared Distance Weighting (ISDW) (Shepard, 1968) to create an interpolation surface behind the network. In order to uniquely determine locations of flows for the Inverse Squared Distance Weighting, differences in the sum of in- and out-flows at node levels have been used as input.

It is seen that traffic is generally attracted to larger roads with motorways (see Figure 2) receiving the highest additional flows. This is both seen for the motorway bypass, but also for the two motorways going through the western border of the map. In the latter case, it is clearly seen that the added flows can be explained by the parallel alternatives using parallel secondary road being less attractive.

Many tertiary roads gains higher flows by attracting traffic from nearby smaller residential roads, with a prominent example being observed in the upper part of

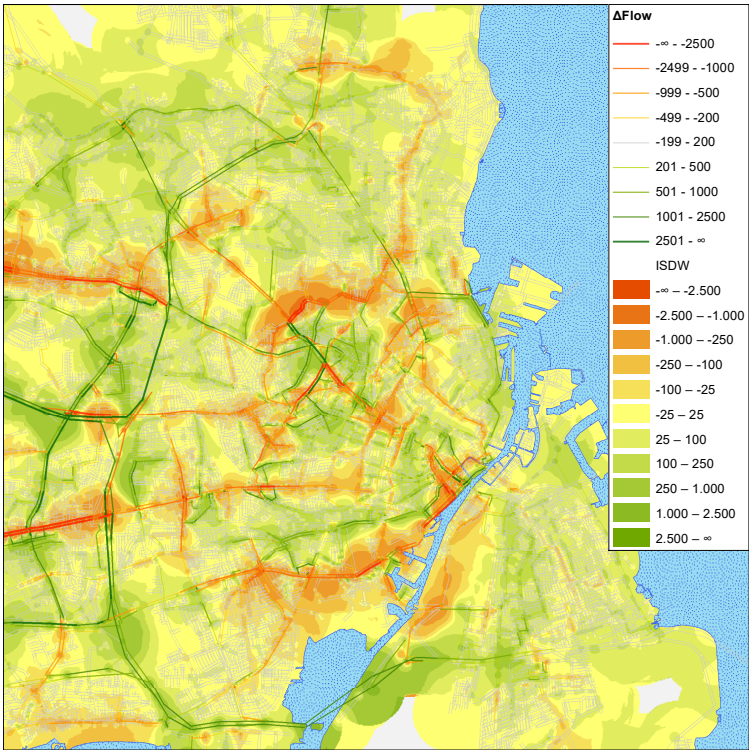


Fig. 5 Car flow differences between the original node model and the right-of-way model when only including cars. Negative numbers imply lower car flows when using the right-of-way model.

the central part of the map. Using the original node model, many agents found it beneficial to use a shortcut of lower ranked links which is now no longer beneficial.

Another tendency is that traffic is more likely to avoid the central part of the city, as the main arterial roads going towards the city – and the city centre itself – are seen to have lower flows when modelling right-of-way. This shows that neglecting to model right-of-way does not only underestimate travel time, it also causes the flow distribution between urban roads and motorways to be biased towards the urban roads going through the city centre.

These findings are in line with Nielsen et al. (1998a, 2002) and the expectations of the impacts of the right-of-way model, as the right-of-way model is designed to add excess travel times to vehicles passing through intersections via non-prioritised roads.

Whereas the previous map showed the effect of including right-of-way for car traffic, Figure 6, shows the car/truck flow differences between the car/truck right-

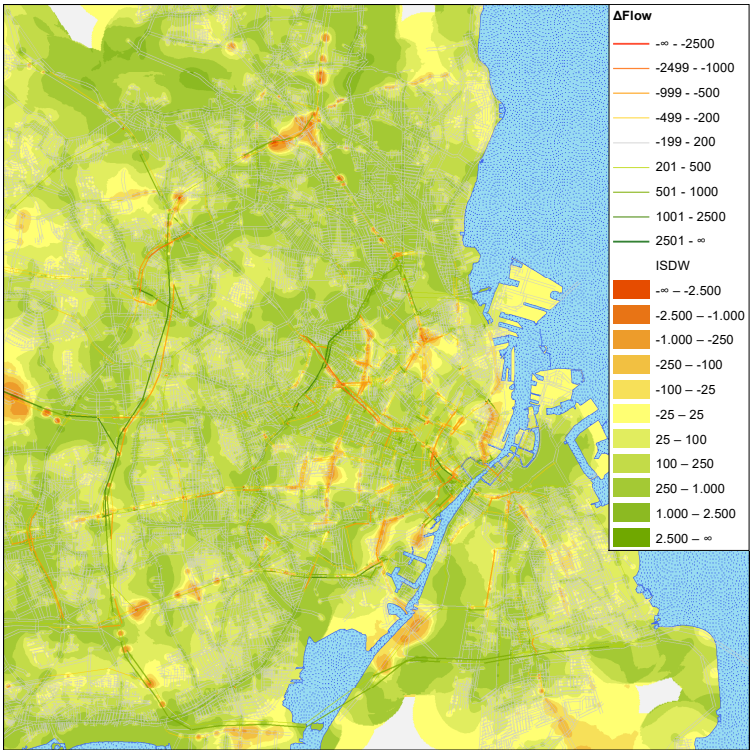


Fig. 6 Car flow differences for the right-of-way model with and without bicycles. Negative numbers imply lower car flows when including bicycle traffic.

of-way scenario, and the joint tri-modal right-of-way scenario. In this regard, all the flow changes are derived marginal effects from having to yield for bicycles.

A tendency seems to be that cars are more likely to avoid the city centre when modelling right-of-way and including bicycle traffic. This makes sense as the bicycle traffic is most profound in this area. Instead, some car users seem to avoid the central part of the city altogether by using the motorway bypass, although to a less extend that what is seen in Figure 5. The inner ring road (of type "Primary" in Figure 2) also gets higher flows when adding bicycles to the model.

This shows that traffic through the central part of the city will still be overestimated in models that only includes intersection delays caused by intra-modal interactions, as there also is a marginal effect of inter-modal interactions caused by conflicting bicycle traffic. In consequence, multi-modality in the case of bicycle traffic and motorised traffic must be dealt with jointly, as treating them separately would bias the resulting traffic flows.

A map showing the full car flow differences between using the original MATSim node model and using the right-of-way node model while also simulating bicycle traffic is included as a supplementary Figure 7 in Appendix A. As such, Figure 7 can be seen as the sum of the effects from Figures 5 and 6.

5 Discussion and future research

The previous sections showed that disregarding conflicts at intersections in traffic assignment models will generally inflict a bias, in the sense that travel times through the urban core are underestimated. This is in line with previous studies on intersection modelling for car traffic such as Nielsen et al. (1998a, 2002). The fact that bicycle traffic reduces the traffic through the city centre even further, may have undesirable effects on how cities are formed. If transport models show that cars are not willing to use bypasses, there is a risk that policy makers may adopt the idea that cars have to go through the city centre, although it might better to separate car and bicycle traffic by guiding a more cars around the city.

Travel times of bicycle traffic did not change much when adding car traffic to the simulation – at least not in absolute terms. But even though the sheer added travel time was low, waiting at intersections is still particular burdensome for cyclists as even short waits requires braking, (partially) dismounting, and using a lot of energy to regain speed. In this regard it would be relevant to add other relevant attributes to the scoring functions as the current setup only includes parameters for "free flow" and congested travel time.

One of the reasons why excess travel time for cyclists is low in the current study, could be due to traffic signals not being included in the model. Traffic signals often causes waiting time at intersections, but on the other hand also frequently allow traffic to cross busy roads. Although traffic signals in Copenhagen are coordinated so that cyclists face green waves, including traffic signals to the model would most likely increase the time of bicycle traffic. The implementation of traffic lights is methodologically straightforward, as it can be done by simply updating t_m of all movements facing a red light, and would have practically zero impact on the computation times of the model. Determining the phases, offsets, and which nodes and links to include in each signalised intersection would, however, require a lot

of manual work as such information is not available in OpenStreetMap (OpenStreetMap, 2020), and no previous models of the region utilising such information are based on OpenStreetMap.

The proposed methodology still fulfils a valuable task, although traffic lights are not included in the study. Firstly, because it is very relevant in applications where signalised intersections are included, since omitting to do so could cause travellers to avoid traffic lights, as the excess travel time from a small detour around the traffic signal will be lower – when allowing passing through conflicting traffic – than corresponding the waiting time the traveller would have to incur at the traffic light. Secondly, particularly for appraisal purposes concerning major changes in the network, the model could be a pragmatic solution to capture some of the delays at intersections without having to deal with the cumbersome task of manually configuring traffic signals for such scenarios. As such, comparing the differences between results of the proposed method and scenarios with traffic lights implemented would be relevant for future research.

In addition to being a well-suited method for evaluating projects believed to have a large multi-modal impact, the model would be particularly relevant to evaluate the effects of changes in right-of-way laws, such as lowered/increased prioritisation of bicycle traffic. Using the model to investigate the effect of such policies would also be an interesting research topic for future research. Finally, it would also be relevant to investigate how increased demand for cars and bicycles influence congested travel time with and without modelling of right-of-way.

6 Conclusions

This study formulated and applied a method for simulating cars and bicycles separately along links while modelling in great detail the delays caused by intra-modal as well as inter-modal conflicts at intersections. Intersection delays have never been included in a dynamic traffic assignment models with realistic simulation of bicycle traffic, and a model for inter-modal conflicts at intersections has never integrated in a large-scale agent-based traffic assignment model. The network nodes were initially categorised into one of five different node types, allowing each of the nodes to process crossing traffic differently while taking multi-modal right-of-way into account. The methodology was implemented in MATSim (Horni et al., 2016), and applied to a large-scale case study of Metropolitan Copenhagen with a total of 4,593,059 trips of bicycles, cars, and trucks assigned to a network of 572,935 links and 144,060 nodes. The case study showed that despite increasing the computation time slightly (2-16 %), the suggested node modelling methodology still allows large-scale scenarios to be simulated within feasible computation time.

The results demonstrated that the marginal effects on bicycle traffic when adding right-of-way and cars to the simulation was almost negligible, although a travel time increase equivalent to a noteworthy part of the congested time was found for bicycles in Copenhagen in the morning rush hour.

However, right-of-way had substantial impact on motorised traffic. Firstly, it was shown that travel times were considerably underestimated when neglecting to include travel time delays caused by conflicts at intersections, especially in the central part of the city when also including bicycle traffic. Secondly, car flows were

seen to be reduced in the central part of the city, whereas bypasses, especially the bypass motorway, received additional flow compared to the baseline scenario.

Although no previous studies have investigated the effect of multi-modal right-of-way in agent-based transport simulation, these findings seem intuitive and in line with somewhat similar studies such as Nielsen et al. (1998a, 2002), who also found that turn delays caused by other motorised traffic would deviate traffic towards alternatives avoiding the city centre. But this study further underlines that the effect of bicycle traffic most also be included, as the inter-modal conflicts account for the majority of the intersection delays for cars travelling in the urban core and has considerable impact on flow distribution.

The study has shown that it is possible to model intersections in agent-based traffic simulation in much greater detail and much more realistically than what is normally done without causing noteworthy overhead in computation times. It further highlights the importance of treating multi-modality jointly rather than separated in urban traffic assignment models, even in cities with a high level of segregation between transport modes as the impact of inter-modal conflicts at intersections are too influential to ignore.

References

Agarwal A, Lämmel G, Nagel K (2018) Incorporating within link dynamics in an agent-based computationally faster and scalable queue model. Transportmetrica A: Transport Science 14(5-6):520–541, DOI 10.1080/23249935.2017.1364802, URL https://www.tandfonline.com/doi/full/10.1080/23249935.2017.1364802

Agarwal A, Ziemke D, Nagel K (2019) Bicycle superhighway: An environmentally sustainable policy for urban transport. Transportation Research Part A: Policy and Practice 137:519–540, DOI 10.1016/j.tra.2019.06.015, URL https://linkinghub.elsevier.com/retrieve/pii/S096585641731162X

Allen DP, Hummer JE, Rouphail NM, Milazzo JS (1998) Effect of Bicycles on Capacity of Signalized Intersections. Transportation Research Record: Journal of the Transportation Research Board 1646(1):87–95, DOI 10.3141/1646-11, URL http://journals.sagepub.com/doi/10.3141/1646-11

Axhausen KW, Horni A, Herrmann HJ (2015) Final Report: The Risk for a Gridlock and the Macroscopic Fundamental Diagram. Technical report. Institute for Transport Planning and Systems, ETH Zurich

Ben-Dor G, Ben-Elia E, Benenson I (2020) Spatiotemporal Implications of Population Downscaling: A MATSim Study of Sioux Falls Morning Peak Traffic. Procedia Computer Science 170:720–725, DOI 10.1016/j.procs.2020.03.165, URL https://doi.org/10.1016/j.procs.2020.03.165

Brilon W, Miltner T (2005) Capacity at intersections without traffic signals. Transportation Research Record (1920):32–40, DOI 10.3141/1920-04

Buch TS, Greibe P (2015) Analysis of Bicycle Traffic on One-Way Bicycle Tracks of Different Width. In: European Transport Conference 2015, Association for European Transport (AET), Frankfurt, Germany, URL https://aetransport.org/en-gb/past-etc-papers/conference-papers-2015?abstractId=4418&state=b

Charypar D (2008) Efficient algorithms for the microsimulation of travel behavior in very large scenarios. PhD thesis, ETH Zurich

Chen J, Wang W, Li Z, Jiang H, Chen X, Zhu S (2014) Dispersion Effect in Left-Turning Bicycle Traffic and Its Influence on Capacity of Left-Turning Vehicles at Signalized Intersections. Transportation Research Record: Journal of the Transportation Research Board 2468(1):38–46, DOI 10.3141/2468-05, URL http://journals.sagepub.com/doi/10.3141/2468-05

Chen X, Shao C, Vue H (2007) Influence of bicycle traffic on capacity of typical signalized intersection. Tsinghua Science and Technology 12(2):198–203, DOI 10.1016/S1007-0214(07)70028-7, URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6076200

Chen XM, Shao CF, Hao Y (2008) Influence of pedestrian traffic on capacity of right-turning movements at signalized intersections. Transportation Research Record (2073):114–124, DOI 10.3141/2073-13

Chen Y, He Y, Sun X (2015) Impact of Pedestrian Traffic on Saturation Rate of Protected Left-Turn at Urban Intersections. Open Journal of Applied Sciences 05(01):22–31, DOI 10.4236/ojapps.2015.51003

Chiu Yc, Bottom J, Mahut M, Paz A, Balakrishna R, Waller T, Hicks J (2011) Dynamic Traffic Assignment. June

Dandl F, Bracher B, Bogenberger K (2017) Microsimulation of an autonomous taxi-system in Munich. In: 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), IEEE, pp 833–838, DOI 10.1109/MTITS.2017.8005628, URL http://ieeexplore.ieee.org/document/8005628/

Fellendorf M (1994) VISSIM: A microscopic Simulation Tool to Evaluate Actuated Signal Control including Bus Priority. 64th Institute of Transportation Engineers Annual Meeting pp 1–9

Flötteröd G (2016) Queueing Representation of Kinematic Waves. In: The Multi-Agent Transport Simulation MATSim, Ubiquity Press, London, chap 50, pp 347–352, DOI 10.5334/baw.50, URL http://www.ubiquitypress.com/site/chapters/10.5334/baw.50/

Gawron C (1998) An iterative algorithm to determine the dynamic user equilibrium in a traffic simulation model. International Journal of Modern Physics C 9(03):393–407, DOI 10.1142/S0129183198000303

Grether D (2014) Extension of a Multi-Agent Transport Simulation for Traffic Signal Control and Air Transport Systems. PhD thesis, Technische Universität Berlin, URL http://dx.doi.org/10.14279/depositonce-4084

Grether D, Thunig T (2016) Traffic Signals and Lanes. In: The Multi-Agent Transport Simulation MATSim, Ubiquity Press, London, chap 12, pp 83–88, DOI 10.5334/baw.12, URL http://www.ubiquitypress.com/site/chapters/10.5334/baw.12/

Guggisberg D (2020) Convergence towards equilibrium in an agent-based transport simulation. MSc Thesis. Institute for Transport Planning and Systems, ETH Zurich.

Guo Y, Yu Q, Zhang Y, Rong J (2012) Effect of Bicycles on the Saturation Flow Rate of Turning Vehicles at Signalized Intersections. Journal of Transportation Engineering 138(1):21–30, DOI 10.1061/(ASCE)TE.1943-5436.0000317, URL http://ascelibrary.org/doi/10.1061/%28ASCE%29TE.1943-5436.0000317

Horni A, Nagel K, Axhausen KW (eds) (2016) The Multi-Agent Transport Simulation MATSim. Ubiquity Press, London, DOI 10.5334/baw, URL https://www.ubiquitypress.com/site/books/10.5334/baw/

Kjems S, Paag H (2019) COMPASS : Ny trafikmodel for hovedstadsområdet. In: Artikler fra Trafikdage på Aalborg Universitet (Proceedings from the Annual Transport Conference at Aalborg University)

Krajzewicz D, Erdmann J, Härri J, Spyropoulos T (2014) Including Pedestrian and Bicycle Traffic in the Traffic Simulation SUMO. 10th ITS European Congress Helsinki (June):10

Li H, Deng W, Tian Z, Hu P (2009) Capacities of unsignalized intersections under mixed vehicular and nonmotorized traffic conditions. Transportation Research Record (2130):129–137, DOI 10.3141/2130-16

Li H, Tian Z, Deng W (2011) Capacity of Multilane All-Way Stop-Controlled Intersections Based on the Conflict Technique. Transportation Research Record: Journal of the Transportation Research Board 2257(1):111–120, DOI 10.3141/2257-13, URL http://journals.sagepub.com/doi/10.3141/2257-13

Llorca C, Moeckel R (2019) Effects of scaling down the population for agent-based traffic simulations. Procedia Computer Science 151(2018):782–787, DOI 10.1016/j.procs.2019.04.106, URL https://doi.org/10.1016/j.procs.2019.04.106

Milazzo JS, Rouphail NM, Hummer JE, Allen DP (1998) Effect of Pedestrians on Capacity of Signalized Intersections. Transportation Research Record: Journal of the Transportation Research Board 1646(1):37–46, DOI 10.3141/1646-05, URL http://journals.sagepub.com/doi/10.3141/1646-05

Mondal S, Gupta A (2020) A review of methodological approaches for saturation flow estimation at signalized intersections. Canadian Journal of Civil Engineering 47(3):237–247, DOI 10.1139/cjce-2018-0696

Newell G (1993) A simplified theory of kinematic waves in highway traffic, part I: General theory. Transportation Research Part B: Methodological 27(4):281–287, DOI 10.1016/0191-2615(93)90038-C, URL http://linkinghub.elsevier.com/retrieve/pii/019126159390038C

Nielsen OA, Frederiksen RD, Simonsen N (1998a) Stochastic user equilibrium traffic assignment with turn-delays in intersections. International Transactions in Operational Research 5(6):555–568, DOI DOI:10.1016/S0969-6016(98)00043-4, URL http://www.sciencedirect.com.ezproxy.library.wisc.edu/science/article/B6VGM-3VKBHD9-D/2/14981fdfb464976fdd534feb1392a105

Nielsen OA, Frederiksen RD, Simonsen N (1998b) Using expert system rules to establish data for intersections and turns in road networks. International Transactions in Operational Research 5(6):569–581, DOI 10.1111/j.1475-3995.1998.tb00138.x

Nielsen OA, Daly A, Frederiksen RD (2002) A Stochastic Route Choice Model for Car Travellers in the Copenhagen Region. Networks and Spatial Economics 2(4):327–346, DOI 10.1023/A:1020895427428

Niittymäki J, Pursula M (1997) Saturation Flows at Signal-Group-Controlled Traffic Signals. Transportation Research Record: Journal of the Transportation Research Board 1572(1):24–32, DOI 10.3141/1572-04, URL http://journals.sagepub.com/doi/10.3141/1572-04

OpenStreetMap (2020) https://www.openstreetmap.org. URL https://www.openstreetmap.org

Paulsen M, Nagel K (2019) Large-Scale Assignment of Congested Bicycle Traffic Using Speed Heterogeneous Agents. Procedia Computer Science 151:820–825, DOI 10.1016/j.procs.2019.04.112, URL https://linkinghub.elsevier.com/retrieve/pii/S1877050919305769

Paulsen M, Rasmussen TK, Nielsen OA (2018) Output variability caused by random seeds in a multi-agent transport simulation model. Procedia Computer Science 130:850–857, DOI 10.1016/j.procs.2018.04.078, URL https://linkinghub.elsevier.com/retrieve/pii/S187705091830440X

Paulsen M, Rasmussen TK, Nielsen OA (2019) Fast or forced to follow: A speed heterogeneous approach to congested multi-lane bicycle traffic simulation. Transportation Research Part B: Methodological 127:72–98, DOI 10.1016/j.trb.2019.07.002, URL https://linkinghub.elsevier.com/retrieve/pii/S0191261518310336

Peeta S, Ziliaskopoulos AK (2001) Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future. Networks and Spatial Economics 1(3/4):233–265, DOI 10.1023/A:1012827724856, URL http://link.springer.com/10.1023/A:1012827724856

Prato CG, Rasmussen TK, Nielsen OA, Watling DP (2013) A disaggregate pseudo-dynamic assignment for the activity-based model of the Greater Copenhagen Area. In: Joao V (ed) 13th World Conference on Transport Research (WCTR), Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, pp 1–19

Preethi P, Ashalatha R (2018) Modelling saturation flow rate and right turn adjustment factor using area occupancy concept. Case Studies on Transport Policy 6(1):63–71, DOI 10.1016/j.cstp.2017.11.001, URL https://www.sciencedirect.com/science/article/pii/S2213624X17303243

Raney B, Cetin N, Völlmy A, Vrtic M, Axhausen K, Nagel K (2003) An Agent-Based Microsimulation Model of Swiss Travel: First Results. Networks and Spatial Economics 3(1):23–41, DOI 10.1023/A:1022096916806

Rieser M, Nagel K (2008) Network breakdown "at the edge of chaos" in multi-agent traffic simulations. European Physical Journal B 63(3):321–327, DOI 10.1140/epjb/e2008-00153-6

Roshani M, Bargegol I (2017) Effect of Pedestrians on the Saturation Flow Rate of Right Turn Movements at Signalized Intersection - Case Study from Rasht City. IOP Conference Series: Materials Science and Engineering 245(4), DOI 10.1088/1757-899X/245/4/042032

Shepard D (1968) Two- dimensional interpolation function for irregularly- spaced data. In: Proceedings of the 1968 23rd Acm National Conference, ACM Press, pp 517–524, DOI 10.1145/800186.810616

Technical University of Denmark (2020) Danish National Travel Survey 2006-2020. Center for Transport Analytics, DTU Management, DOI 10.11581/dtu:00000035, URL https://www.cta.man.dtu.dk/english/national-travel-survey

Thunig T, Scheffler R, Strehler M, Nagel K (2019) Optimization and simulation of fixed-time traffic signal control in real-world applications. Procedia Computer Science 151:826–833, DOI 10.1016/j.procs.2019.04.113, URL https://www.sciencedirect.com/science/article/pii/S1877050919305770

Wierbos MJ, Knoop VL, Hänseler FS, Hoogendoorn SP (2019) A Macroscopic Flow Model for Mixed Bicycle–Car Traffic. Transportmetrica A: Transport Science pp 1–14, DOI 10.1080/23249935.2019.1708512, URL https://www.

tandfonline.com/doi/full/10.1080/23249935.2019.1708512
Zilske M, Neumann A, Nagel K (2011) OpenStreetMap For Traffic Simulation. M Schmidt, G Gartner (Eds), Proceedings of the 1st European State of the Map – OpenStreetMap conference, no 11-10 pp 126–134

A Supplementary figures

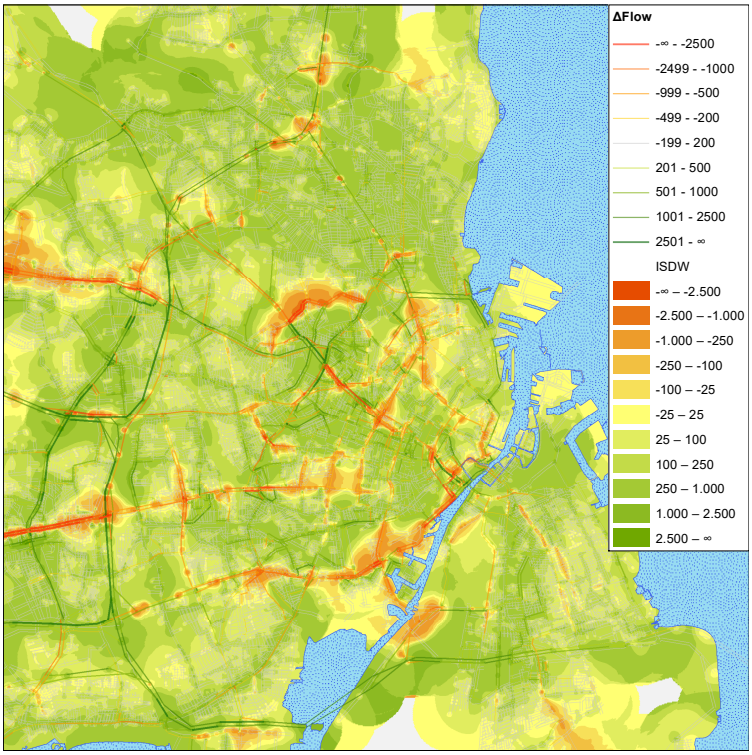


Fig. 7 Car flow differences between the uni-modal car scenario with the original node model and the tri-modal right-of-way model scenario. Negative numbers imply lower car flows in the latter.

7 Conclusions

The PhD thesis has led to a further understanding of how agent-based modelling can be used for various types of mesoscopic simulation of multi-modal urban traffic. The contributions across three topics presented in this thesis are important to researchers for further development of such models, but also gives rise to more general implications in transport planning, for instance regarding how to prioritise urban spaces in a way that facilitates efficient transport in cities.

The following sections (Sections 7.1-7.3) present the main conclusions within the three research areas of the thesis: i) Output variability in agent-based simulation of transport systems, ii) Agent-based passenger delay modelling with real-time information, and iii) Agent-based simulation of bicycle traffic and interaction with cars. Section 7.4 summarises the policy implications spawned by the five articles and the PhD thesis as a whole, whereas the final Section 7.5 outlines directions for future research.

7.1 Output variability in agent-based simulation of transport systems

Part I of the thesis consisting of Paper 1 raises the question of the how much the output of agent-based transport simulation models differ when changing the random seed. Whereas other sources of uncertainties can be limited by providing better model assumptions or input, the variability originating from the choice of random seeds is a fully inherent part of such models. Although the phenomenon has previously been documented in the literature for decades (Nagel, 1997), no previous studies has taken on the task of scrutinising whether the variability of disaggregate measures such as individual link flow is a problem of real concern or simply an academic curiosity.

The paper examines link flow variability in depth through a large-scale MATSim (Horni et al., 2016) case study of Santiago de Chile (Kickhöfer et al., 2016) with 665,201 agents. Using out-of-the-box configurations the scenario is run 100 times with a different random seed in each run. In line with existing literature (Lawe et al., 2009; Ziems et al., 2011), the analysis based on the variability in daily flows across the 100 seeds for each of the 22,981 links finds that the coefficients of variation were below 5% for the majority of links. However, despite the low average coefficients of variation, some links – primarily links with low traffic volumes but also some with high – deviate more than 15% from their grand total average daily traffic flow for certain seeds. Luckily, this effect is found to be almost fully independent of the selected random seed on an aggregate level, meaning that the study does not find any seeds that are systematically closer to or further from the

mean of the distribution. Finally, it is shown that the variation across runs with different seeds is generally much higher than variation between the last and penultimate iteration of each seed. This is interesting as it provides empirical support for the suggestion of Flötteröd (2016) to recognise that the output of a single run should not be regarded as a unique solution, but rather a single realisation from the set of possible solutions, why it is more appropriate to base analyses on outputs across several model runs using different seeds. The paper goes even further by recommending doing so even if it comes at the expense of reducing the number of iterations per seed.

7.2 Agent-based passenger delay modelling with real-time information

Paper 2 forming Part II of the thesis focuses on passenger delays in public transport systems. Whereas the literature has an abundance of studies on vehicle delays, the extend of literature on passenger-oriented perspectives leaves much to be desired (Parbo et al., 2016). This is somewhat remarkable, as the experiences of the users ideally should be central when determining whether a transport system is successful.

The paper formulates an agent-based, adaptive passenger path choice model for simulating the en-route decisions of public transport passengers as they travel through the network while receiving real-time information about the public transport system. The paper differentiates itself from other similar studies by using 65 days of actual Automated Vehicle Location (AVL) data for trains and buses as direct input for vehicles delays, previously only done for uni-modal train networks in Nielsen et al. (2009); Lijesen (2014) and the in many ways inadequate study by Paulsen et al. (2018). For each of the 65 days with available data corresponding paths are modelled for all 801,719 public transport trips in the public transport system of Metropolitan Copenhagen offering a wide range of alternatives through transfers.

The computation times of the model are kept low due to several methodological and technical optimisations, making it fast enough to be large-scale operational on a day-to-day basis or in real-time. An analysis of occurrences of on-board notifications of better alternatives emerging shows that such a system is generally most useful at stations with a variety of high-classed alternatives. However, the paper does not find it particularly favourable to be able to also search for better alternatives while walking or being on board public transport vehicles, as although the mean passenger delay can be reduced marginally, it comes at the expense of slightly increasing the standard deviation and often leads to what eventually turns out to be suboptimal paths. Spatial effects are also investigated, showing that passenger delays are generally lower when travelling to or from the central parts of Copenhagen, and that passenger redistribution away from a particular unreliable railway line (Kystbanen) is substantial. The study finds that receiving real-time

information, either pre-trip but ideally during the trip is key to keeping passenger delays at a reasonable level, as average delays are reduced from 10.4 minutes to 5.3 (pre-trip) or 3.5 (en-route) minutes by doing so. In all cases, passenger delays are still way above the corresponding vehicle delays, stressing that using vehicle delays as a measure of the performance of the system is insufficient from a passenger perspective.

7.3 Agent-based simulation of bicycle traffic and interaction with cars

Part III containing Papers 3, 4, and 5 is the final and largest part of the thesis. It develops a methodology for realistic and efficient simulation of high intense bicycle traffic based on heterogeneity across cyclists, and integrates it in a dedicated bicycle traffic assignment model as well as a joint traffic assignment model for bicycle and car traffic, both applied in large-scale case studies of Metropolitan Copenhagen. The contributions and main findings of each of the three associated papers are summarised in the following.

The first of these papers (Paper 3) is in many ways the underlying theoretical foundation for the two subsequent papers. In line with Twaddle et al. (2014) the literature review finds that previous bicycle simulation approaches were either too simple to be realistic or too computationally demanding to be feasible in large-scale applications. As such, the paper develops a computational efficient agent-based simulation model dedicated for bicycle traffic based on a set of simple assumptions. The assumptions are all tailor made for bicycle traffic and acknowledges the large heterogeneity among cyclists and the need for explicit modelling of overtaking. Two of the defining assumptions concerning desired speed and preferred headway distances. These are assumed to vary across cyclists, and suitable distributions – Johnson's S_U Distribution for desired speeds and a speed dependent symmetric Beta Distribution for headway distances – are estimated accordingly based on observed bicycle traffic data from low intensity traffic. Another important assumption is that bicycle links can be divided into pseudolanes among which agents explicitly have to choose between on-the-go when entering links.

The model is put to the test through small scale experiments of one simulation hour on two networks of three serial links with traffic intensities varying between 50 and 10,000 cyclists per hour. As intended, the model has the ability to often delay cyclists with high desired speeds, who are delayed even at moderate traffic flows, whereas cyclists with lower desired speeds generally require larger traffic flows to be witness excess travel time. The corresponding fundamental diagrams are compared to those from observed bicycle traffic of 3,763 cyclists on the very busy Queen Louise's Bridge collected through one hour of state-of-the-art video tracking data, and are found to coincide extremely well with the fundamental diagrams of the simulation experiments. The research highlights the importance of recognising and explicitly modelling overtaking and speed heterogeneity

when simulating congested bicycle traffic, and demonstrates that such a model can be computationally efficient and estimated entirely based on low volume bicycle traffic data.

Whereas Paper 3 presumes that its simulation model is large-scale applicable based on running times of small scale experiments, this is eventually proven in Paper 4 (and again in Paper 5 with even lower running times) where the methodology is implemented in MATSim (Horni et al., 2016). The paper (Paper 4) begins by exposing the literature gap regarding proper traffic assignment of dedicated bicycle traffic, and closes the gap by combining aforementioned methodology with a route choice model in a framework where travel times are influenced by the demand and route choices respond to the travel times. The resulting model is applied to a large-scale case study of Metropolitan Copenhagen with 1,082,958 bicycle trips spread across 547,085 agents with individualised desired speeds and headway distance preferences. The study shows that an iterative approach with a feedback is in fact necessary as some cyclists end up choosing different routes than in the initial iteration – avoiding busy corridors with few lanes, and being attracted to corridors with many lanes. Excess travel times are found to be small (0.12 minutes on average), which can be partially explained by excess capacity in the network when rerouting. Although a fictitious scenario with single-lane infrastructure shows congestion would be considerable if the infrastructure was poor, the study backs that bicycle traffic is highly resistant to congestion.

The final paper, Paper 5, extends the methodology of the two previous papers by including car traffic and multi-modal intersection modelling. The paper formulates a node model that simulates traffic in accordance with right-of-way laws, ensuring that vehicles and cyclists must yield for intra-modal and inter-modal conflicting movements of higher priority. The model categorises every node of the network into one of five overall node types which processes incoming traffic in different ways. The method is applied to a large-scale scenario of Metropolitan Copenhagen with more than 4.5 million trips spread across cars (3,210,685), trucks (299,416), and bicycles (1,082,958) and a network consisting of 572,935 and 144,060 nodes. Scenarios with the original MATSim node model as well as the right-of-way model formulated in the paper are each run with three different populations: Cyclists only, cars and trucks only, and all of the three. The model is demonstrated to be large-scale applicable with running times only slightly higher (2-16 %) than the original model. The aggregated travel times show that including right-of-way at nodes increases travel times considerably for car and truck traffic, especially within the city where conflicting bicycle traffic is frequent and inter-modal conflicts is a larger source of added travel time at intersections than the intra-modal conflicts from other cars. Furthermore, a spatial analysis of flow differences illustrates that neglecting to model right-of-way at intersections not only underestimates travel times but also overestimates the traffic going through the city centre. As such, the study underlines the importance of modelling multi-modality in urban areas jointly.

7.4 Policy implications

The PhD thesis has contributed within three different areas of agent-based transport simulation. A selection of the most relevant associated policy implications is given below.

Firstly, the findings of Paper 1 can be used as important guidelines concerning how to use agent-based transport models when performing project appraisal. Given that the results can be generalised, the conclusions discourage analysing results on the basis of a single model run, unless only evaluating very aggregate measures. Instead, from a transport modelling perspective it seems to be favourable to run several seeds with fewer iterations than many iterations with a single seed, and then present outputs as distributions. This may, however, have undesired implications as model predictions with uncertain outcomes to a large degree have a risk of being disregarded by decision makers that deem them incomprehensible (te Brömmelstroet et al., 2017). Alternatively, in order to still be able to present results as a single number, using the mean across different seeds instead of the output of a single run as suggested by Wegener (2011); Flötteröd (2016) could become customary. It may also indicate that such models, in general, are better suited for evaluating wider policy actions such as road pricing than for instance appraising projects expected to only influence a limited geographical area, and where the appraisal requires exact disaggregate estimations. This issue and how to proceed with this topic in a research perspective is discussed further in Section 7.5.

Secondly, Paper 2 contributes to a better understanding of the importance of passenger real-time information in public transport networks with a large number of alternatives. In line with previous studies such as Nuzzolo et al. (2001); Cats et al. (2011); Zargayouna et al. (2018), it is shown that providing sufficient real-time information to passengers is a useful means for lowering passenger delays, in turn, increasing passenger satisfaction in metropolitan public transport systems. Previous findings of Nielsen et al. (2009) that passenger delays are generally much larger than vehicle delays were confirmed in a large-scale multi-modal setting by the study, suggesting that there is a large potential in this regard. However, the paper also showed that there is a limit to how much passengers can benefit from retrieving real-time information more frequently – at least with the simple real-time predictions used in the study, see also Section 7.5.

Thirdly, the research on simulation of bicycle traffic has underlined the importance of acknowledging and modelling cyclists as a heterogeneous group. As such, although traffic on dedicated bicycle tracks has very large capacities (Paper 3), Paper 4 shows that if the infrastructure is not built to accommodate overtaking, even bicycle traffic may be subject to significant congestion. In order to ease overtaking even further, it may be beneficial – as it is already done on some bicycle tracks in Copenhagen - to mark a lane on the left hand side of infrastructure intended for overtaking. Having in mind how space efficient bicycle traffic is and how small a proportion of the road space it requires to provide excellent

facilities compared to car lanes, on-street parking, etc., wide bicycle tracks is an easy and affordable way to future-proof reliable and sustainable transport in otherwise congested urban environments. As such, it is highly recommended to ensure that cyclists have such high quality infrastructure available for more or less the entirety of their journeys.

Fourthly, Paper 5 underlines the importance of modelling multi-modality jointly when evaluating projects that are expected to influence the central part of cities where the share of cyclists is high. Bicycle traffic is shown to have serious impacts on the accessibility of car traffic in such areas why failing to include such inter-modal interactions biases traffic flows by overestimating urban car traffic. Such bias is problematic as it may cause policy makers to believe that there is no alternative to car traffic in central parts of the cities. The findings also show that taking additional measures towards a higher separation of car and bicycle traffic, e.g. by providing isolated bicycle tracks that shortcut through the city, would not only benefit cyclists, but would also eliminate the number of shared intersections, hence allowing car traffic to run smoother.

However, the largest policy implication has a more abstract nature. The thesis has developed a bicycle traffic assignment simulation model capable of modelling realistic travel times for bicycle traffic at a metropolitan scale. This means that for any projected intervention that may be suspected to influence the bicycle traffic, the impacts on bicycle traffic should no longer fly under the radar, but instead be evaluated and included in the appraisal of the project.

7.5 Future research

Directions for future research have already been outlined in the individual papers of the thesis. However, a few areas are worth highlighting.

As pointed out in Paper 1 and earlier in reviews by Wegener (2011); Bazzan and Klügl (2014), the stochastic and interdependent nature of such models is a challenge that need to be addressed thoroughly. There has already been some progress within the field since publication of Paper 1, as Guggisberg (2020) investigated between-run variability and related topics in MATSim (Horni et al., 2016). The study is yet another important step towards understanding the variability of MATSim (and agent-based traffic simulation models in general), and points towards using different approaches from Fourie et al. (2013); Flötteröd (2019) for reaching more consistent final states. Although such approaches might reduce variability, the outputs of stochastic simulations are still inherently stochastic for most scenarios. How to deal with such uncertainty in a way that allows uncertain results to be comprehended and interpreted by non-researchers is a matter that should receive much more attention than it has had so far. A more tangible but still relevant topic, is whether the effects of local infrastructure projects are able to shine through the variability when consistently evaluating the same project using different seeds.

In terms of passenger delay modelling it would be very interesting to investigate to which extent more advanced predictive real-time information methods such as Corman and Kecman (2018); Petersen et al. (2019) could reduce passengers delays even further. Although such have shown to predict vehicle arrivals more accurately, their effect on door-to-door passenger delays still needs to be evaluated. The passenger delay model formulated in Paper 2 would be an appropriate tool for doing so.

Even though this thesis has pioneered the topic of demand-sensitive bicycle traffic assignment models (Papers 3, 4, and 5), it has barely scratched the surface. In order to close the gap to the capabilities of equivalent models of car traffic, significant advances still need to be made. For instance including wider set of variables in the route choice part (e.g. gradient, land-use, and surface), exploring how desired speed is influenced by such variables, influence by weather, etc. There is also a need for further understanding of the underlying correlation and causation structures of trip length and desired speeds in order to be fully able to model the impact of large network improvements such as bicycle super highway projects.

Finally, the developed methodology is highly suited for evaluating the effects of increasing penetration rates of other types of bicycles such as e-bikes and speed pedelecs, as these can easily be included in the model. In this regard, it would be interesting to use the model to determine to which extent such can reach their full potential when used on congested bicycle networks, and whether they might be a space efficient and sustainable alternative to car traffic in urban areas.

References

Bazzan, A. L. C. and F. Klügl (2014). "A review on agent-based technology for traffic and transportation". In: *The Knowledge Engineering Review* 29.3, pp. 375–403. DOI: 10.1017/S0269888913000118. URL: https://www.cambridge.org/core/product/identifier/S0269888913000118/type/journal_article.

Cats, O., H. N. Koutsopoulos, W. Burghout, and T. Toledo (2011). "Effect of Real-Time Transit Information on Dynamic Path Choice of Passengers". In: *Transportation Research Record: Journal of the Transportation Research Board* 2217.1, pp. 46–54. DOI: 10.3141/2217-06. URL: http://journals.sagepub.com/doi/10.3141/2217-06.

Corman, F. and P. Kecman (2018). "Stochastic prediction of train delays in real-time using Bayesian networks". In: *Transportation Research Part C: Emerging Technologies* 95, pp. 599–615. DOI: 10.1016/j.trc.2018.08.003. URL: https://linkinghub.elsevier.com/retrieve/pii/S0968090X18311021.

Flötteröd, G. (2016). "MATSim as a Monte-Carlo Engine". In: *The Multi-Agent Transport Simulation MATSim*. Ed. by A. Horni, K. Nagel, and K. W. Axhausen. London: Ubiquity Press. Chap. 48, pp. 361–370.

Flötteröd, G. (2019). *A method for particle-based traffic assignment*. Unpublished work. Linköping University, VTI.

Fourie, P. J., J. Illenberger, and K. Nagel (2013). "Increased Convergence Rates in Multiagent Transport Simulations with Pseudosimulation". In: *Transportation Research Record: Journal of the Transportation Research Board* 2343.1, pp. 68–76. DOI: 10.3141/2343-09. URL: http://journals.sagepub.com/doi/10.3141/2343-09.

Guggisberg, D. (2020). *Convergence towards equilibrium in an agent-based transport simulation*. MSc Thesis. Institute for Transport Planning and Systems, ETH Zurich.

Horni, A., K. Nagel, and K. W. Axhausen, eds. (2016). *The Multi-Agent Transport Simulation MATSim*. London: Ubiquity Press. DOI: 10.5334/baw. URL: https://www.ubiquitypress.com/site/books/10.5334/baw/.

Kickhöfer, B., D. Hosse, K. Turner, and A. Tirachini (2016). *Creating an open MATSim scenario from open data: The case of Santiago de Chile*. VSP Working Paper 16-02. TU Berlin. URL: https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2016/16-02/.

Lawe, S., J. Lobb, A. Sadek, S. Huang, and C. Xie (2009). "TRANSIMS Implementation in Chittenden County, Vermont". In: *Transportation Research Record: Journal of the Transportation Research Board* 2132, pp. 113–121. DOI: 10.3141/2132-13. URL: http://trrjournalonline.trb.org/doi/10.3141/2132-13.

Lijesen, M. G. (2014). "Optimal Traveler Responses to Stochastic Delays in Public Transport". In: *Transportation Science* 48.2, pp. 256–264. DOI: 10.1287/trsc.2013.0465. URL: http://pubsonline.informs.org/doi/abs/10.1287/trsc.2013.0465.

Nagel, K. (1997). "Experiences with iterated traffic microsimulations in Dallas". In: *Traffic and granular flow '97*. Ed. by M. Schreckenberg and D. E. Wolf. Heidelberg: Springer, pp. 199–214. arXiv: 9712001v1 [adap-org]. URL: https://arxiv.org/pdf/adap-org/9712001.pdf.

Nielsen, O. A., O. Landex, and R. D. Frederiksen (2009). "Passenger delay models for rail networks". In: *Schedule-Based Modeling of Transportation Networks*. Boston, MA: Springer US. Chap. 2, pp. 27–49. DOI: 10.1007/978-0-387-84812-9_2. URL: http://link.springer.com/10.1007/978-0-387-84812-9_2.

Nuzzolo, A., F. Russo, and U. Crisalli (2001). "A Doubly Dynamic Schedule-based Assignment Model for Transit Networks". In: *Transportation Science* 35.3, pp. 268–285. DOI: 10.1287/trsc.35.3.268.10149. URL: http://pubsonline.informs.org/doi/abs/10.1287/trsc.35.3.268.10149.

Parbo, J., O. A. Nielsen, and C. G. Prato (2016). "Passenger Perspectives in Railway Timetabling: A Literature Review". In: *Transport Reviews* 36.4, pp. 500–526. DOI: 10.1080/01441647.2015.1113574. URL: https://www.tandfonline.com/doi/full/10.1080/01441647.2015.1113574.

Paulsen, M., T. K. Rasmussen, and O. A. Nielsen (2018). "Modelling Railway-Induced Passenger Delays in Multi-Modal Public Transport Networks". In: *Proceedings for the*

14th Conference on Advanced Systems in Public Transport and Transit Data 2018. Ed. by M. D. Hickman. Brisbane, Australia.

Petersen, N. C., F. Rodrigues, and F. C. Pereira (2019). "Multi-output bus travel time prediction with convolutional LSTM neural network". In: *Expert Systems with Applications* 120, pp. 426–435. DOI: 10.1016/j.eswa.2018.11.028. URL: https://linkinghub.elsevier.com/retrieve/pii/S0957417418307486.

Te Brömmelstroet, M., M. Skou Nicolaisen, B. Büttner, and A. Ferreira (2017). "Experiences with transportation models: An international survey of planning practices". In: *Transport Policy* 58.May 2016, pp. 10–18. DOI: 10.1016/j.tranpol.2017.04.007. URL: http://dx.doi.org/10.1016/j.tranpol.2017.04.007.

Twaddle, H., T. Schendzielorz, and O. Fakler (2014). "Bicycles in Urban Areas". In: *Transportation Research Record: Journal of the Transportation Research Board* 2434.1, pp. 140–146. DOI: 10.3141/2434-17. URL: http://journals.sagepub.com/doi/10.3141/2434-17.

Wegener, M. (2011). "From macro to micro - how much micro is too much?" In: *Transport Reviews* 31.2, pp. 161–177. DOI: 10.1080/01441647.2010.532883.

Zargayouna, M., A. Othman, G. Scemama, and B. Zeddini (2018). "Multiagent Simulation of Real-Time Passenger Information on Transit Networks". In: *IEEE Intelligent Transportation Systems Magazine* 12.2, pp. 50–63. DOI: 10.1109/MITS.2018.2879166. URL: https://ieeexplore.ieee.org/document/8532100/.

Ziems, S. E., B. Sana, J. Plotz, and R. M. Pendyala (2011). "Stochastic Variability in Microsimulation modeling Results and Convergence og Corridor-Level Characteristics". In: *Transportation Research Board 90th Annual Meeting*. Vol. 11-3560. URL: http://rampendyala.weebly.com/uploads/5/0/5/4/5054275/stochasticvariation_ziemssanaplotzpendyala_trb2011.pdf.

Advanced transport models capable of estimating traffic patterns and travel times are important tools when planning for sustainable mobility, especially in congested, urban contexts.

Through five papers, this thesis contributes to the field of large-scale agent-based traffic simulation by increasing the understanding of uncertainties of such models and developing new methodologies for modelling detailed individual behaviour in a computationally efficient fashion. Firstly, the inherent output variability caused by the dependence on pseudo-random numbers is analysed through repeated large-scale experiments. Secondly, a model for simulating adaptive passenger behaviour under presence of real-time information of public transport vehicle delays is formulated and implemented for a large-scale case study of Metropolitan Copenhagen. Finally, the thesis develops a methodology for realistic simulation of congested bicycle traffic based on speed heterogeneity, which is later implemented in a large-scale agent-based traffic assignment model allowing realistic simulation of bicycle as well as car traffic and their mutual interaction at intersections for the entire Metropolitan Copenhagen.

DTU Management
Department of Management, Technology and Economics
Technical University of Denmark

Akademivej
Building 358
DK-2800 Kongens Lyngby
Tel. +45 45 25 48 00

www.man.dtu.dk